



UNIVERSIDAD TÉCNICA NACIONAL

SEDE CENTRAL

INGENIERÍA DEL SOFTWARE

LICENCIATURA EN INGENIERÍA DEL SOFTWARE

ESTUDIO TÉCNICO DE LAS CARACTERÍSTICAS Y LA OFERTA DE SERVICIOS

TECNOLÓGICOS DIGITALES BASADOS EN TECNOLOGÍA DE *MACHINE*

***LEARNING* EN LA GRAN ÁREA METROPOLITANA DE COSTA RICA EN EL**

PERÍODO MAYO 2017 A JUNIO 2018.

MARCO ANTONIO ZÚÑIGA UGALDE

**ALAJUELA, COSTA RICA
MAYO, 2019**

DECLARACIÓN JURADA

Yo, Marco Antonio Zúñiga Ugalde, mayor, soltero, estudiante de la Carrera de Ingeniería del Software, de la Universidad Técnica Nacional, domiciliado en Alajuela, Monserrat, portador de la cédula de identidad número 2-0697-0733, en este acto, debidamente apercebido y entendido de las penas y consecuencias con las que se castiga, en el Código Penal, el delito de perjurio, ante quienes se constituyen en el Tribunal Examinador de mi trabajo de Tesis para optar por el título de licenciatura en Ingeniería del Software, juro solemnemente que mi trabajo de investigación "Estudio técnico de las características y la oferta de servicios tecnológicos digitales basados en tecnología de machine learning en la gran área metropolitana de Costa Rica en el período mayo 2017 a junio 2018" es una obra original que ha respetado todo lo preceptuado por las Leyes Penales así como la Ley de Derechos de Autor y Derechos Conexos, número 6683 de 14 de octubre de 1982 y sus reformas, publicada en La Gaceta número 226 de 25 de noviembre de 1982; incluyendo el numeral 70 de dicha ley que advierte: artículo 70º: Es permitido citar a un autor transcribiendo los pasajes pertinentes siempre que estos no sean tantos y seguidos, que puedan considerarse como una producción simulada y sustancial, que redunde en perjuicio del autor y de la obra original. Asimismo, estoy advertido que la Universidad Técnica Nacional se reserva el derecho de protocolizar este documento ante Notario Público. En fe de lo anterior firmo en la ciudad de Alajuela, el día 10 del mes de 7 del año 2019.



Marco Antonio Zúñiga Ugalde

Cédula de identidad: 2-0697-0733

DEDICATORIA

Dedico este trabajo a mis seres queridos más cercanos que son fuente de mi seguridad y confianza. Por desarrollar en mí un apoyo elemental y puro para terminar este trabajo de investigación.

Marco A. Zúñiga Ugalde

AGRADECIMIENTOS

La vida se encuentra llena de obstáculos y retos, que solo se batallan en el espacio mental, es una lucha que conlleva caídas y aprendizajes, es un reto en particular que nutre con conocimientos y experiencias que forman el profesionalismo y concierne la vida y el futuro.

Le agradezco a mis profesores por sus esfuerzos, por siempre forjar en mí un buen profesional y no rendirse.

Agradezco a mis padres por siempre ser fuente de apoyo para continuar con mi educación y enseñanzas de vida.

Gracias a la vida por este nuevo triunfo, gracias a todas las personas que me apoyaron y creyeron en la realización de esta tesis.

Marco A. Zúñiga Ugalde

Trabajo Final de Graduación

Acta No. 002

Acta de la sesión **No. 002**, del día Miércoles 19 de junio del 2019 a partir de las 18:00 horas , en periodo del segundo cuatrimestre, y en la que el Tribunal Evaluador recibe la sustentación del proyecto de graduación, realizado por el estudiante: **Marco Zúñiga Ugalde**, portador de la cédula: **206970733**, quién opta por el Grado Académico de Licenciatura en Ingeniería del Software, sita, en la Universidad Técnica Nacional, presentando el trabajo final de graduación con el tema:

Estudio técnico de las características y la oferta de servicios tecnológicos digitales basados en tecnología de *machine learning* en la gran área metropolitana de Costa Rica en el período mayo 2017 a junio 2018.

Preside el Tribunal la señora, directora de carrera de Ingeniería del Software, MGr. Ana Cecilia Odio Ugalde, junto con la participación del **MSc. Jairo Brenes Delgado**, tutor del trabajo final de graduación, **MAu. Luis Guillermo Alvarado Quesada y MGr. Joaquín Artavía Chaves**, lectores del trabajo final de graduación.

La señora presidente del Tribunal manifiesta que los miembros del mismo leyeron el informe, que el estudiante acogió las recomendaciones indicadas y en consecuencia se procede a recibir la sustentación correspondiente, en la que el estudiante realiza su exposición, sujeto al tiempo establecido. Terminada la misma, se procede a externar comentarios pertinentes al trabajo presentado, se formulan preguntas que fueron respondidas por parte del estudiante de manera exitosa.

Concluida la sustentación, el Tribunal, solicita a los presentes retirarse de la sala para proceder a la votación secreta. La votación da como resultado: Aprobado. Con nota de 9.7 nueve punto siete.

De nuevo en la sala, la señora presidente les comunica el resultado declarando que ya son: Licenciados en Ingeniería del Software, a la vez, indica que conforme a la normativa existente, deben revisar el Reglamento de Trabajos Finales de Graduación de la Universidad Técnica Nacional (disponible en la página web de la UTN), específicamente en el Capítulo IV, Artículos 39 y 40, donde se indica el

procedimiento a seguir para efectuar la entrega de los ejemplares físicos y digitales. Se les recuerda también la obligación de presentarse al ACTO DE GRADUACIÓN, al que serán convocados oportunamente. Se cierra la sesión a las: 19:30 horas del 19 de junio del presente.



Marco Zúñiga Ugalde
Estudiante



MSC. Jairo Brenes
Delgado
Miembro del Tribunal
Evaluador
Tutor



MAU. Luis Guillermo
Alvarado Quesada
Miembro del Tribunal
Evaluador
Lector



MSC. Joaquín Artavia
Chaves
Miembro del Tribunal
Evaluador
Lector



MGt. Ana Cecilia Odio
Ugalde
Directora de Carrera

Añexo III

**CARTA DE AUTORIZACIÓN PARA USO Y MANEJO DE LOS TRABAJOS
FINALES DE GRADUACIÓN UNIVERSIDAD TÉCNICA NACIONAL**

(Trabajo Individual)

Alajuela, 26/06/19.

Señores Vicerrectoría de Investigación

Sistema Integrado de Bibliotecas y Recursos Digitales

Estimados señores:

Yo Marco Antonio Zúñiga Ugalde portador (a) de la cédula de identidad número 2-0697-0733. En mi calidad de autor (a) del trabajo de graduación titulado: ESTUDIO TÉCNICO DE LAS CARACTERÍSTICAS Y LA OFERTA DE SERVICIOS TECNOLÓGICOS DIGITALES BASADOS EN TECNOLOGÍA DE MACHINE LEARNING EN LA GRAN ÁREA METROPOLITANA DE COSTA RICA EN EL PERÍODO MAYO 2017 A JUNIO 2018.

El cual se presenta bajo la modalidad de:

Proyecto de Graduación

Tesis de Graduación

Presentado en la fecha 26/06/19, autorizo a la Universidad Técnica Nacional, sede central, para que mi trabajo pueda ser manejado de la siguiente manera:

Autorizo	SI	No
Conservación de ejemplares para préstamo y consulta física en biblioteca	X	
Inclusión en el catálogo digital del SIBIREDI (Cita catalográfica)	X	
Comunicación y divulgación a través del Repositorio Institucional	X	
Resumen (Describe en forma breve el contenido del documento)	X	

Consulta electrónica con texto protegido	X	
Descarga electrónica del documento en texto completo protegido	X	
Inclusión en bases de datos y sitios web que se encuentren en convenio con la Universidad Técnica Nacional contando con las mismas condiciones y limitaciones aquí establecidas.	X	

Por otra parte declaro que el trabajo que aquí presento es de plena autoría, es un esfuerzo realizado de forma personal, académica e intelectual con plenos elementos de originalidad y creatividad. Garantizó que no contiene citas, ni transcripciones de forma indebida que puedan devenir en plagio, pues se ha utilizado la normativa vigente de la American Psychological Association (APA). Las citas y transcripciones utilizadas se realizan en el marco de respeto a las obras de terceros. La responsabilidad directa en el diseño y presentación son de competencia exclusiva, por tanto, eximo de toda responsabilidad a la Universidad Técnica Nacional. Consciente de que las autorizaciones no reprimen mis derechos patrimoniales como autor del trabajo. Confío en la que Universidad Técnica Nacional respete y haga respetar mis derechos de propiedad intelectual.

Firma del estudiante: _____

Cédula: 206970733

Día: 26/06/2019



Nota de validación del tutor

Por este medio, en calidad de tutor del documento "Análisis de la oferta del procesamiento de datos que utiliza machine learning en la gran área metropolitana de costa rica en el periodo 2017-2018 para la determinación de las características tecnológicas de software utilizadas", doy fe de que los ajustes y variaciones propuestas por los lectores, han sido tomados en cuenta en la consistencia y redacción técnica del documento final.

Por lo anterior es que doy mi anuencia para que el documento pase a la etapa final de la defensa, previa revisión filológica.

Sin más por el momento, suscribo

Jairo Brenes D.

MSc. Jairo Brenes Delgado

Cédula de identidad 1-1118-0915

Número de colegiado 2235

Cc.- Archivo

NOTA DE VALIDACIÓN DEL LECTOR

Por este medio, en calidad de lector del documento "**Estudio técnico de las características y la oferta de servicios tecnológicos digitales basados en tecnología de *machine learning* en la gran área metropolitana de Costa Rica en el período mayo 2017 a junio 2018**", realizado por el estudiante Marco Antonio Zúñiga Ugalde, cédula de identidad 2-0697-0733, doy fe de que los ajustes y variaciones propuestas por los lectores, han sido tomados en cuenta en la consistencia y redacción técnica del documento final.

Por lo anterior es que doy mi anuencia para que el documento pase a la etapa final de la defensa, previa revisión filológica.

Sin más por el momento, suscribo



MGT Joaquín Alberto Artavia Chaves

Cédula de identidad: 204900534

Número de colegiado: 1190

Cc.- Archivo

NOTA DE VALIDACIÓN DEL LECTOR

Por este medio, en calidad de lector del documento “**Estudio técnico de las características y la oferta de servicios tecnológicos digitales basados en tecnología de *machine learning* en la gran área metropolitana de Costa Rica en el período mayo 2017 a junio 2018**”, realizado por el estudiante Marco Antonio Zúñiga Ugalde, cédula de identidad 2-0697-0733, doy fe de que los ajustes y variaciones propuestas por los lectores, han sido tomados en cuenta en la consistencia y redacción técnica del documento final.

Por lo anterior es que doy mi anuencia para que el documento pase a la etapa final de la defensa, previa revisión filológica.

Sin más por el momento, suscribo

Luis Guillermo Alvarado Quesada

Cédula de identidad: 107160408

Número de colegiado: CPIC 875 / CPCECR 35474

Cc.- Archivo

15 de mayo, 2019

Señores
Universidad Técnica Nacional
Sede Central
Ingeniería del Software
Licenciatura en Ingeniería del Software

Estimados señores

Leí y corregí el trabajo final de graduación denominado "**Estudio técnico de las características y la oferta de servicios tecnológicos digitales basados en tecnología de machine learning en la gran área metropolitana de costa rica en el periodo mayo 2017 a junio 2018**", documento presentado por el estudiante, Marco Antonio Zúñiga Ugalde, para optar por el grado académico de Licenciatura en Ingeniería del Software.

Corregí el trabajo en aspectos tales como: concordancia y construcción de párrafos, vicios del lenguaje que se trasladan a lo escrito, ortografía, puntuación y otros aspectos relacionados con el campo filológico y desde este punto de vista considero que este documento, está listo para ser presentado, por cuanto cumple con los requisitos establecidos.

Cordialmente,



Licda. Victoria Eugenia Gaitán Sibaja
Cédula 2-288-1488
Número de colegiada 2779
Asociación Costarricense de Filólogos
Carné 0015ACFIL

TABLA DE CONTENIDOS

1 Capítulo I Introducción y delimitación.....	1
1.1 Justificación.....	2
1.2 Estado del arte	6
1.3 Objetivo General	17
1.4 Objetivos Específicos	17
1.5 Problema.....	18
1.6 Alcance	20
1.7 Limitaciones	21
1.8 Matriz de congruencia	22
2 Capítulo II Marco Teórico.....	2
2.1 La inteligencia artificial y su impacto en la vida moderna.....	28
2.2 <i>Machine learning</i> , tecnología del futuro	30
2.3 Situación del mercado de servicios digitales relacionados con <i>machine learning</i> en Costa Rica.....	31
2.4 Servicios tecnológicos digitales.	38
2.4.1 Sistemas digitales	39

2.5 Empresas proveedoras de servicios digitales.....	40
2.6 Características de <i>machine learning</i>	40
2.6.1 Framework de Machine learning	41
2.6.2 Estilos de machine learning.	76
2.6.3 Tipos de enfoques de <i>machine learning</i>	81
2.6.4 Métodos de machine learning.	84
2.6.5 Casos de uso específico de <i>machine learning</i>	95
2.6.6 Retos de implementación de <i>machine learning</i>	102
2.6.7 Tipos de consumo de <i>machine learning</i>	103
2.7 Características organizativas de <i>machine learning</i>	104
2.7.1 Necesidades principales para desarrollar <i>machine learning</i>	105
2.7.2 Proyección del desarrollo tecnológico de <i>machine learning</i> en Costa Rica.	105
2.7.3 Recurso humano.....	107
2.7.4 Capacitación de personal.....	109
2.7.5 Áreas económicas de vinculación impactadas con el machine learning.	114
3 Capítulo III Marco Metodológico.....	30
3.1 Tipo de investigación.....	118

3.2 Enfoque de investigación	120
3.3 Población	121
3.4 Fuentes de información	126
3.5 Sujetos de información	127
3.6 Planeación para el proceso de recolección de información.	128
3.6.1 Planteamiento de objetivos de la visita.	128
3.6.2 Población Objetivo	128
3.6.3 Establecimiento del método de medición.	128
3.6.4 Instrumento de medición.....	129
3.6.5 Preparación de los investigadores de campo.....	131
3.6.6 Prueba piloto.....	133
3.6.7 Organización y análisis de la información recolectada.	134
3.7 Temporalidad	134
3.8 Matriz metodológica	135
4 Capítulo IV Análisis situacional	120
4.1 Oferta de servicios tecnológicos digitales basados en machine learning, área metropolitana de CR.	144

4.1.1 Resumen y análisis de la información recolectada.....	144
4.2 Informe sobre la relevancia entre las características técnicas, funcionales y organizativa respecto de las áreas económicas.	172
4.3 Análisis de la academia.....	175
4.3.1 Universidades estatales de Costa Rica.....	176
4.3.2 Centros privados de enseñanza en correlación con <i>machine learning</i>	188
5 Capítulo V Conclusiones y Recomendaciones.....	147
5.1 Conclusiones.....	192
5.1.1 Definición de la oferta de servicios tecnológicos con <i>machine learning</i> e identificación de las empresas.	192
5.1.2 Características técnicas.....	193
5.1.3 Características funcionales.....	194
5.1.4 Características organizativas.....	194
5.1.5 Academia.....	197
5.2 Recomendaciones.....	198
5.2.1 Definición de la oferta de servicios tecnológicos con <i>machine learning</i> e identificación de las empresas.	198
5.2.2 Características técnicas.....	198

5.2.3 Características funcionales	199
5.2.4 Características organizativas	200
5.2.5 Academia.....	200
6 Capítulo VI Guía introductoria al <i>machine learning</i> en empresas	197
6.1 Fase exploratoria.....	205
6.1.1 Determinación de los objetivos de negocio.	205
6.1.2 Análisis de la competencia, una buena estrategia.....	206
6.1.3 Recomendaciones sobre infraestructura para el desarrollo de <i>machine learning</i>	207
6.1.4 Recomendaciones sobre recurso humano por contratar.	208
6.2 Fase de diseño, experimentación e implementación.	209
6.2.1 Etapas base para desarrollar <i>machine learning</i>	209
7 Referencias Bibliográficas.....	214
8 Apéndice.....	237
9 Anexo	257
10 CRONOGRAMA	265

ÍNDICE DE CUADROS

Tabla 1 Algunas definiciones de inteligencia artificial, organizadas en cuatro categorías.....	7
Tabla 2 Matriz de congruencia.....	26
Tabla 3 Beneficios de Azure Machine Learning Studio.....	64
Tabla 4 Beneficios de Watson Machine Learning.....	73
Tabla 5 Definición de NLP, NLU y NLG.....	100
Tabla 6 Cursos en línea para capacitación sobre machine learning.....	113
Tabla 7 Matrícula para el Master en Ciencias de Machine learning de la Universidad de Carnegie Mellon.....	114
Tabla 8 Definición de la población con indicios de servicios basados en machine learning.....	126
Tabla 9: Objetivo 1 de la Matriz Metodológica.....	136
Tabla 10 Objetivo 2 de la Matriz Metodológica.....	137
Tabla 11 Objetivo 3 de la Matriz Metodológica.....	140
Tabla 12 Objetivo 4 de la Matriz Metodológica.....	142
Tabla 13 Cantidad de empresas encuestadas.....	148

Tabla 14 Respuesta de Neustar con impacto en el área económica de mercadeo o marketing digital.....	173
Tabla 15 Respuesta de la empresa Singularities y su impacto en las áreas económicas como la financiera y logística.	174
Tabla 16 Cursos optativos Escuela de Ingeniería Eléctrica, UCR.	177
Tabla 17 Laboratorios del Programa de Posgrado en Ingeniería Eléctrica (PPIE).	180
Tabla 18 Cursos del énfasis de ciencias de la computación, UCR.	183
Tabla 19 Cursos optativos en planes de Bachillerato y Licenciatura con referencia a machine learning.	184
Tabla 20 Cursos optativos en planes de Bachillerato y Licenciatura con referencia a machine learning.	185
Tabla 21 El plan de Bachillerato en ingeniería en computación, relevancia de materia con machine learning.	186
Tabla 22 Respuesta de Neustar con impacto en el área económica de mercadeo o marketing digital.....	211

ÍNDICE DE GRÁFICOS

Gráfico 1 Gráfico de participación de la encuesta.	148
Gráfico 2 Frameworks más utilizados por las empresas participantes	150
Gráfico 3 Métodos de machine learning más utilizados	153
Gráfico 4 Estilos de machine learning	155
Gráfico 5 Tipos de enfoque de machine learning	156
Gráfico 6 Áreas económicas impactadas por el machine learning	157
Gráfico 7 Casos de uso específicos de machine learning.	159
Gráfico 8 Resumen por categoría de casos de uso más utilizados	160
Gráfico 9 Caso de uso en que se utiliza Machine learning por nicho de negocio.	161
Gráfico 10 Retos en la implementación de productos o servicios con machine learning.	163
Gráfico 11 Tipo de consumo de machine learning	166
Gráfico 12 Tipos de profesiones contratados para desarrollo de Machine learning.....	169
Gráfico 13 Métodos de capacitación utilizados por el recurso humano encargado del desarrollo de machine learning en las empresas	171

ÍNDICE DE FIGURAS

Figura 1. Naturaleza multidisciplinaria del machine learning	9
Figura 2 Machine learning tiene un amplio potencial en todas las industrias y casos de uso, McKinsey Global Institute analysis.	12
Figura 3. Firmas locales brindan servicios de inteligencia artificial, El Financiero Costa Rica, 2017.	19
Figura 4. Inteligencia artificial en América Latina, Statista.....	29
Figura 5 Áreas de fortaleza, alertas y temas críticos. Estado de Nacional 2014.	33
Figura 6 La industria tecnológica en Costa Rica.	35
Figura 7. Servicios y planes de firmas informáticas 2017.....	37
Figura 8 Como funciona un framework y su interpretación visual.....	42
Figura 9 Concepto de flujo de uso de librerías.	44
Figura 10 Ejemplo de funcionamiento de TensorFlow con DeepLearning.....	49
Figura 11 Conjunto de librerías que compone Apache Spark.	50
Figura 12 Beneficios de Amazon Machine Learning	55
Figura 13 Ejemplo de flujo utilizando Amazon Machine Learning	56
Figura 14 Frameworks que se pueden optimizar, Amazon SageMaker.	58
Figura 15 Capacidad de optimización de TensorFlow con Amazon SageMaker.	59

Figura 16 Interfaz de la plataforma Azure Machine Learning Studio	60
Figura 17 Flujo de trabajo utilizando Azure Machine Learning Studio.....	61
Figura 18 Beneficios de Microsoft Azure Machine Learning Services.....	65
Figura 19 Flujo de trabajo en Azure Machine Learning Services.....	66
Figura 20 Productos que forman parte de Cloud AI de Google.....	67
Figura 21 Paquetes del Cloud AutoML.....	68
Figura 22 Detección de imágenes con AutoML Vision y su interfaz.....	69
Figura 23 AutoML Natural Language y el análisis de texto.....	70
Figura 24 Flujo de trabajo de alto nivel de Google Cloud Machine Learning.....	71
Figura 25 Frameworks más utilizados según Google Cloud para los servicios de machine learning.....	72
Figura 26 Interfaz de IBM Watson Machine Learning.....	74
Figura 27 Técnicas de Machine Learning en Watson Machine Learning.....	75
Figura 28 Comparación de frameworks basados en la nube.....	76
Figura 29 Diagrama de flujo del estilo supervisado.....	78
Figura 30 Diagrama de flujo del estilo no supervisado.....	79
Figura 31 Diagrama de flujo del estilo por refuerzo.....	81

Figura 32 Diferencia entre el enfoque clasificadorio y regresivo en la ciencia médica...	84
Figura 33 Ilustración abstracta de cómo funciona un algoritmo de regresión.	86
Figura 34 Representación del funciona miento de un algoritmo basado en instancias.	88
Figura 35 Representación de un algoritmo de árbol de decisiones.	89
Figura 36 Fórmula del teorema de Bayes.	91
Figura 37 Representación del funcionamiento del algoritmo de agrupamiento.....	92
Figura 38 Redes neuronales organizadas en capas que constan de una serie de nodos interconectados. Las redes pueden tener decenas o cientos de capas ocultas. ...	94
Figura 39 La mezcla de técnicas pueden resolver un caso de uso.	97
Figura 40 Diagrama de arquitectura de chatbots.	101
Figura 41 Estudio de ServiceNow sobre los cambios organizaciones para adoptar el machine learning.	109
Figura 42 Dominios de disrupción del machine learning.	116
Figura 43 Frameworks más utilizados según GitHub, StackOverflow y búsquedas de Google.	152
Figura 44 Top 10 Métodos de Machine learning más usados por Científicos de Datos, 2017	154
Figura 45. Equipos de investigación Pris-LAB EIE UCR.	181

Figura 46. Equipos de investigación Pris-LAB EIE UCR. (UCR E. d., 2018) 182

ÍNDICE DE APÉNDICE

Apéndice A Respuesta de CAMTIC sobre la investigación.	237
Apéndice B Respuesta CINDE sobre la investigación.....	238
Apéndice C Respuesta del CONICIT sobre la investigación.	239
Apéndice D Respuesta MICITT sobre la investigación.....	240
Apéndice E Respuesta de la empresa CRUX de la encuesta.	243
Apéndice F Respuesta empresa Grupo Prides sobre la encuesta.	243
Apéndice G Respuesta empresa Logosoft sobre la encuesta.	244
Apéndice H Respuesta de la empresa MediaLAB sobre la encuesta.	247
Apéndice I Respuesta de la empresa SimMachines sobre la encuesta.	247
Apéndice J Respuesta de la empresa Xchematic sobre la encuesta.	248
Apéndice K Volante informativo de tesis.	249
Apéndice L Cuestionario para la recolección de información.	251
Apéndice M Formulario de Google con las preguntas del cuestionario.	254
Apéndice N Machote de Carta de confidencialidad.	255
Apéndice O Diagrama de flujo del método de medición.	256

ÍNDICE DE ANEXO

Anexo 1 Visión por computadora. Plan de estudio Ingeniería Eléctrica, UCR.	257
Anexo 2 Inteligencia Artificial, Ciencias de la computación e informática, UCR.	258
Anexo 3 Bloque optativo en los planes de Bachillerato y Licenciatura, Computación e informática, UCR	259
Anexo 4 Bloque optativo en los planes de Bachillerato y Licenciatura, Computación e informática, UCR	260
Anexo 5 Inteligencia artificial, Ingeniería en sistemas de información, UNA.....	261
Anexo 6 Minería de datos 1 y 2, Ingeniería en sistemas de información, UNA	262
Anexo 7 Inteligencia artificial, Bachillerato Ingeniería en Computación, TEC.....	263
Anexo 8 Mapa de algoritmos de machine learning.....	264

GLOSARIO DE TÉRMINOS

AI

La inteligencia artificial (IA), también llamada inteligencia computacional, es la inteligencia exhibida por máquinas, 7

Amazon S3

Sistema de almacenamiento de datos en la nube por Amazon Web Services, 56

AWS

Amazon Web Services, 57

Big Data

Se refiere a conjuntos de datos o combinaciones de conjuntos de datos cuyo tamaño (volumen), complejidad (variabilidad) y velocidad de crecimiento (velocidad) dificultan su captura, gestión, procesamiento o análisis mediante tecnologías y herramientas convencionales, tales como bases de datos relacionales y estadísticas convencionales o paquetes de visualización, dentro del tiempo necesario para que sean útiles., 95

Bot

Un bot (aféresis de robot) es un programa informático que efectúa automáticamente tareas., 101

Bsc

Bachelor of Science es un título de grado en ciencias muy extendido en los países anglófonos o anglosajones., 15

Business Intelligence

Inteligencia de negocios, 95

call centers

Centros de servicios al cliente, 194

CAMTIC

Cámara de Tecnologías de Información y Comunicación, 2

CCSS

Caja Costarricense de Seguro Social, 35

CEO

Chief Executive Officer, tal reza su denominación en idioma inglés o Director

Ejecutivo, Consejero Delegado, Presidente Ejecutivo o Principal Oficial Ejecutivo,

98

chatbot

Un chatbot es simplemente un software que te permite comunicarte con tus clientes

dentro de las aplicaciones de mensajería., 98

científicos de datos

Profesional formado en las ciencias matemáticas y las estadísticas que domina la programación y sus diferentes lenguajes, ciencias de la computación y analítica., 169

CONICIT

Consejo Nacional para Investigaciones Científicas y Tecnológicas, 2

Facebook

Red social en internet., 122

Formulario de Google

Formularios en línea que permiten la recolección de información por medio de la plataforma de Google., 131

Frameworks

En el desarrollo de software, un entorno de trabajo es una estructura conceptual y tecnológica de asistencia definida, normalmente, con artefactos o módulos concretos de software, que puede servir de base para la organización y desarrollo de software., 5

IBM

International Business Machines Corporation es una reconocida empresa multinacional estadounidense de tecnología y consultoría con sede en Armonk, Nueva York., 72

KDNuggets

Es un sitio que se centra en lo importante e interesante en IA, Análisis, Big Data, Minería de datos, Ciencia de datos y Aprendizaje automático., 154

M. Sc.

Master of Science es un título de grado, 15

machine learning

El aprendizaje automático o aprendizaje de máquinas (del inglés, "*Machine Learning*") subcampo de las ciencias de la computación y una rama de la inteligencia artificial, cuyo objetivo es desarrollar técnicas que permitan a las computadoras aprendan., 2

MathWorks

es una corporación privada estadounidense que se especializa en software de computación matemática., 154

The MathWorks, Inc. es una corporación privada estadounidense que se especializa en software de computación matemática., 93

MATLAB

es una herramienta de software matemático que ofrece un entorno de desarrollo integrado con un lenguaje de programación propio. Está disponible para las plataformas Unix, Windows, Mac OS X y GNU/Linux, 153

MBIT School

MBIT School es una escuela especializada en Big Data, Data Science e Inteligencia Artificial, 95

MICITT

Ministerio de Ciencia Tecnología y Telecomunicaciones de Costa Rica, 2

NLG

Natural Language Generation (Generación de lenguaje natural), 99

NLU

Natural Language Understanding(Entendimiento de lenguaje natural), 98

Phyton

Se trata de un lenguaje de programación multiparadigma, ya que soporta orientación a objetos, programación imperativa y, en menor medida, programación funcional., 65

PNL

El procesamiento del lenguaje natural, 98

Promidat

Programa Iberoamericano de Formación en Minería de Datos, 188

redes neuronales de *deep learning*

redes neuronales de deep learning, 94

Simulink

es un entorno de programación visual, que funciona sobre el entorno de programación Matlab, 153

startups

Startup es una gran empresa en su etapa temprana a diferencia de una Pyme, la Startup se basa en un negocio que será escalable más rápida y fácilmente, haciendo uso de tecnologías digitales., 169

TEC

Tecnológico de Costa Rica, 15

UCR

Universidad de Costa Rica, 15

UNA

Universidad Nacional de Costa Rica, 15

RESUMEN EJECUTIVO

El objetivo de esta investigación es elaborar un estudio técnico de la oferta de servicios tecnológicos digitales con *machine learning* por medio de las empresas proveedoras que permita el análisis de sus características principales en la Gran Área Metropolitana de Costa Rica, durante el periodo de mayo 2017 a junio 2018.

El capítulo 1. Incluye los objetivos generales y específicos de la investigación, estado del arte, la justificación, los alcances y las limitaciones que conlleva.

El capítulo 2. Presenta el marco teórico, este contiene información referente con el tema de estudio.

El capítulo 3. Expone el proceso metodológico llevado a cabo en la investigación, explica también las variables de investigación, tipo de investigación, enfoque investigativo, sujetos de información y los instrumentos que se aplican en la recolección de la información.

En el capítulo 4. Se analiza y se clasifica la información recolectada a partir de la aplicación de los instrumentos de recopilación de la información diseñados.

Finalmente, el capítulo 5. Presenta las conclusiones y recomendaciones que se obtienen de la investigación, donde se concluye que según las características técnicas, funcionales y organizativas del *machine learning*, el estado de desarrollo de esta tecnología se encuentra en una fase inicial y de crecimiento, sin llegar a su completo desarrollo, según las características encontradas.

INCIDE DE PALABRAS CLAVES

- Machine learning.
- Inteligencia artificial.
- Servicios tecnológicos digitales.
- Estudio técnico.
- Estado de desarrollo tecnológico.
- Características técnicas, funcionales y organizativas.
- Guía introductoria al *machine learning* en empresas.
- Análisis de la academia.

Capítulo I

Introducción y delimitación

1.1 Justificación

Para tener un concepto de *machine learning* se cita al Dr. Román Yampolskiy, Universidad de Louisville (2018): “*Machine learning* es la ciencia de hacer que las computadoras aprendan tan bien como los humanos o mejor”.

Esta capacidad de aprendizaje y generación de conocimiento de manera artificial está permitiendo grandes avances tecnológicos, es un campo importante que investigar.

Esta investigación se realiza debido a la inexistencia de investigaciones sobre el estado de desarrollo de la tecnología *machine learning* en las empresas en la Gran Área Metropolitana de Costa Rica. Es decir, un primer paso para analizar las características del *machine learning* que es desarrollado por empresas en el país. Esto se justifica con las consultas a CAMTIC, CINDE, CONICIT y MICITT que se encargan de la investigación, desarrollo tecnológico y promotoras de negocios.(Apéndice A, Apéndice B, Apéndice C y Apéndice D.)

La investigación hace referencia del análisis de la oferta de los servicios de *machine learning* de las empresas proveedoras en Costa Rica durante el periodo mayo 2017 a junio 2018. Dentro de la investigación se tiene como objetivo la identificación y el análisis de las principales características tecnológicas de *machine learning* . Además, permite compartir información oportuna y relevante sobre empresas de tecnologías digitales localizadas en el territorio costarricense que se dedican a este esfuerzo económico y social.

Estas empresas satisfacen una demanda tecnológica digital por medio de la oferta de un producto o servicio que forma parte de la cadena de producción o de consumo de los clientes. Es por ello por lo que se deben analizar puntualmente las características de la tecnología *machine learning* desde la producción y la oferta, es decir desde su punto de desarrollo del conocimiento.

Se cita el concepto de oferta en general, según Kotler, Armstrong, Cámara y Cruz, autores del libro "Marketing" (2012), plantean la siguiente definición:

“Las necesidades y los deseos de los consumidores se satisfacen con las ofertas de mercado, es decir, cierta combinación de productos, servicios, información y experiencias que se ofrece a un mercado para satisfacer una necesidad o un deseo. Las ofertas de mercado no se limitan a productos físicos. También incluyen servicios, actividades o beneficios que se ofrecen en venta, que son básicamente intangibles y que no derivan en la posesión de algo”. (pág. 6)

Es importante conocer sobre la oferta de servicios tecnológicos digitales con *machine learning*, porque determina el avance tecnológico del país y la ventaja competitiva que genera la utilización de esta tecnología en productos y servicios de gran diversidad.

La investigación se enfoca en un sector de gran crecimiento en el nivel mundial en los últimos años y constituye uno de los principales avances y

novedades, con lo cual es posible conocer las diferentes dificultades y obstáculos que enfrentan las empresas proveedoras de servicios tecnológicos digitales con *machine learning*.

Con ello, esta pretende encontrar cuáles son las principales características del *machine learning* que se encuentran en la oferta de servicios tecnológicos digitales de las empresas identificadas. Así, los resultados se apoyan en técnicas de investigación adecuadas, además, que la información suministrada para su desarrollo es de carácter certero, verídico y confiable.

También, permite al lector y entes interesados la determinación de los servicios y de los sectores económicos impactados por esta tecnología en Costa Rica.

Es importante entender los antecedentes y el por qué esta tecnología llegó a un punto donde es factible su utilización, se destaca, por ejemplo, la cantidad de datos digitales de sistemas computacionales almacenados en bases de datos que se generan actualmente en las empresas se incrementa en forma exponencial, por lo que su extracción se torna en una ventaja competitiva que no se puede menospreciar.

Es una oportunidad a la que se le debe prestar especial atención, pues esta extracción es usada como oferta de un servicio que genera una mayor utilidad, en comparación a tener simplemente estos datos almacenados.

Con cantidades inmensas de datos es imposible el procesamiento por una sola persona, de manera que pueda emitir conclusiones y menos todavía, hacer predicciones. Con el apoyo de diversos algoritmos, se pueden detectar patrones de comportamiento, contando con las variables que se le proporcionan.

Hoy, la evolución de la tecnología ha logrado mayores capacidades de cómputo que permiten utilizar algoritmos de predicción de modelos para procesar los grandes volúmenes de información, situación que puede ser aprovechada mediante tecnologías como la descrita en este documento.

Específicamente, con esta investigación se pretende evidenciar el estado de desarrollo de la tecnología de *machine learning* por medio de las principales características técnicas y otras características que tiene relevancia con *machine learning* como los *Frameworks*¹, métodos, estilos, tipos de enfoques de *machine learning*, áreas económicas impactadas, casos de uso específicos, ventajas, retos a la hora de implementar *machine learning*, tipo de consumo local o exterior, necesidades o retos al desarrollar proyectos con *machine learning*, proyección a

¹ Framework: Un marco de software es una plataforma concreta o conceptual donde el código común con funcionalidad genérica puede ser selectivamente especializado o anulado por desarrolladores o usuarios. Janssen, C. (n.d.) Software Framework. <http://www.techopedia.com/definition/14384/software-framework>

futuro del desarrollo de *machine learning*, según los entrevistados y especialistas del tema, recurso humano como generador de conocimiento y profesionales y capacitación en *machine learning*.

1.2 Estado del arte

Para entender que es *machine learning*, se toma la referencia del Dr. Danko Nikolic, CSC y *Max-Planck Institute* (2018): "El *machine learning* es la ciencia que muestra cómo las computadoras actúan sin ser programadas explícitamente, sino que les permite aprender algunos trucos por su cuenta".

Machine learning es un subcampo de la inteligencia artificial, por ello se toma el libro "Artificial Intelligence a modern approach tercera edición" por Stuart Russell y Peter Norvig (2010, pág. 2) en el cual explican ampliamente con diferentes definiciones ya sea racional y humana ver Tabla 1. Explica que existen diferencias entre lo que es pensar y actuar humanamente de maneras racionales o humanas citando varias referencias para tener diferentes perspectivas por lo inmenso que es el concepto de inteligencia artificial.

<p>Pensando humanamente</p> <p>"El nuevo y emocionante esfuerzo para hacer que las computadoras piensen. . . máquinas con mentes, en sentido pleno y literal. "(Haugeland, 1985).</p> <p>"[La automatización de] actividades que asociamos con el pensamiento humano, actividades como la toma de decisiones, la resolución de problemas, el aprendizaje. . . "(Bellman, 1978).</p>	<p>Pensando racionalmente</p> <p>"El estudio de las facultades mentales por medio del uso de modelos computacionales".</p> <p>(Charniak y McDermott, 1985).</p> <p>"El estudio de los cálculos que hacen posible percibir, razonar y actuar".</p> <p>(Winston, 1992).</p>
<p>Actuando humanamente</p> <p>"El arte de crear máquinas que realizan funciones que requieren inteligencia cuando las realiza gente" (Kurzweil, 1990).</p> <p>"El estudio de cómo hacer que las computadoras hagan cosas en las que, por el momento, la gente es mejor" (Rich and Knight, 1991).</p>	<p>Actuando racionalmente</p> <p>"La Inteligencia Computacional es el estudio del diseño de agentes inteligentes." (Poole et al., 1998).</p> <p>"Inteligencia artificial. . . se preocupa por el comportamiento inteligente" en artefactos. "(Nilsson, 1998).</p>

Tabla 1 Algunas definiciones de inteligencia artificial, organizadas en cuatro categorías.

Fuente: *Artificial Intelligence a modern approach tercera edición* por Stuart Russell y Peter Norvig (2010, pág. 2).

La inteligencia artificial es un concepto que engloba al *machine learning*, pues es la forma cómo un ser artificial (máquina o sistema computacional) tiene la capacidad de adquirir conocimiento de manera racional o humana y esta pueda ser utilizada en escenario similares.

Para tener un conocimiento más amplio se citan los siguientes ejemplos de casos de uso: desde conversaciones de voz en texto, etiquetado de fotografías, detección de fraude bancario, medicina de alta precisión, predicción de lesiones, autos autónomos. Se encuentra en métodos informáticos como el análisis avanzado de datos, el procesamiento de lenguaje natural y *machine learning*.

Diane Bryant vicepresidenta ejecutiva y directora general del *Data Center Group* de Intel explica en este artículo publicado en el sitio *NetworkWorld* (2016) que la inteligencia artificial: "...está transformando la forma de trabajar de las empresas, así como nuestra manera de interactuar con el mundo".

Para introducir el concepto de *machine learning* se toma la interpretación de Tom Mitchell (1997) quien en su libro *Machine learning* dice: "El campo del *machine learning* se refiere a la cuestión de cómo construir programas informáticos que mejoren automáticamente con la experiencia." (pág. Prefacio)

En el artículo académico SAS313-2014, "*An Overview of Machine learning with SAS® Enterprise Miner™*" por Patrick Hall, Jared Dean, Ilknur Kaynar Kabul, y Jorge Silva del instituto SAS (2014) mencionan que el campo multidisciplinario que abarca *machine learning*: "ilustra la naturaleza multidisciplinaria de la minería de datos, la ciencia de datos, *machine learning* y áreas relacionadas". (pág. 3)

Esta Figura 1 fue creada originalmente por SAS en un manual de investigación sobre minería de datos de 1998, y el nuevo campo de la ciencia de datos se agregó al artículo académico SAS313-2014.

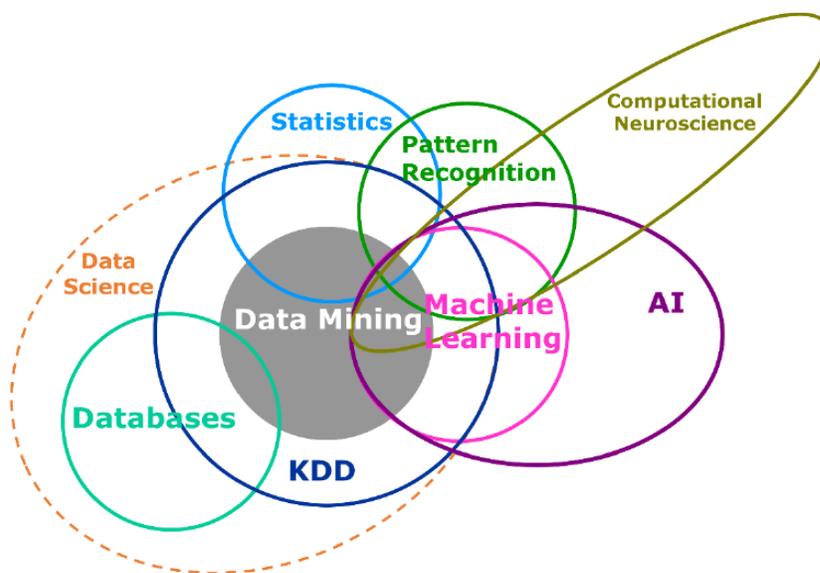


Figura 1. Naturaleza multidisciplinaria del machine learning

Fuente: Artículo académico SAS313-2014 (SAS 2014).

El autor *McKinsey*, en su informe *The Age of Analytics* (2016), menciona que:

“La convergencia de varias tendencias tecnológicas está acelerando el progreso. El volumen de datos continúa duplicándose cada tres años a medida que la información fluye desde plataformas digitales, sensores inalámbricos y miles de millones de teléfonos móviles. La capacidad de almacenamiento de datos ha aumentado, mientras que su costo se ha desplomado. Los científicos de datos ahora tienen una potencia informática sin precedentes a su disposición, y están diseñando algoritmos cada vez más sofisticados. (p. 1)

Para ampliar esta cita y aclarar por qué son tan importante los datos que se producen, es importante referirse a que las capacidades de los datos y el análisis han tenido gran impulso en los últimos años. El volumen de datos disponibles ha aumentado exponencialmente, se han desarrollado algoritmos más efectivos, y el poder computacional y el del almacenamiento han mejorado constantemente.

La convergencia de estas tendencias está alimentando el avance tecnológico y la ruptura en los negocios. Los datos son ahora un activo corporativo importante, vienen de la web, porque hay miles de millones de teléfonos, de sensores, de sistemas de pago, cámaras y una enorme variedad de otras fuentes, y su valor está ligado a su uso final. Mientras que los datos se convertirán progresivamente en productos de alta demanda.

En cuanto a la demanda de profesionales con especialidad en *machine learning*, como lo son los científicos de datos, son altamente demandados y sobrepasan la oferta, por lo que es imprescindible la capacitación de profesionales en el área y así lo justifica el informe de McKinsey (2016), quien menciona:

“Los científicos de los datos, en particular, están en alta demanda. Nuestro informe de 2011 planteó la hipótesis de que la demanda de científicos de datos sobrepasaría la oferta. Esto es lo que vemos actualmente en el mercado de trabajo, a pesar de que las universidades están agregando programas de datos y análisis y que otros tipos de programas de capacitación están proliferando. Los salarios medios de los científicos de datos en los

Estados Unidos aumentaron aproximadamente un 16 por ciento al año de 2012 a 2014. (p. 4)

En la Figura 2 de *McKinsey Global Institute, The Age of Analytics: Competing in data-driven world*, (2016) explica que los usos específicos de la industria combinan la riqueza de datos en conjunto con la mayor oportunidad, las burbujas más grandes en el cuadrante superior derecho. Estas representan áreas donde las organizaciones deben priorizar el uso de *machine learning* y prepararse para una transformación que toma lugar.

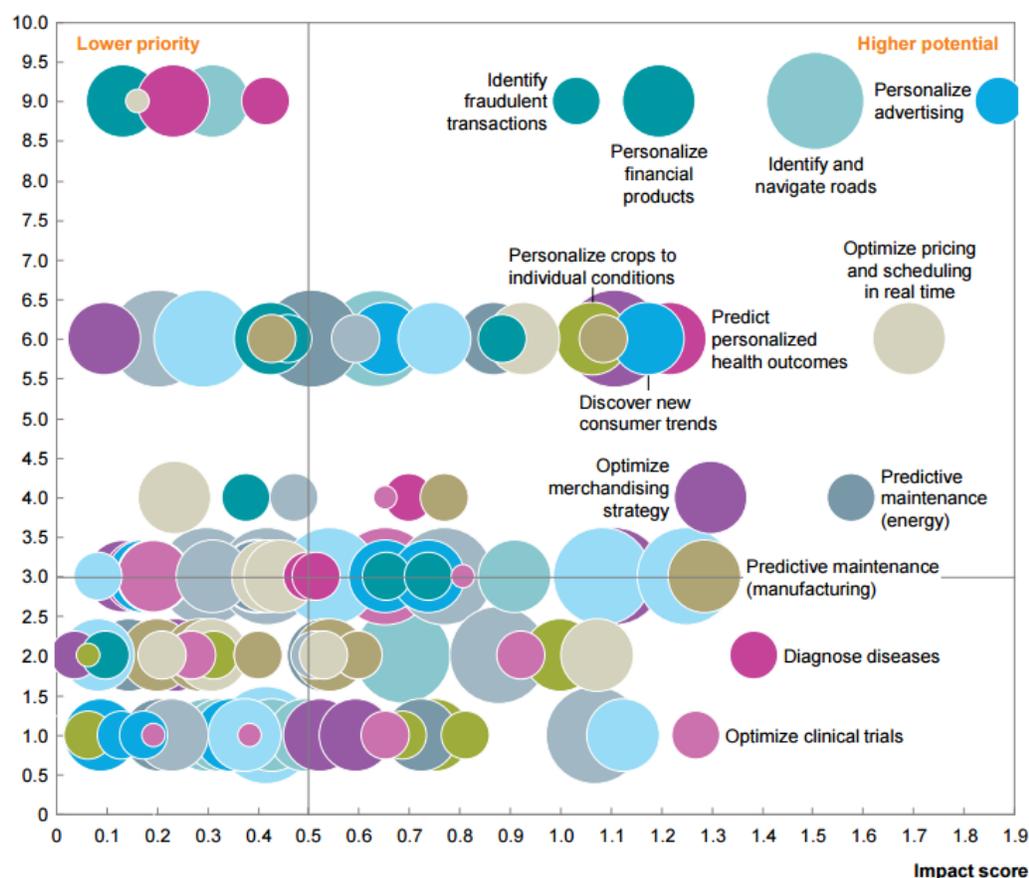
Algunos de los casos de mayor oportunidad incluyen la publicidad personalizada, los vehículos autónomos, la optimización de precios, el enrutamiento y la programación basados en datos en tiempo real de viajes y logística, el pronóstico de resultados de salud personalizados y la optimización de las estrategias de comercialización en el comercio minorista.

Machine learning has broad potential across industries and use cases



Volume

Breadth and frequency of data



SOURCE: McKinsey Global Institute analysis

Figura 2 Machine learning tiene un amplio potencial en todas las industrias y casos de uso, McKinsey Global Institute analysis.

Fuente: McKinsey Global Institute, *The Age of Analytics: Competing in data-driven world*, (2016).

Según, SynergicPartners una firma especializada en *Big data*, *Data Science* y *Data Engineering*, (2016) menciona: “Microsoft crea el “*Distributed Machine learning Toolkit*”, que permite la distribución eficiente de problemas de *machine*

learning en múltiples computadores”. Esto indica que compañías grandes en el nivel mundial ya están pensando en desarrollar servicios o herramientas para generar conocimiento y proyectos con *machine learning*.

Marc Nolla director de iniciativas estratégicas de SAP DGTAL EMEA, menciona cómo el *machine learning* mejora los sistemas de solicitudes de empleo en el periódico digital sobre informática y tecnología *TicBeat* (2016):

“Esta metodología también permite a los solicitantes de empleo, establecer claramente los puestos vacantes que mejor se adaptan a sus capacidades y cualificaciones. El resultado no es solo un proceso de selección cada vez más eficiente, sino una mano de obra más cualificada y motivada, pues hay mayor probabilidad de que los empleados se adapten perfectamente a los trabajos que realizan.

La utilización *machine learning* comienza a ser imprescindible en una realidad social y tecnológica, donde el aumento de los datos no cesa, la posibilidad de analizarlos rigurosamente y extraer variables de estudio realmente útiles y válidas para aplicar en un ámbito determinado.

En relación con el ámbito nacional se cuenta con varias empresas, instituciones académicas que trabajan con *machine learning*, ya sea para venta de servicio o investigaciones científicas.

El diario costarricense La Nación (2016), menciona en su artículo “Crédito se alía con la inteligencia artificial”:

Con la llegada de sistemas de inteligencia artificial las entidades financieras podrán tener un panorama más amplio acerca del perfil de sus clientes. También menciona que un estudio global de Baker & McKenzie junto con Euromoney Thought Leadership en el 2016, indicó que el sector financiero identificó tres áreas principales donde la inteligencia artificial va a ser aplicada en los próximos tres años. El 49% de los encuestados señaló al área de evaluación de riesgos como la de mayor uso, seguida por los análisis financieros (45%) y las inversiones (37%).

También en el sector académico las universidades públicas están desarrollando carreras y proyectos de investigación, cabe destacar es de mucha importancia para el desarrollo tecnológico de Costa Rica.

Según, Jorge Vargas Calvo Ingeniero en Computación de la Universidad de Costa Rica y con una maestría en Ciencias Cognoscitivas y el experto Santiago Núñez, ex director de Tecnologías Digitales del MICITT, en una entrevista hecha por el periódico digital CostaRica2050 (2014):

“Queda claro que en Costa Rica la Inteligencia Artificial avanza más lento que en los países del primer mundo. Se debe en parte a que en el país el número de personas con doctorado en Inteligencia Artificial no llega siquiera a 15, muchas de ellas no pueden dedicarse a tiempo completo a investigar el tema. No obstante, aquí se han logrado resultados muy positivos. Cita como ejemplo el éxito que ha tenido la empresa costarricense de software ArtinSoft aplicando técnicas arcaicas de traducción de lenguas naturales a la

traducción de lenguas artificiales, lo que les genera grandes ahorros a las empresas.

Además de la misma entrevista hecha por el periódico digital CostaRica2050, Jorge Vargas Calvo, menciona (2014):

“En Costa Rica no existe la carrera de IA. Pero en el nivel de pregrado hay un curso obligatorio de Inteligencia Artificial en la UCR y TEC, y ese curso es opcional en la UNA. Como área de especialización no existe como posgrado, pero el programa de Maestría en Ciencias Cognoscitivas, de la Universidad de Costa Rica, sí tiene a la IA como un área de especialización.

En una entrevista no estructurada que se realizó el martes 20 de febrero del 2017 como forma de generar un estado del arte más completo por medio de expertos se contactó con Saúl Calderón Ramírez experto en inteligencia artificial, Bsc. Computer Science, M. Sc. Ingeniero eléctrico, quien menciona que en el Tecnológico de Costa Rica cuentan con distintas iniciativas, desde el reconocimiento de plantas a partir de imágenes de sus hojas, procesamiento de imágenes biomédicas hasta aplicaciones del *machine learning* en agromática. Cuentan con un grupo especializado llamado PARMA (PAttern Recognition and *Machine learning* group), en el cual varios profesores se unen para incentivar iniciativas de investigación en el área en conjunto con los estudiantes.

El grupo PARMA desarrolla proyectos como “Mejora de la segmentación temporal usando DNLM y DBF” en el cual por medio de filtros se pueden mejorar la

lectura de imágenes y disminución de ruido para aplicaciones en la medicina y la biología. También existe el proyecto de “Mejora del filtro DNLM-IFFT utilizando algoritmos genéticos”.

En el caso de la Universidad de Costa Rica, se contempla *machine learning* en el laboratorio PRIS Lab. Este Laboratorio de Reconocimiento de Patrones y Sistemas Inteligentes es una organización interdisciplinaria que forma parte de la Facultad de Ingeniería Eléctrica de la Universidad de Costa Rica.

Fue fundada en 2012 por el Dr. Rer. Nat. Francisco Siles Canales como núcleo de la formación interdisciplinaria. Busca la generación de la investigación científica y la innovación tecnológica, hacia la integración entre la comunidad académica, el gobierno, el Sector Productivo, la Sociedad Civil y el Medio Ambiente, para mejorar la calidad de vida de las personas.

El PRIS Lab tiene muchos proyectos de investigación entre ellos se destaca el proyecto de “Uso de Reconocimiento de patrones y sistemas inteligentes en aplicaciones de la ingeniería biomédica”, M. Sc Denise Dajles, coordinadora del PRISLab de la UCR, menciona que están desarrollando la capacidad de manipular partes robóticas para prótesis por medio señales cerebrales.

Finalmente, se puede indicar que el estado del arte muestra como el *machine learning*, se aplican en servicios y distintos casos de uso, como las grandes empresas de nivel mundial quienes hacen esfuerzos para generar conocimiento y

en el nivel nacional las universidades hacen sus investigaciones específicas de aplicación con esta tecnología.

El análisis del desarrollo de nuevas tecnologías como *machine learning* promueve una ventaja y además crea una disrupción de servicios, con miras al reforzamiento de recursos y aprendizaje académico en las universidades, por lo que es importante examinarlo más a fondo por medio de las empresas que oferten estos servicios.

1.3 Objetivo General

Elaborar un estudio técnico de la oferta de servicios tecnológicos digitales con *machine learning* por medio de las empresas proveedoras que permita el análisis de sus características principales en la Gran Área Metropolitana de Costa Rica durante el periodo mayo 2017 a junio 2018.

1.4 Objetivos Específicos

Establecer la oferta de servicios tecnológicos digitales basados en *machine learning* existentes en la Gran Área Metropolitana de Costa Rica, para lograr la sectorización de las empresas proveedoras de servicios mediante un planteamiento de características técnicas, funcionales y organizativas.

Identificar las empresas proveedoras que ofertan servicios tecnológicos digitales basados en *machine learning*, por medio de recopilación de información, que permita la definición total de la población al periodo estudiado.

Analizar las características técnicas, funcionales y organizativas del *machine learning* que se encuentran en la oferta de servicios tecnológicos digitales de las empresas identificadas para la evaluación del desarrollo tecnológico.

Generar un informe de resultados que permita la comprensión de las características técnicas, funcionales y organizativas de la oferta de servicios tecnológicos digitales basados en *machine learning*, para la vinculación con las áreas económicas impactadas por la tecnología.

1.5 Problema

Las inversiones en el nivel global en inteligencia artificial están en crecimiento como lo demuestra la Figura 3, según la publicación del periódico El Financiero Costa Rica (2017), que dice: “Firmas locales brindan servicios de inteligencia artificial”. He aquí la importancia de generar conocimiento sobre cómo están caracterizadas estas tecnologías.

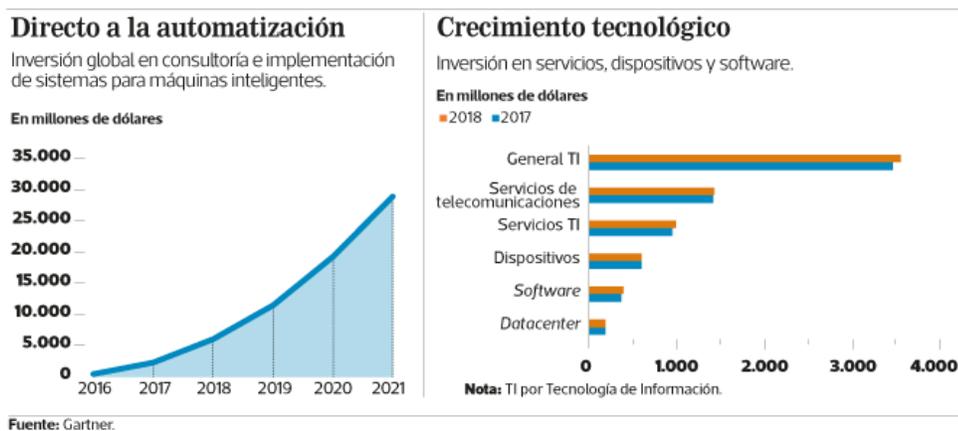


Figura 3. Firms locales brindan servicios de inteligencia artificial, *El Financiero Costa Rica*, 2017.

Fuente: *El Financiero Costa Rica*, 2017.

Se puede ver en la Figura 3 cómo se proyecta un incremento en las inversiones de sistemas para máquinas inteligentes y en las inversiones tecnológicas. Debido al alto crecimiento en el desarrollo de máquinas inteligentes, es oportuno estudiar la oferta de estos servicios por lo que se plantea el problema de elaboración de un estudio técnico de la oferta de servicios tecnológicos digitales basado en *machine learning*, porque esto conlleva una disrupción de los servicios de una economía, la creación y pérdida de empleos, a su vez capacitación de personal profesional especializado.

Por ello, el problema principal es el siguiente:

¿De qué manera las características técnicas, funcionales y organizativas del *machine learning* en las empresas proveedoras que ofertan servicios tecnológicos digitales, afectan el estado de desarrollo de esta tecnología?

Ante esta situación surgen las siguientes interrogantes:

¿Cuál es el estado de desarrollo de la tecnología del *machine learning* en el sector productivo costarricense?

¿Cuáles son los *Frameworks* más utilizados por las empresas proveedoras especialmente para *machine learning*?

¿Cuáles son los métodos de *machine learning* más utilizados por las empresas proveedoras?

¿A cuáles áreas económicas se destinan los productos que utilizan *machine learning*, según las empresas entrevistadas?

¿Qué ventajas tiene para un cliente contar con productos que utilizan *machine learning*?

¿Qué se espera para los siguientes años en materia de desarrollo de *machine learning* en Costa Rica?

1.6 Alcance

Esta investigación pretende que sirva de base informativa para:

Gerentes y funcionarios de las empresas proveedoras de servicios o productos que cuenten con *machine learning*, porque es de suma importancia conocer el estado de la oferta del país para tomar decisiones económicas y estratégicas. También para determinar cuáles áreas económicas son las que demandan.

Estudiantes y graduados de las carreras relacionadas con la rama tecnológica como ingenierías en software, ciencias de la computación, electrónica, entre otros, los cuales pueden requerir información para trabajos de investigación, exámenes o tesis para graduarse.

Empresas extranjeras que se encuentren evaluando la posibilidad de establecerse en el país y que necesiten un panorama general del estado de la economía costarricense y cómo se comporta el crecimiento de la oferta de servicios con *machine learning* en el mercado nacional.

Personas que realicen estudios sobre análisis de datos, que puedan apoyarse en esta investigación para recabar información respecto del tema, que la presente investigación facilite la información necesaria para llevar a cabo los estudios que deseen realizarse en el área de informática.

1.7 Limitaciones

Acceso limitado a las empresas proveedoras de servicios con *machine learning* con el objetivo de recabar información de la oferta de servicios tecnológicos digitales, que permita el análisis de sus características principales, evitando que sea suficiente para la obtención de resultados, así como el impedimento a información confidencial que proporcione datos reales y estadísticos.

Una base de datos fidedigna de empresas que oferten servicios tecnológicos digitales basados en *machine learning* para el análisis en la investigación.

Medios desactualizados de comunicación de las empresas, lo que impide el acercamiento para la aplicación de los instrumentos de recopilación de información.

1.8 Matriz de congruencia

Título	Problema	Objetivo General	Objetivo Específico	Pregunta de investigación
Estudio técnico de las características y la oferta de servicios tecnológicos digitales basados en tecnología de <i>machine learning</i> en la Gran área Metropolitana de Costa Rica en el periodo	¿De qué manera las características técnicas, funcionales y organizativas del <i>machine learning</i> en las empresas proveedoras que ofertan servicios tecnológicos digitales, afectan el estado de	Elaborar un estudio técnico de la oferta de servicios tecnológicos digitales con <i>machine learning</i> por medio de las empresas proveedoras que permita el análisis de sus características	Establecer la oferta de servicios tecnológicos digitales basados en <i>machine learning</i> existentes en la gran área metropolitana de Costa Rica, para lograr la sectorización de las	¿Cuál es el estado de desarrollo de la tecnología del <i>machine learning</i> en el sector productivo costarricense? ¿Cuáles son los <i>Frameworks</i> más utilizados por las

<p>mayo 2017 a junio 2018.</p>	<p>desarrollo de esta tecnología?</p>	<p>principales en la Gran Área Metropolitana de Costa Rica durante el periodo mayo 2017 a junio 2018.</p>	<p>empresas proveedoras de servicios mediante un planteamiento de características técnicas, funcionales y organizativas.</p> <p>Identificar las empresas proveedoras que ofertan servicios tecnológicos digitales basados en <i>machine learning</i>, por medio de recopilación de información,</p>	<p>empresas proveedoras especialmente para <i>machine learning</i>?</p> <p>¿Cuáles son los métodos de <i>machine learning</i> más utilizados por las empresas proveedoras?</p> <p>¿A cuáles áreas económicas se destinan los productos que utilizan <i>machine</i></p>
--------------------------------	---------------------------------------	---	---	--

		que permita la definición total de la población al periodo estudiado.	<i>learning</i> , según las empresas entrevistadas?
		Analizar las características técnicas, funcionales y organizativas del <i>machine learning</i> que se encuentran en la oferta de servicios tecnológicos digitales de las empresas identificadas para la evaluación del	¿Con qué ventajas puede contar un cliente que utiliza productos con <i>machine learning</i> ? ¿Qué se espera para los siguientes años en materia de desarrollo de <i>machine</i>

			<p>desarrollo tecnológico.</p> <p>Generar un informe de resultados que permita la comprensión de las características técnicas, funcionales y organizativas de la oferta de servicios tecnológicos digitales basados en <i>machine learning</i>, para la vinculación</p>	<p><i>learning</i> en Costa Rica?</p>
--	--	--	---	---------------------------------------

			con las áreas económicas impactadas por la tecnología.	
--	--	--	--	--

Tabla 2 Matriz de congruencia

Fuente: *Elaboración propia.*

Capítulo II

Marco Teórico

2.1 La inteligencia artificial y su impacto en la vida moderna.

Para tener una mejor introducción al tema de inteligencia artificial, en adelante se citan algunos autores que mencionan el concepto y su impacto en la vida moderna.

El concepto de inteligencia artificial según Luger and Stubblefield, (1993) : "La rama de la informática que se ocupa de la automatización del comportamiento inteligente". (pág. 1)

Los autores Luger y Stubblefield explican que, la inteligencia artificial es la rama de la informática donde la define como un sistema de información que actúan de manera racional y automatizada. Es decir, un sistema inteligente actúa por medio de ciertos modelos o leyes.

La inteligencia artificial, como herramienta para la automatización de procesos, puede generar disrupción en el nivel económico, además impacta a los empleos que existen. Para sustentar lo anterior, según el periódico digital español Navarra Información, en su publicación con el título "Casi la mitad de los trabajos actuales desaparecerá en diez años, según SAP" del 2017 se menciona que:

"En el Fórum de Innovación de SAP celebrado en Madrid, Franck Cohen ha incidido en la "realidad" de que "todo va a ser automatizado" y ha mostrado su preocupación por que las autoridades europeas y las universidades no sean lo suficientemente conscientes del impacto de la inteligencia artificial de cara al empleo futuro".

En el caso de América Latina ya se estima un crecimiento de los ingresos en la inteligencia artificial en el área, según el sitio Merca 2.0 sobre los números que arroja un estudio (2017). Esto se puede ver en la Figura 4,.

INTELIGENCIA ARTIFICIAL EN AMÉRICA LATINA

Tractica estimó que los ingresos de la inteligencia artificial crecerán dentro del mercado de aplicaciones empresariales en América Latina de manera constante, tal como se muestra en la siguiente gráfica.

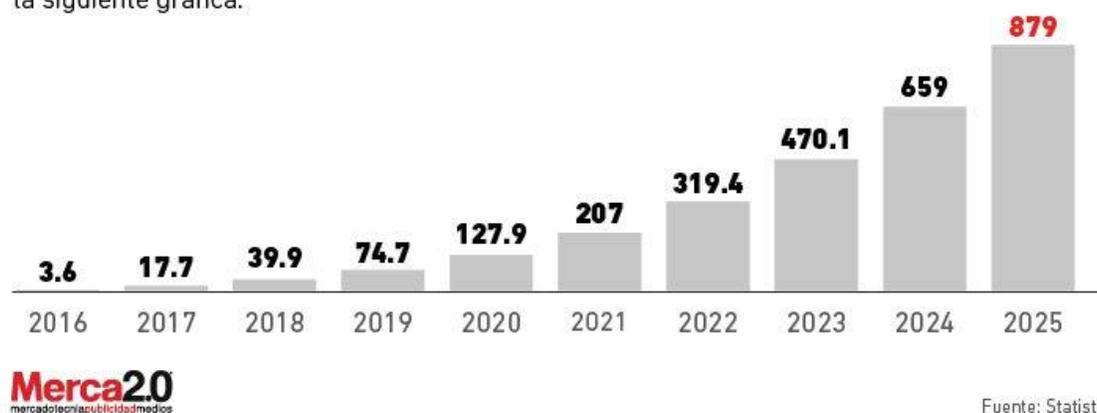


Figura 4. Inteligencia artificial en América Latina, Statista.

Fuente: Inteligencia artificial en América latina, qué números hay hasta ahora, Merca 2.0, (2017)

Como se nota la Figura 4 sobre el crecimiento que pronostica Statista portal estadístico español, se puede apreciar cómo la inteligencia artificial estará en el mercado de aplicaciones empresariales lo cual supondría un alto interés en el desarrollo tecnológico y apostando a la automatización.

Cada vez hay más sistemas inteligentes en la vida moderna y los puestos de trabajo podrían ser aún más especializados en el futuro. La eficiencia de las máquinas inteligentes es algo que no se puede evitar, estará cada vez más presente en la población y en la vida moderna.

2.2 *Machine learning*, tecnología del futuro

A manera de introducción, se analiza el concepto de *machine learning*, sus orígenes, aportes de nivel mundial y características internas. Para sustentar este marco teórico se recurre a varios autores y empresas dedicadas al *machine learning* en el nivel internacional, permitiendo clarificar esta innovadora y fascinante tecnología.

Existen muchas definiciones de *machine learning*, por lo cual se van a citar dos definiciones que clarifican su concepto: el primero es de Tom Mitchell (1997) quien en su libro *machine learning* proporciona una en la primera línea del prefacio: “El campo del *machine learning* se refiere a la cuestión de cómo construir programas informáticos que mejoren automáticamente con la experiencia.” (pág. prefacio)

La definición citada en el párrafo anterior es una definición corta, simple y es la base para entender este campo tecnológico. Pero deja incógnitas sobre el funcionamiento interno, por lo que Tom Mitchell formaliza el concepto en el mismo libro, con la siguiente definición (1997): “Se dice que un programa de computadora aprende de la experiencia E respecto de alguna clase de tarea T y midiendo su desempeño P , si su desempeño en la tarea T , medido por P , mejora con la experiencia E .” (pág. prefacio)

Para mayor claridad de estos términos, se establece la siguiente analogía de creación propia: “El ser humano para aprender a conducir una bicicleta puede caer varias veces hasta que logra mantener el equilibrio y como resultado controlar la

bicicleta; esto es la esencia pura del *machine learning*, donde sistemas de información o máquinas pueden aprender de experiencias y mejorar su desempeño”.

Según, la publicación por la revista Mundo Ejecutivo (2017) el 27 abril 2017, “*Machine learning*, Una expresión de la inteligencia artificial” dice Héctor Cobo, director regional, SAS México y Caribe:

“*Machine learning* se está convirtiendo en un factor que detona una profunda innovación, dado que su aplicación contribuye a mejorar procesos, generar nuevas oportunidades de negocio (antes imperceptibles o cuya detección no era sencilla) y definir nuevos parámetros empresariales (por ejemplo, la forma en que se consigue la satisfacción del cliente). Todo esto desde la propia solución tecnológica, sin que una instancia humana tenga que estudiar los datos, valorarlos y generar una respuesta.

Dado que el *machine learning* es un factor de innovación es pertinente conocer la situación del mercado costarricense específicamente en los servicios digitales en relación con la inteligencia artificial y otras ramas relacionadas.

2.3 Situación del mercado de servicios digitales relacionados con *machine learning* en Costa Rica.

Las tecnologías son de vital importancia para cualquier nación, especialmente en un mundo globalizado y donde las plataformas de servicios

digitales son sumamente necesarias. La evolución tecnológica se acelera de manera exponencial y surgen nuevas áreas tecnológicas.

La estabilidad económica de Costa Rica en el futuro depende de la acción de los involucrados en el crecimiento; en otras palabras, la población, el Gobierno, el entorno y uno responsables externos como la tecnología.

El primer informe Estado de la Ciencia, la Tecnología y la Innovación, realizado por el Programa Estado de la Nación (PEN) (2014) , del Consejo Nacional de Rectores (CONARE), con el auspicio del Banco Interamericano de Desarrollo (BID) y la Asociación Estrategia Siglo XXI: “Propone hacer más visible el papel central de la ciencia y la tecnología para el futuro del país y aportar información que sirva de apoyo para la toma de decisiones en este campo”. (pág. 23)

Las dimensiones analizadas en el Primer Informe Estado de la Ciencia, la Tecnología y la Innovación son recursos humanos, producción científico tecnológica, infraestructura científico-tecnológica y sistema de apoyo.

Según, el estado de la ciencia, la tecnología y la innovación, Figura 5, muestra una situación preocupante en temas de ciencias, tecnología e innovación.

Cuadro 0.12

Áreas de fortaleza relativa, alertas y temas críticos en ciencia, tecnología e innovación en Costa Rica

Fortalezas relativas	Premio laboral	<input checked="" type="checkbox"/>	
	Perfil de la diáspora científica	<input checked="" type="checkbox"/>	
	Uso compartido de la infraestructura	<input checked="" type="checkbox"/>	
Alertas	Recursos humanos para el desarrollo	<input type="checkbox"/>	
	Oferta de recursos humanos	<input type="checkbox"/>	
	Estructura ocupacional	<input type="checkbox"/>	
	Comunidades científicas	<input type="checkbox"/>	
	Fuga o movilidad de cerebros	<input type="checkbox"/>	
	Reinserción de la diáspora científica	<input type="checkbox"/>	
	Liderazgo en la apropiación de conocimiento	<input type="checkbox"/>	
	Vinculación de las comunidades de investigación	<input type="checkbox"/>	
	Idoneidad de la infraestructura	<input type="checkbox"/>	
	Perfil de las políticas públicas	<input type="checkbox"/>	
	Cumplimiento del <i>Plan de Medio Siglo</i>	<input type="checkbox"/>	
	Temas críticos	Competencias de los jóvenes	<input checked="" type="checkbox"/>
		Liderazgo en la producción de conocimiento	<input checked="" type="checkbox"/>
Correspondencia entre el conocimiento generado y la política pública		<input checked="" type="checkbox"/>	
Incentivos para la producción y la vinculación		<input checked="" type="checkbox"/>	
Suficiencia de la inversión en I+D		<input checked="" type="checkbox"/>	
Entorno para los emprendimientos innovadores		<input checked="" type="checkbox"/>	

Fortaleza Alerta Área crítica

Figura 5 Áreas de fortaleza, alertas y temas críticos. Estado de Nación 2014.

Fuente: Estado de la Nación, 2014.

Los grandes desafíos que menciona el estudio, como la atracción de jóvenes hacia áreas de ciencias y tecnologías desde la secundaria, es un tema relevante, dado que en la educación superior conviene fomentar las carreras relacionadas con la ciencia, tecnología e innovación, mediante incentivos para que las universidades aumenten la matrícula en disciplinas consideradas prioritarias y que tengan demanda en el mercado laboral.

Otra área de preocupación son las comunidades científicas, aunque Costa Rica tiene una notable posesión de expertos en múltiples campos, sus comunidades científicas muestran problemas de relevo generacional, baja redundancia y falta de colaboración interdisciplinaria e interinstitucional.

El Primer Informe Estado de la Ciencia, la Tecnología y la Innovación (2014) señala que la formación de profesionales y técnicos altamente calificados será un tema importante: “A corto plazo Costa Rica debe aumentar la cantidad y las destrezas de sus científicos, ingenieros y técnicos medios y superiores, en un marco en el que prevalezca la equidad de género”. (pág. 55)

Con la finalidad de incremento progresivo de los cupos de matrícula en las carreras universitarias de mayor demanda proyectada, como las ingenierías y la Informática.

Según, la Caja Costarricense de Seguro Social (2017), hay un aumento de negocios y empleo en el sector local. Entre 2010 y 2016, la cantidad de empresas en varias áreas tecnológicas creció 23%, a 1.067. En el mismo periodo el empleo del sector aumentó 33% a 28.941 empleos reportados en planillas esto demuestra un alta de profesional calificado. Ver Figura 6.

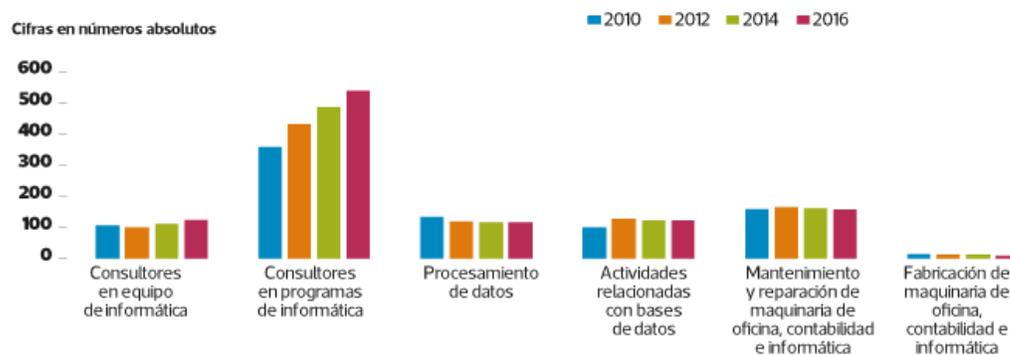
La industria tecnológica en Costa Rica

Cantidad de empresas y empleo por áreas de actividad.

EMPRESAS INFORMÁTICAS

Empresas según rama de actividad.

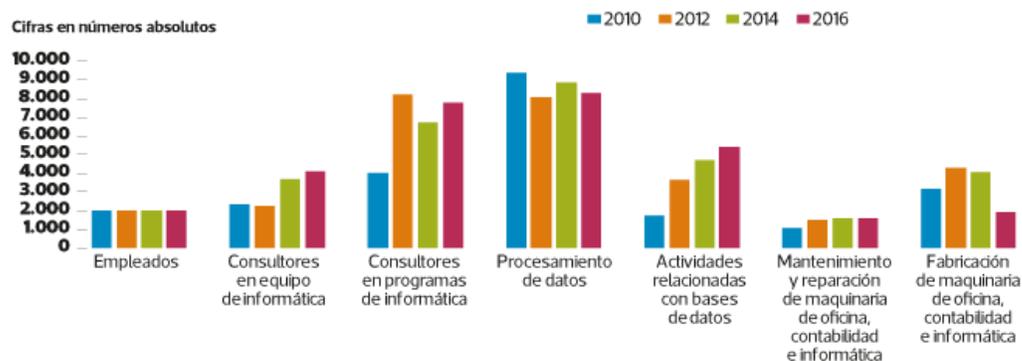
Cifras en números absolutos



PLANILLAS

Empleados según rama de actividad.

Cifras en números absolutos



Fuente: CCSS, Estadísticas.

Figura 6 La industria tecnológica en Costa Rica.

Fuente: CCSS, Departamento de Estadísticas 2016.

Se puede observar en la Figura 6, un incremento de las empresas del área de consultores en programas de informática. Un dato de interés de este informe de la CCSS, ha sido el aumento en la cantidad de empleados del área de procesamiento de datos a pesar de tener menor cantidad de empresas en esa rama tecnológica.

El periódico El Financiero (2017) en su publicación sobre “*Nuevas tecnologías impulsarán empleo de informáticas ticas en 2017*”, menciona la planificación de cómo van a modificar sus servicios para el año 2017, cabe destacar que muchas de estas empresas implementaran dentro de sus servicios la tecnología de *machine learning*, esto se puede ver en la Figura 7 tomada de la publicación del periódico.

Portafolios para crecer

Servicios y planes de firmas para 2017 y su efecto en sus planillas:

Firma	Áreas e Impacto
Avantica	Desarrollos en la nube, servicios para facilitar desarrollo y entrega continua, pruebas automatizadas de software , así como big data , computación cognitiva y aprendizaje de máquina. Espera pasar de 600 a 700 empleados en sus seis sedes al final del 2017.
Freeway	Continuará centrado en soluciones de nube para la atención de los clientes en las empresas. Empieza a percibir oportunidades en inteligencia artificial aplicada a ventas y mercadeo. Tiene 150 personas en tres países y espera aumentar 50% en el año.
Mobilize	Seguirá con adopción de la nube e inteligencia artificial para desarrollo y modernización de apps. Identifican oportunidades en big data para apps . Crecería entre 20-25% su actual planilla de 150 personas.
Novacomp	Trabaja en áreas de fabrica de software , outsourcing en software , inteligencia de negocios e infraestructura, entre otros. Tiene expectativas de alcanzar 400 colaboradores en el 2017.
Proximity	Continúa con plataformas de comercio electrónico de alta complejidad de implementación e integración. Exploran mercados para big data , IoT , aprendizaje de máquina, inteligencia artificial y otras. Esperan aumentar 50% la planilla.
Softland	Continúa con apps y traslado de plataformas en la nube, así como analítica de datos. Están investigando oportunidades en IoT e inteligencia artificial. Se estima que la actual planilla de 100 personas en el país tendrá un "leve crecimiento".

Figura 7. Servicios y planes de firmas informáticas 2017.

Fuente: *El Financiero, (Nuevas tecnologías impulsarán empleo de informáticas ticas en 2017, 2017)*

Cabe destacar las palabras de Rodrigo Briceño, Gerente de Ventas sénior de Avantica, en una entrevista para el periódico El Financiero (2017):“*Big data*, computación cognitiva y *machine learning* son también áreas que están cobrando mayor interés de cara a nuestros clientes y prospectos”.

Un ejemplo es el avance en desarrollo de *chatbots* en Costa Rica se ve en el desarrollo de Layla, un nuevo *chatbot* por Impesa y Equifax, enfocado en simplificar los trámites de la banca convencional.

En un reportaje del periódico El Financiero, “*Nuevo chatbot puede analizar su información para gestionarle créditos*” (2018), menciona cómo funciona Layla: “Este novedoso robot se basa en dos tecnologías principales. El procesamiento del lenguaje natural que le permite identificar el lenguaje escrito y hablado de un ser humano y las técnicas de *machine learning*”.

Precisamente muchas empresas están investigando oportunidades en el área de inteligencia artificial, por su valor agregado y su demanda como sistemas inteligentes. Esto demuestra que es un salto y un cambio en la oferta de servicios.

Para especificar el enfoque de esta investigación se agregan las siguientes definiciones sobre los servicios tecnológicos digitales para una mayor comprensión.

2.4 Servicios tecnológicos digitales.

Como parte de los objetivos de esta investigación se presenta la definición de servicios tecnológicos digitales, para comprender el tipo de oferta que se está analizando y tener una mejor sectorización.

Lovelock, C. (2009) define los servicios como:

“Actividades económicas que se ofrecen de una parte a otra, las cuales generalmente utilizan desempeños basados en el tiempo para obtener los

resultados deseados en los propios receptores, en objetos o en otros bienes de los que los compradores son responsables”. (pág. 15)

Para una mejor definición de los tipos de servicios tecnológicos digitales, se presentan los siguientes párrafos que permitirán una definición clara sobre el tema.

2.4.1 Sistemas digitales

Para aclarar el término anterior sobre las tecnologías digitales, se toma la definición dada por Tocci R. (2006):

“Un sistema digital es una combinación de dispositivos diseñados para manipular información lógica de cantidades físicas que se representan en forma digital; es decir, las cantidades solo pueden tomar valores discretos. Estos dispositivos suelen ser electrónicos, pero también pueden ser mecánicos, magnéticos o neumáticos. Algunos de los sistemas digitales más familiares incluyen computadoras y calculadoras digitales, equipos de audio y video digitales y el sistema telefónico”. (pág. 5)

Dado lo anterior se quiere delimitar a sistemas computacionales que están dentro del campo de los sistemas digitales como lo puede ser un software basado en *machine learning*, como lo menciona en la cita anterior, las computadoras son los sistemas digitales más familiares.

2.5 Empresas proveedoras de servicios digitales.

Como aspecto inicial de este tema se toma la definición de proveedor, según *Debitoor* empresa financiera (2018): “Se trata de la persona que surte a otras empresas con existencias necesarias para el desarrollo de la actividad.”

También es importante entender el tipo de proveedor que en este caso es un proveedor de servicios en el área de tecnología. Se cita el sitio sobre economía *Economía Simple* (2018): “Los proveedores de recursos responden a las necesidades económicas de la empresa y por último los proveedores de servicios atienden a las necesidades del cliente.”

Por lo tanto, una empresa proveedora de servicios tecnológicos es aquella que surte a otras empresas o clientes una necesidad de servicios tecnológicos. Además, estas empresas dentro de sus servicios pueden contener servicios digitales con *machine learning* que son de interés en esta investigación.

2.6 Características de *machine learning*

Machine learning cuenta con una gran cantidad de características por ser un campo multidisciplinario. Para que sea más clara su comprensión se definen las características técnicas, funcionales y organizativas, como lo son:

Características técnicas

- *Frameworks*
- Métodos

- Estilos
- Tipos de enfoques

Características funcionales

- Casos de uso específico
- Ventajas
- Retos de implementación
- Tipos de consumo

Características organizativas

- Necesidades principales para desarrollar *machine learning*
- Proyección del desarrollo tecnológico de *machine learning* en Costa Rica.
- Recurso humano
- Capacitación de personal

En los siguientes apartados se encontrará información relevante sobre las características que fueron listadas anteriormente.

2.6.1 Framework de Machine learning

Un *framework* o marco de trabajo es un término utilizado para referirse al conjunto de bibliotecas de software que se usan para desarrollar una aplicación de manera estandarizada, es decir, sirven como base para el desarrollo de software o aplicación, ya que contienen las herramientas, prácticas y criterios para un problema en particular.

Un *framework* está compuesto por varias librerías que han sido diseñadas para un problema en particular. Para conceptualizar la definición se usa la Figura 8 tomada de la publicación titulada “*Libraries vs. Frameworks*” de Bobby Brennan en el medio digital *Medium.com* (2017), que explica como es el flujo de uso de un framework. Un *framework* utiliza un código lógico programático (“Your code”), el cual posee una lista de comandos por seguir, el *framework* contiene las librerías que ayudarán a ejecutar el código inicial.

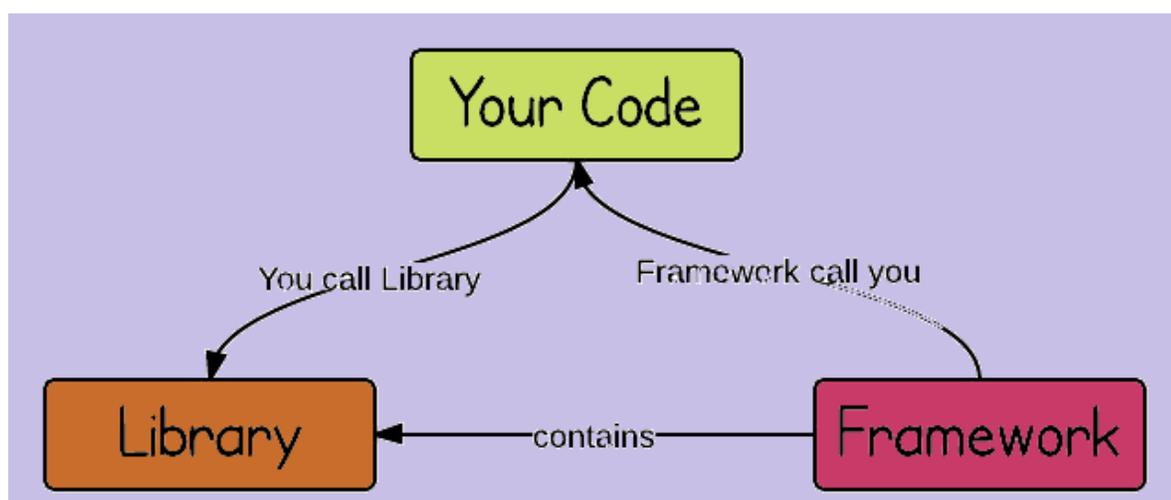


Figura 8 Como funciona un framework y su interpretación visual.

Fuente: Bobby Brennan, *Libraries vs. Frameworks*, *Medium.com* (2017).

En los siguientes párrafos se explican los diferentes tipos de *frameworks* para exponer el tema.

2.6.1.1 Librerías o bibliotecas de *machine learning*

Se define una librería o bibliotecas según Baris Simsek (2004): “Las bibliotecas permiten a los programadores escribir códigos más pequeños y modulares.”

Una librería o biblioteca de *machine learning* es la forma modular de un código diseñado para problemas específicos de desarrollo de aplicaciones o software con *machine learning*.

En la Figura 9 de elaboración propia, se puede observar cómo funcionan las librerías y como pueden ser llamadas por un programa o sistema computacional. Esto es respaldado por la definición de Tony Reno (2017), en el sitio web *Quora* donde define una librería como:

“Un código que se escribió anteriormente, al que se puede recurrir al crear su propio código. Es un trabajo que anteriormente fue realizado por alguien más que se puede ser reutilizar. Y puede hacer uso de él en su propio código, legalmente, tal como fue escrito para ser usado por otros.

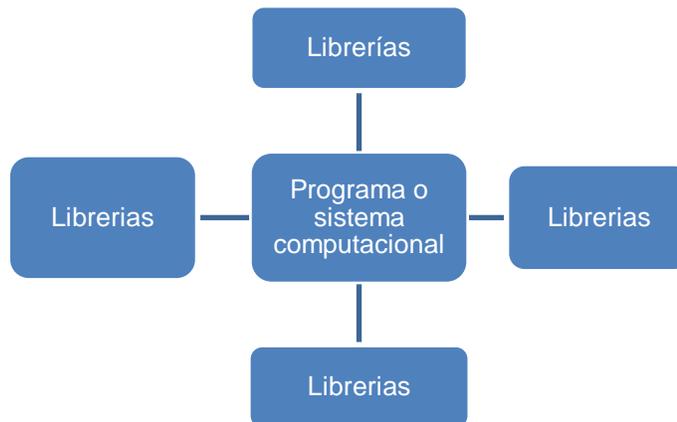


Figura 9 Concepto de flujo de uso de librerías.

Fuente: *Elaboración propia.*

2.6.1.2 Open source

Existen *frameworks* de código abierto o también llamados *Open Source* por lo que se extrae la definición del sitio *OpenSource* (2018) menciona que:

"Código fuente" es la parte del software que la mayoría de los usuarios de computadoras nunca ven; es el código que los programadores de computadoras pueden manipular para cambiar la forma de software, un "programa" o "aplicación" - trabaja. Los programadores que tienen acceso al código fuente de un programa de computadora y pueden mejorar ese programa agregándole características o arreglando partes que no siempre funcionan correctamente".

A continuación, se llevará a cabo un análisis de diferentes *frameworks*, con la finalidad de que el lector tenga una visión más extensa del tema:

- *TensorFlow*
- Apache Spark
- Scikit-learn

2.6.1.2.1 *TensorFlow*

TensorFlow es uno de los *frameworks* más utilizados en *machine learning* dado a su gran respaldo que tiene de *Google*, su flexibilidad y rapidez. Además, esta biblioteca permite resolver la mayoría de los problemas en el campo del *machine learning*.

Se define como:

*TensorFlow*TM Es una biblioteca de software libre que se utiliza para realizar cálculos numéricos mediante diagramas de flujo de datos. Los nodos de los diagramas representan operaciones matemáticas y las aristas reflejan las matrices de datos multidimensionales (tensores) comunicadas entre ellas. Gracias a la flexibilidad de la arquitectura, solo necesitas una API para desplegar el sistema informático de una o varias CPU o GPU en un escritorio, servidor o dispositivo móvil. En su origen, *TensorFlow* fue fruto del trabajo de investigadores e ingenieros de Google Brain Team que formaban parte de la organización de investigación del aprendizaje automático de Google. Su objetivo era realizar investigaciones en el campo del aprendizaje automático

y las redes neuronales profundas. A pesar de que este era su propósito inicial, se trata de un sistema lo bastante general como para poder aplicarse en muchos otros campos. (TensorFlow, 2018)

Dentro de sus ventajas principales:

- Portabilidad
- Diferenciación automática
- Rendimiento

Rick Blaisdell (2017), en el artículo “*Meeting the most popular deep learning framework – TensorFlow*”, publicado en su sitio web *Rickscloud.com* menciona que: “Una de las principales preocupaciones que tienen las empresas son la portabilidad. Sin embargo, *TensorFlow* supera este desafío. La plataforma se ejecuta en CPU, GPU, servidores y plataformas de computación móvil, lo que significa que no está limitado por hardware.”

Además, manifiesta que la diferenciación automática es una gran ventaja, así lo expresa en el artículo: “La función de diferenciación automática mejora los algoritmos de *machine learning* basados en gradientes, lo que le permite diseñar la arquitectura de su modelo predictivo y combinarla con su función objetivo y agregarle datos.”

Para entender la diferenciación automática y por qué es una ventaja que consigue *TensorFlow*, se cita una explicación muy completa por parte de Atilim Gunes Baydin (2018), menciona que:

La diferenciación automática (DA), también llamada diferenciación algorítmica o simplemente "autodif.", es una familia de técnicas similar, pero más general que la propagación hacia atrás para evaluar de manera eficiente y precisa los derivados de funciones numéricas expresadas como programas de computadora. DA es un campo pequeño, pero establecido con aplicaciones en áreas que incluyen dinámica de fluidos computacional, ciencias atmosféricas y optimización de diseño de ingeniería. Hasta hace muy poco, los campos del *machine learning* y la DA no se conocían entre sí y, en algunos casos, se habían descubierto de forma independiente los resultados de cada uno. A pesar de su relevancia, la DA de propósito general ha faltado en la caja de herramientas de *machine learning*, una situación que cambia lentamente con su adopción continua bajo los nombres de "gráficos computacionales dinámicos" y "programación diferenciable".

En cuanto al rendimiento de *TensorFlow*, Rick Blaisdell (2017), en el artículo "*Meeting the most popular deep learning framework – TensorFlow*", publicado en su sitio web *Rickscloud.com* expresa que: "la mejora del rendimiento siempre ha estado en la lista de todos los propietarios de negocios, y *TensorFlow* aparentemente lo ofrece. La plataforma admite colas,

subprocesos y cómputo asíncrono que le permite sacar el máximo provecho de su hardware disponible.”

Entre sus principales casos de uso, según la empresa de servicios de *machine learning* Exastax (2017) en la publicación “Los cinco principales casos de uso de *TensorFlow*”, se listan los siguientes:

- Reconocimiento de voz / sonido
- Aplicaciones basadas en texto
- Reconocimiento de imágenes
- Series de tiempo
- Detección de video

Para ejemplificar el concepto se toma la publicación en el sitio *Mesosphere* (2017) que explica cómo funciona *TensorFlow*, en la Figura 10, la capa de entrada es responsable de encontrar patrones de contraste local, la capa 1 oculta es responsable de encontrar rasgos faciales individuales basados en esos contrastes y la capa 2 oculta es responsable de identificar caras enteras basadas en esos rasgos faciales.

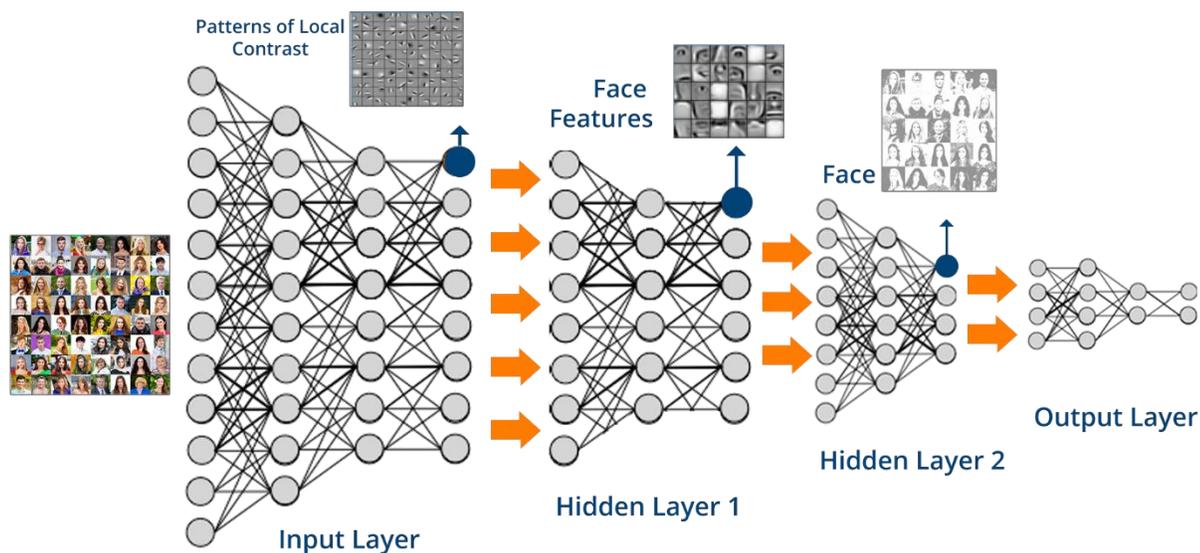


Figura 10 Ejemplo de funcionamiento de TensorFlow con DeepLearning

Fuente: *Distributed TensorFlow on Any Infrastructure with Mesosphere*, Mesosphere (2017)

2.6.1.2.2 Apache Spark

Apache Spark (Foundation, 2014): “es un *framework* de clúster de código abierto para el análisis de datos a gran escala, rápido y flexible”.

Apache Spark es reconocido por su velocidad de procesamiento y fácil uso, por lo que se ha vuelto muy popular. Este *framework* es propiedad de *Apache Software Foundation* en el marco de código abierto.

Para entender mejor el concepto de *Apache Spark*, Abraham Requena Mesa un científico de datos en España menciona (2018):

“*Apache Spark* es un sistema de computación que se basa en *Hadoop Map Reduce* y que, principalmente, permite dividir o paralelizar el trabajo, ya que

normalmente se instala en un clúster de máquina. La idea es que tengamos n máquinas, por ejemplo, diez máquinas, y cada una de esas instancias va a tener instalada una versión de *Apache Spark*”.

Apache Spark en otras palabras, permite al usuario dividir el trabajo con una cierta cantidad de computadores, cada computador se encarga de procesar una parte del trabajo en paralelo y al final unificar los resultados. Se puede utilizar para diversas tareas, como el *machine learning*.

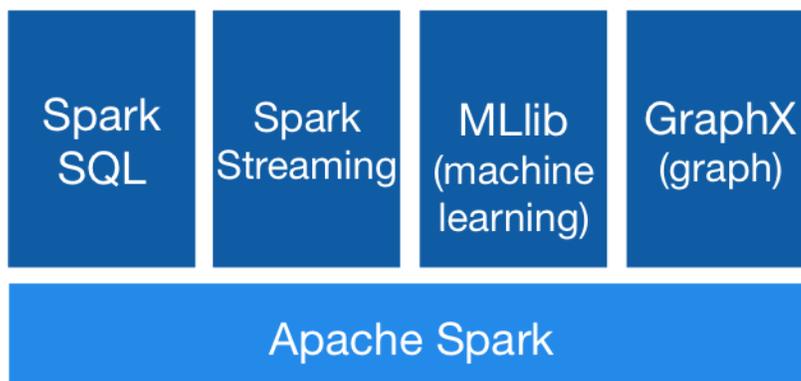


Figura 11 Conjunto de librerías que compone Apache Spark.

Fuente: Sitio oficial de Apache Spark (Spark, 2018)

En la *Figura 11*, se puede observar el término *MLlib*, el cual es la librería destinada a *machine learning* en *Apache Spark*.

Ari Amster (2016), en el sitio Qubole sobre “Los Mejores Casos De Uso De Apache Spark”, menciona la utilidad y casos de uso de la librería:

MLlib puede trabajar en áreas tales como agrupación, clasificación y reducción de dimensionalidad, entre muchas otras. Todo esto permite utilizar Spark para algunas funciones de *big data* muy comunes, como inteligencia

predictiva, segmentación de clientes con fines de marketing y análisis de sentimientos. Las compañías que usan un motor de recomendación encontrarán que *Spark* hace el trabajo rápidamente.

2.6.1.2.3 Scikit-learn

Se define *Scikit-learn*, según el sitio *Scikit-learn.org* (2018) como: “Herramientas simples y eficientes para la minería de datos y el análisis de datos. De fácil acceso, reutilizable en diversos contextos, construido en *NumPy*, *SciPy* y *matplotlib* y de código abierto, utilizable comercialmente.”

Es uno de los *frameworks* usados para *machine learning* basado en el lenguaje de programación *Python*, que provee algoritmos supervisados y no supervisados. Es una de las grandes opciones por científicos de datos, esto lo menciona en el artículo de *Data Science Plus* (2018):

Scikit-learn ofrece una gama de algoritmos de aprendizaje supervisados y sin supervisión por medio de una interfaz consistente en *Python*. *Scikit-learn* se basa en el *SciPy*. Entonces, antes de que puedas usar *scikit-learn*, debes instalar *SciPy*. Los módulos o extensiones para *SciPy* se denominan convencionalmente como *SciKits*. El módulo, como tal, proporciona algoritmos de aprendizaje y se conoce como *scikit-learn*.

George Serif (2018), en su publicación “*An Introduction to Scikit Learn: The Gold Standard of Python Machine Learning*”, explica la función de cada librería que contiene *Scikit-learn*:

- NumbPy: Para cualquier trabajo con matrices, especialmente operaciones matemáticas.
- SciPy: La computación científica y técnica.
- Matplotlib: Visualización de datos
- IPython: Consola interactiva para Python
- Sympy: Matemáticas simbólicas
- Pandas: Manejo, manipulación y análisis de datos.

Estas librerías utilizadas en conjunto son sumamente poderosas, además Ben Lorica (2015), en su artículo “Seis razones por las que recomiendo *scikit-learn*” explica el gran respaldo de la comunidad:

Scikit-learn incluye herramientas para muchas de las tareas estándar de *machine learning* (como agrupamiento, clasificación, regresión, entre otros). Y dado que *scikit-learn* es desarrollado por una gran comunidad de desarrolladores y expertos en *machine learning*, las nuevas técnicas tienden a incluirse en un plazo bastante breve.

Algunos casos de uso tomados del sitio *Scikit-Learn.org* (2018) como es el caso de la empresa INRIA:

Utilizamos *scikit-learn* para apoyar la investigación básica de vanguardia en muchos equipos: *Parietal* de neuroimagen, *Lear* de visión por ordenador, los *Visage* para análisis de imágenes médicas, *Privatics* para la seguridad. El proyecto es una herramienta fantástica para abordar las aplicaciones difíciles del *machine learning* en un entorno académico, pues es versátil y eficaz, pero es fácil de usar y está bien documentado, lo que lo hace muy adecuado para los estudiantes de posgrado.

Scikit-learn es una herramienta enfocada en *machine learning*, con un acompañamiento de una comunidad científica importante que lo vuelve muy popular, además cualquiera puede descargar el *framework* por lo que se vuelve una gran herramienta para principiantes en el tema.

2.6.1.3 Frameworks basados en la nube.

Existen *frameworks* o servicios basados en la nube que utilizan el modelo de infraestructura como servicio.

Según, TechTarget (2018) explica el modelo de infraestructura como servicio:

“Infraestructura como servicio (IaaS) es una forma de computación en la nube que proporciona recursos informáticos virtualizados a través de internet. IaaS es una de las tres categorías principales de servicios de computación en la nube, junto con el software como servicio (SaaS) y la plataforma como servicio (PaaS).

Se analizarán los *frameworks* basados en la nube, entre los más populares se encuentran:

- Amazon Machine Learning
- Microsoft Azure Machine Learning
- Google Cloud Machine Learning
- IBM Watson Machine Learning

2.6.1.3.1 Amazon Machine Learning

Amazon Machine Learning utiliza el modelo de infraestructura como servicio y proporciona una serie de herramientas específicamente de *machine learning* a través de internet. Para tener una definición *Amazon* menciona (2018):

“*Amazon Machine Learning (Amazon ML)* es un servicio potente y basado en la nube que permite a los desarrolladores usar fácilmente la tecnología de aprendizaje automático independientemente de su nivel de cualificación. *Amazon ML* ofrece asistentes y herramientas de visualización que le guían por el proceso de creación de modelos de aprendizaje automático (ML) sin tener que aprender algoritmos y tecnologías ML complejos. Una vez elaborados los modelos, *Amazon ML* permite obtener fácilmente predicciones para una aplicación mediante *API* sencillas y sin tener que escribir código de generación de predicciones personalizado ni administrar ninguna infraestructura.

En la Figura 12, se pueden observar los beneficios que lista *Amazon Machine Learning* (2019).

Cree modelos de aprendizaje automático con facilidad	De modelos a predicciones en segundos	Servicio de generación de predicciones escalable y de alto desempeño
<p>Los asistentes y las API de Amazon Machine Learning facilitan a los desarrolladores la creación y el ajuste de modelos de aprendizaje automático de datos almacenados en Amazon S3, Amazon Redshift o Amazon RDS, así como la consulta de estos modelos para obtener predicciones. Los procesadores de datos, algoritmos de ML escalables, datos interactivos, herramientas de visualización de modelos y alertas de calidad integrados en el servicio le permiten crear y refinar sus modelos con rapidez.</p>	<p>Amazon Machine Learning es un servicio administrado que facilita la creación, la implementación y la monitorización completas de modelos. Una vez que el modelo está listo, puede generar predicciones para sus aplicaciones con rapidez y confianza, eliminando el tiempo y la inversión necesarios para crear, escalar y mantener una infraestructura de aprendizaje automático.</p>	<p>Las API de predicción de Amazon Machine Learning se pueden usar para generar miles de millones de predicciones para sus aplicaciones. Puede solicitar predicciones para numerosos registros de datos a la vez mediante el API de predicciones en lote o utilizar el API en tiempo real para obtener predicciones para registros de datos individuales y utilizarlas en aplicaciones web, móviles o de escritorio interactivas.</p>

Figura 12 Beneficios de Amazon Machine Learning

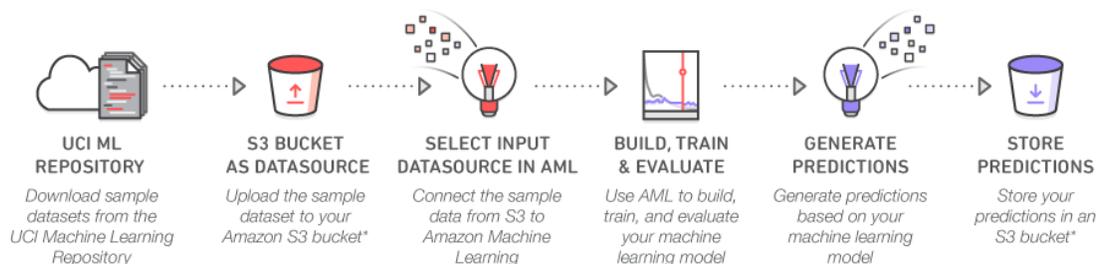
Fuente: *Amazon Machine Learning, Beneficios (2019).*

En cuanto a los casos de uso en los que los servicios de *Amazon Machine learning* se utilizan son:

- Detección del fraude
- Personalización del contenido
- Modelos de propensión para campañas de marketing.
- Clasificación de documentos
- Predicción de renovación de clientes.
- Recomendación de soluciones automática para el servicio de soporte al cliente.

En la Figura 13 muestra un proyecto de ejemplo y su flujo. Esto fue tomado de *Amazon Web Services (2018)*, “Crear un modelo de aprendizaje automático” que logra lo siguiente:

“Este proyecto, descargará datos de clientes de muestra disponibles de manera gratuita y cargará los datos en un *bucket* de *Amazon S3* para crear una fuente de datos. A continuación, creará un modelo de *machine learning* a partir de la fuente de datos, desde la cual luego podrá evaluar y ajustar el desempeño del modelo de *machine learning* y utilizarlo para generar predicciones.



* And don't forget to clean up your S3 bucket after you finish the project to avoid additional storage charges!

Figura 13 Ejemplo de flujo utilizando Amazon Machine Learning

Fuente: *Crear un modelo de aprendizaje automático, Amazon Web Services (2018)*

El ejemplo anterior es una de las muchas formas de utilizar las herramientas de *machine learning* de *Amazon Web Services*, pero para facilitar estas tareas existe el *Amazon SageMaker*, en los siguientes párrafos se amplía más el tema.

2.6.1.3.1.1 Amazon SageMaker

Amazon Web Services (2019), dentro de sus servicios de *machine learning* creo una herramienta llamada *Amazon SageMaker* que permite lo siguiente:

“Para simplificar el *machine learning* para desarrolladores y científicos de datos, AWS ofrece ahora *Amazon SageMaker*, un servicio totalmente administrado con características avanzadas de desarrollo, formación técnica y alojamiento que permite a los desarrolladores centrarse en la creación, formación y ajuste de modelos de *machine learning* sin tener que preocuparse por la infraestructura ni la administración del sistema.

En cuanto a la estructura de *Amazon SageMaker* (2019) cuenta con tres módulos, construcción, capacitación e implementación, que se explican en la siguiente cita:

El módulo de construcción ofrece un entorno alojado para trabajar con sus datos, experimentar con algoritmos y visualizar sus salidas. El módulo de capacitación le permite un modelo de capacitación y ajuste de un clic a alta escala y bajo costo. El módulo de implementación le ofrece un entorno administrado para que aloje y pruebe fácilmente los modelos para lograr una inferencia segura y con baja latencia.

Amazon SageMaker permite una optimización de los *frameworks* y algoritmos como los que se muestran en la Figura 14. Esto lo menciona *Amazon SageMaker* (2018) , en su sitio web específico:

Amazon SageMaker configura y optimiza automáticamente *TensorFlow*, *Apache MXNet*, *PyTorch*, *Chainer*, *Scikit-learn*, *SparkML*, *Horovod*, *Keras* y *Gluon*. Los algoritmos de aprendizaje automático de uso habitual se integran y se adaptan al escalado, velocidad y precisión de más de cien modelos y algoritmos adicionales ya entrenados, disponibles en *AWS Marketplace*. También puede integrar cualquier otro algoritmo o marco de trabajo en un contenedor *Docker* para incorporarlo.



Figura 14 Frameworks que se pueden optimizar, *Amazon SageMaker*.

Fuente: *Amazon SageMaker* (2018).

La plataforma en la nube de *Amazon SageMaker* cuenta con una gran potencia de hardware que permite la optimización y mayor velocidad de procesamiento de los *frameworks* y algoritmos. En la Figura 15 se visualiza la optimización que se logra, gracias a la capacidad de computación que cuenta *Amazon* en la nube.

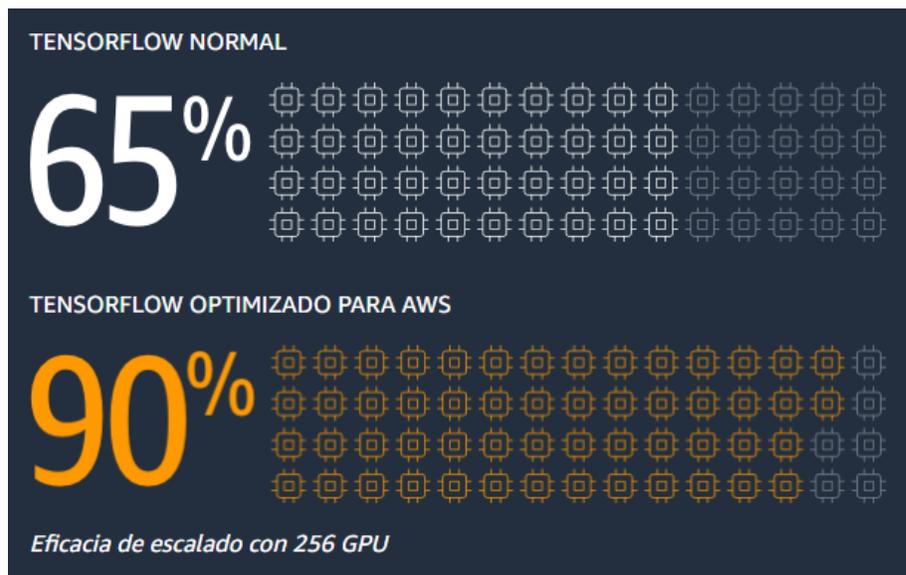


Figura 15 Capacidad de optimización de TensorFlow con Amazon SageMaker.

Fuente: Amazon SageMaker (2018).

2.6.1.3.2 Microsoft Azure Machine Learning

Está dividido en dos, *Azure Machine Learning Studio* y *Azure Machine Learning Services* basados en la nube y con el modelo de infraestructura como servicio.

Estos se desarrollarán en los siguientes apartados para lograr la diferenciación de estos.

2.6.1.3.2.1 Azure Machine Learning Studio.

Según, *Microsoft* define *Azure Machine Learning Studio* (2018):

“Azure Machine Learning Studio le ofrece un espacio de trabajo visual e interactivo para construir, probar e iterar fácilmente en un modelo de análisis predictivo. Arrastra y suelta conjuntos de datos y módulos de análisis en un lienzo interactivo, conectándolos para formar un experimento, que se ejecuta

en Machine Learning Studio. Para repetir en el diseño de su modelo, edite el experimento, guarde una copia si lo desea y vuelva a ejecutarla. Cuando esté listo, puede convertir su experimento de entrenamiento en un experimento predictivo y luego publicarlo como un servicio web para que otros puedan acceder a su modelo.

En la Figura 16, se puede observar la interfaz de la plataforma *Azure Machine Learning Studio* y cómo se generan los modelos por medio del arrastre de figuras ubicadas a la izquierda, para conectarlas entre sí. Esto permite a los desarrolladores crear modelos de manera ágil.

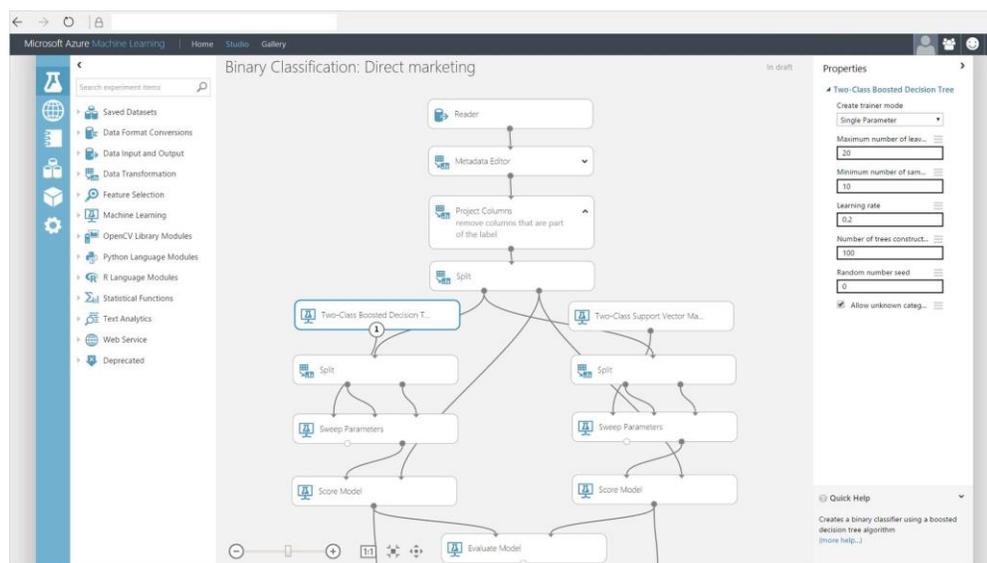


Figura 16 Interfaz de la plataforma *Azure Machine Learning Studio*

Fuente: Microsoft Documentation, (Microsoft, *What is Azure Machine Learning Studio?*, 2018).

Azure Machine Learning Studio permite a los desarrolladores enfocarse en el desarrollo de modelos y la publicación de APIs para operar en la web o generar ganancias en el mercado de API de Microsoft.

Se puede observar dicho flujo en la Figura 17, tomada de la publicación titulada “*Full Microsoft Azure Machine Learning Studio Predictive Analysis Software Review*”, por el sitio Skyoese (2018).

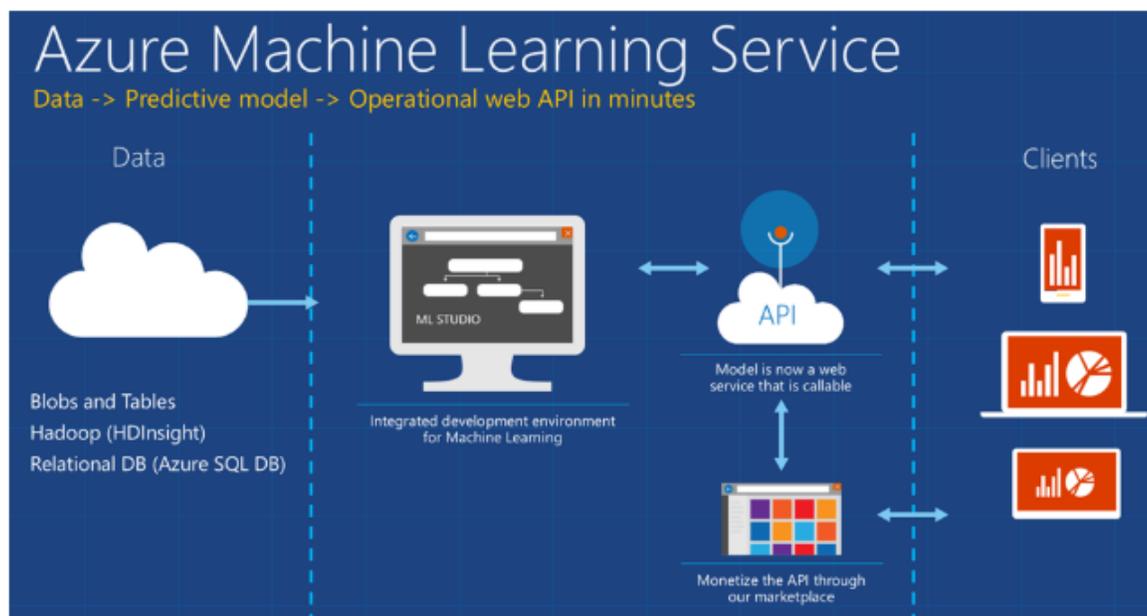


Figura 17 Flujo de trabajo utilizando Azure Machine Learning Studio.

Fuente: *Full Microsoft Azure Machine Learning Studio Predictive Analysis Software Review*, Skyoese (2018).

A modo de exponer de una manera clara los beneficios de *Microsoft Azure Machine Learning Studio*, se crea una tabla de elaboración propia con los beneficios publicados en el sitio *Finances Online Reviews for Business*, que presenta una revisión de la plataforma de Microsoft Azure Machine Learning (2019), esto se puede ver en la Tabla 3.

Beneficio	Descripción
Plataforma interactiva	<p>Azure Machine Learning Studio no es una plataforma común. Hace que el desarrollo de modelos sea más divertido y fácil con su función de arrastrar y soltar que funciona en conjuntos de datos y módulos de análisis. Estos movimientos se utilizan para formar experimentos que se ejecutan en el programa. Editar, guardar una copia, ejecutar, convertir y publicar experimentos también es un proceso sencillo. Debido a sus conexiones visuales, no se requiere una programación pesada y los modelos se completan a un ritmo más rápido.</p>
Soluciones compartidas	<p>Azure Machine Learning Studio reúne a desarrolladores de diferentes lugares en la Galería de Inteligencia de Cortana. Es una sección de la plataforma donde los usuarios pueden compartir y descubrir soluciones que pueden ser de utilidad para otros y para ellos mismos. La comunidad también es donde los usuarios pueden consultar a sus compañeros programadores y dónde pueden trabajar juntos para llegar a una respuesta a un problema.</p>

Despliegue Web	Los modelos diseñados dentro de Azure Machine Learning Studio se pueden enviar directamente a la web, a los que pueden acceder personas que pueden encontrar beneficiosas para sus propios proyectos. Esto permite a los usuarios enviar datos al modelo y recibir predicciones de este, que pueden aplicar a sus propios procesos. Sin embargo, en caso de cambios, los desarrolladores aún pueden actualizar los modelos publicados y administrarlos por medio del portal de servicios web de la plataforma.
Acceso AI	Con Azure Machine Learning Studio, los desarrolladores tienen una biblioteca de algoritmos de inteligencia artificial a su alcance. Pueden usar estas API cognitivas para mejorar su modelado predictivo y obtener resultados más precisos para transformar la forma en que procesan los datos de inteligencia empresarial.

Todo en la nube	Azure Machine Learning Studio es accesible para cualquier persona con una conexión a Internet. No es necesario instalar programas en dispositivos, lo que significa que puede implementarse y utilizarse en cualquier computadora de escritorio o portátil. Esta movilidad permite a los usuarios llevar su trabajo con ellos dondequiera que vayan para que puedan terminar los proyectos a su conveniencia.
-----------------	---

Tabla 3 Beneficios de Azure Machine Learning Studio.

Fuente: *Finances Online Reviews for Business, Microsoft Azure Machine Learning (2019)*

2.6.1.3.2.2 Microsoft Azure Machine Learning Services.

Por otro lado, existe *Microsoft Azure Machine Learning Services* que se define como (2018):

“El servicio *Azure Machine Learning* es un servicio en la nube que puede utilizar para desarrollar e implementar modelos de *machine learning*. Con el servicio *Azure Machine Learning*, puede hacer un seguimiento de sus modelos a medida que los construye, entrena, implementa y administra, todo en la amplia escala que proporciona la nube. El servicio de *machine learning* de *Azure* es totalmente compatible con las tecnologías de código abierto, por lo que puede usar decenas de miles de paquetes de *Python* de código abierto con componentes de aprendizaje automático como *TensorFlow* y *scikit-learn*”.

En la Figura 18, se pueden observar los beneficios que ofrece *Microsoft Azure Machine Learning Services* publicados en el sitio web de *Microsoft* (2019). Entre ellos destaca soporte para *frameworks* como *TensorFlow* y *Scikit-learn*, además el gran soporte para los entornos de desarrollo basados en *Python*.



Aprendizaje automático automatizado

Identificar algoritmos adecuados e hiperparámetros más rápido.



Despliegue simple

Implemente modelos en las instalaciones, en la nube y en el borde con unas pocas líneas de código.



DevOps para aprendizaje automático

Aumente la productividad con el seguimiento de experimentos, la gestión y supervisión de modelos, CI / CD integrado y las líneas de aprendizaje automático.



Cálculo gestionado

Entrene los modelos con facilidad y reduzca los costos al escalar automáticamente los potentes clústeres de GPU.



Python SDK de Tool-agnostic

El servicio de aprendizaje automático de Azure se integra con cualquier entorno de Python, incluidos Visual Studio Code, portátiles Jupyter y PyCharm.



Soporte para frameworks de código abierto.

Use sus herramientas y marcos de trabajo de aprendizaje automático favoritos, como PyTorch, TensorFlow y scikit-learn.

Figura 18 Beneficios de Microsoft Azure Machine Learning Services.

Fuente: Microsoft Azure Machine Learning Services (2019)

Como se puede ver en la Figura 19 tomada de la documentación de *Microsoft*, *Microsoft Documentation*, (2018), el flujo explica que primeramente se prepara la información guardada en alguna base de datos, luego se crea el modelo desde un entorno de desarrollo como puede ser *Jupyter* o *Visual Studio Code*, es decir desde un ordenador o estación personal se puede construir el modelo con el lenguaje de programación *Python*, después este modelo se prueba y se entrena con

los servicios de *Azure Machine Learning*, finalmente se despliega en la *web* para su consumo y monitoreo.

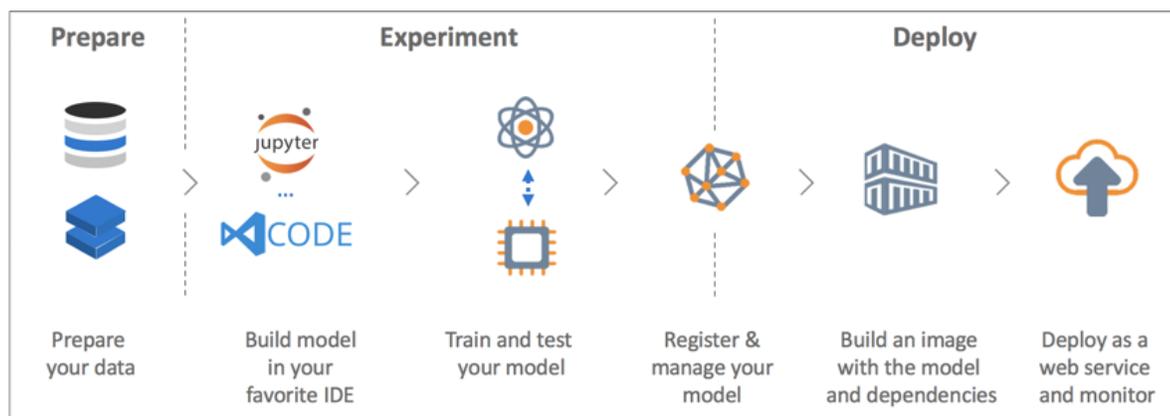


Figura 19 Flujo de trabajo en Azure Machine Learning Services.

Fuente: Microsoft Documentation, (2018)

2.6.1.3.3 Google Cloud Machine Learning.

Google ofrece una serie de servicios enfocados en inteligencias artificial y *machine learning* que los engloba con el nombre de *Cloud AI*, así lo explica en el sitio de Google Cloud (2019):

“*Cloud AI* ofrece servicios modernos de *machine learning*, con modelos pre-entrenados y un servicio para generar sus propios modelos personalizados. Nuestro servicio de ML basado en redes neuronales tiene un mejor rendimiento de entrenamiento y mayor precisión en comparación con otros sistemas de aprendizaje profundo. Nuestros servicios son rápidos, escalables y fáciles de usar. Las principales aplicaciones de Google utilizan el *machine learning* en la nube, incluidas fotos (búsqueda de imágenes),

traductor, bandeja de entrada (respuesta inteligente) y la aplicación de Google (búsqueda por voz).

En la Figura 20, se listan los productos que forman parte de la gama Cloud AI de Google orientada a soluciones de inteligencia artificial y específicamente en *machine learning*.

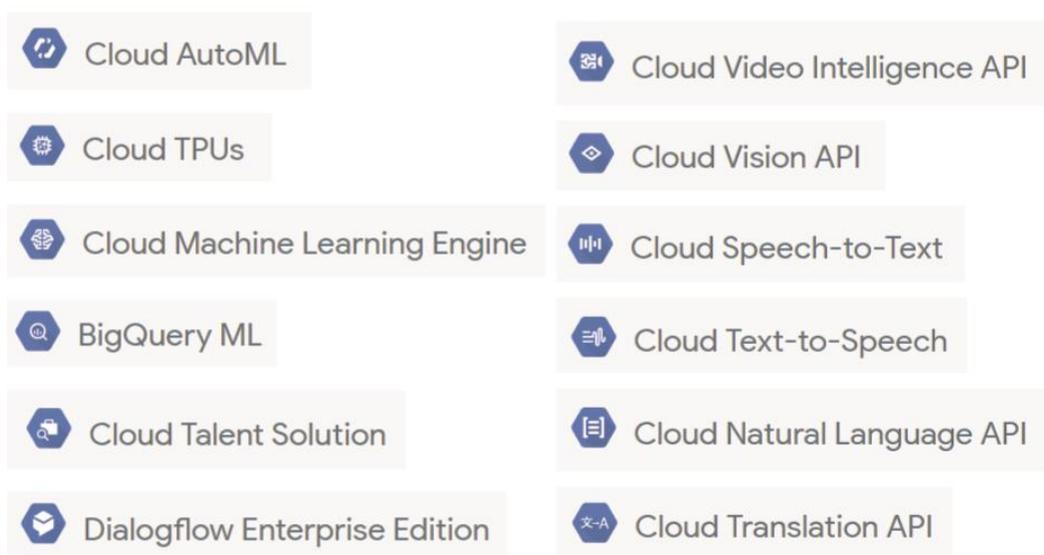


Figura 20 Productos que forman parte de Cloud AI de Google.

Fuente: Google Cloud (2019).

Para ampliar la Figura 20, *Cloud AutoML* es uno de los productos más nuevos que implementa *Google Cloud* (2019), cuenta con funciones muy interesantes, en su documentación menciona lo siguiente:

“*Cloud AutoML* es un paquete de productos de aprendizaje automático que permite a los desarrolladores con poca experiencia en la materia preparar modelos de alta calidad adaptados a sus necesidades empresariales

específicas; todo ello gracias a la tecnología de búsqueda con arquitectura neuronal y al aprendizaje por transferencia de última generación de Google”.

Los paquetes que mencionan en la cita anterior se pueden observar en la Figura 21, entre ellos se encuentran *AutoML Vision* para el reconocimiento de imágenes, *AutoML Natural Language* para la categorización de texto y *AutoML Translation* para generar modelos que permitan traducir idiomas.



Figura 21 Paquetes del Cloud AutoML.

Fuente: *Cloud AutoML, Google Cloud (2019).*

En la Figura 22, se puede observar cómo funciona el *AutoML Vision*, según la publicación en el sitio web *Towards Data Science* por Yufeng G (2018), en este ejemplo se utilizan imágenes y se asignan etiquetas para entrenar, evaluar y por último crear predicciones sobre futuras imágenes e identificar cuando es una silla, una bicicleta o una mesa.

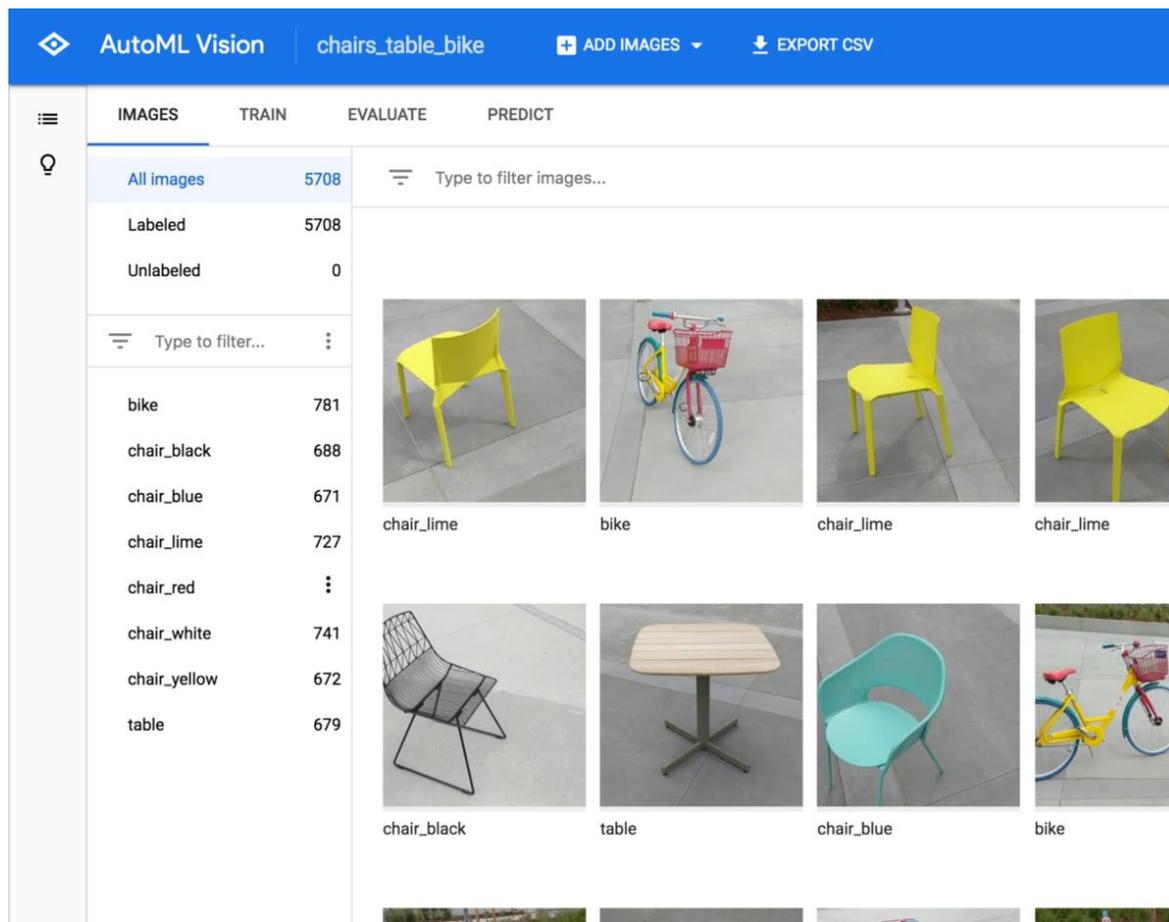


Figura 22 Detección de imágenes con AutoML Vision y su interfaz.

Fuente: *Towards Data Science, Cloud AutoML, (2018).*

En la documentación publicada por Google Cloud (2019), explica cómo funciona el *AutoML Natural Language*, el cual permite analizar textos digitales para clasificar o descubrir el sentimiento oculto.

En la Figura 23 se puede observar cómo el *AutoML Natural Language* aplicado en una reseña de un supermercado ficticio, en el cual se clasifica el texto en categorías como: “Buen servicio”, “Sugerencia” y “Solicitud de información” para

conocer la calidad de servicio. Los números en la imagen representan que tan certero es el modelo sobre las etiquetas correctas.

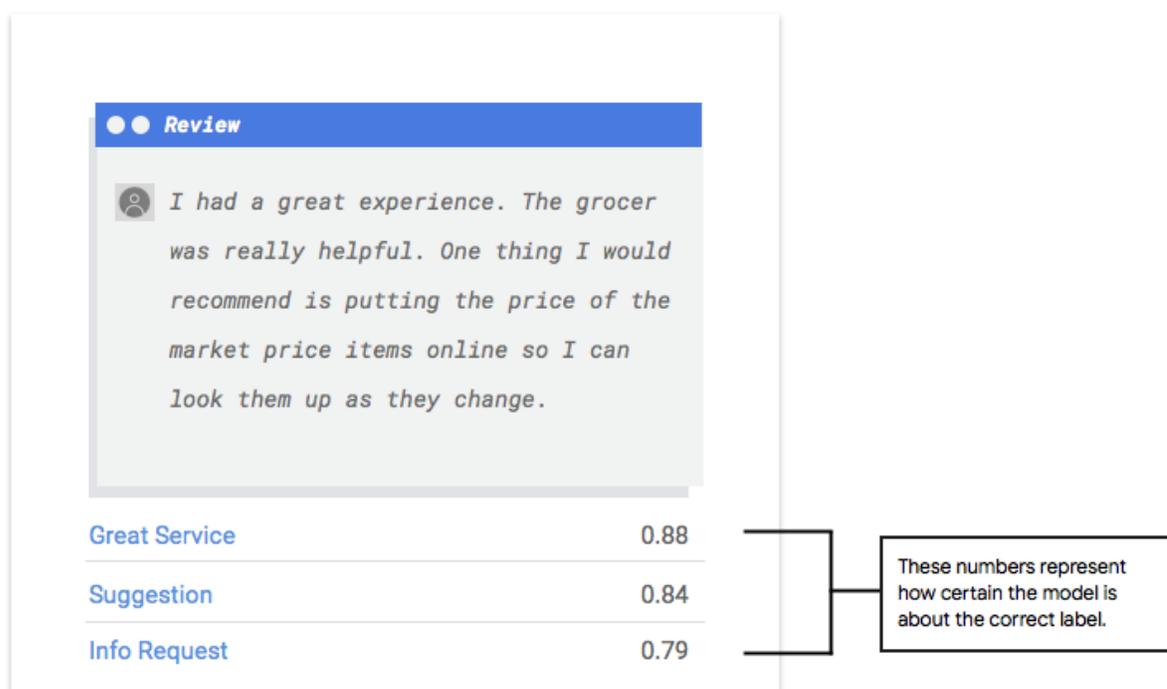


Figura 23 AutoML Natural Language y el análisis de texto.

Fuente: AutoML Natural Language, Google Cloud (2019).

Otro producto que destaca es el *Cloud Machine Learning Engine*, según el sitio *Google Cloud* (2018) lo define como:

“Un servicio administrado que permite que los desarrolladores y científicos de datos creen modelos de *machine learning* superiores para la producción. *Cloud ML Engine* ofrece servicios de entrenamiento y predicción que se pueden usar juntos o individualmente. *Cloud ML Engine* es un servicio comprobado que usan las empresas para resolver problemas que van desde la identificación de las nubes en imágenes satelitales, hasta garantizar la

inocuidad de los alimentos y responder correos electrónicos de los clientes cuatro veces más rápido.

La Figura 24 tomada del sitio web *Google Cloud*; **Error! No se encuentra el origen de la referencia.**, explica que los datos se preparan, luego se codifica el modelo y los servicios de *Google Cloud* se encargan de entrenar, evaluar y mejorar el modelo, además de permitir el despliegue en la nube para obtener predicciones, monitoreo y manejo de versión de los modelos creados.

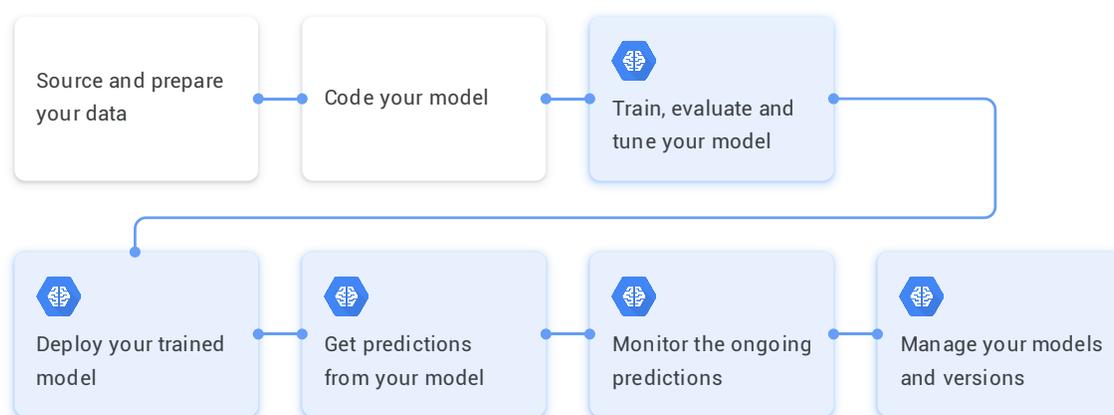


Figura 24 Flujo de trabajo de alto nivel de Google Cloud Machine Learning.

Fuente: Google Cloud, (2018)

Además, *Google Cloud Machine Learning* permite el uso de diversos *frameworks* para *machine learning*, como *TensorFlow*, *Keras*, *XGBoost* y *scikit-learn* como se puede ver en la Figura 25.



Figura 25 Frameworks más utilizados según Google Cloud para los servicios de machine learning.

Fuente: *Google Cloud Machine Learning (2018).*

2.6.1.3.4 IBM Watson Machine Learning

Según, el sitio de *IBM* (2018) define *IBM Watson Studio*:

“IBM Watson Studio es un entorno de colaboración con herramientas de inteligencia artificial que usted y su equipo pueden usar para recopilar y preparar datos de capacitación, y para diseñar, capacitar e implementar modelos de aprendizaje automático”.

Para una mejor lectura de los beneficios de *IBM Watson Machine Learning* se utiliza la Tabla 4 que contiene información tomada de la publicada en el sitio web de *IBM* (2019).

Beneficio	Descripción
Escalable	<i>Watson Machine Learning</i> hace que sea fácil y rentable implementar los activos de <i>machine learning</i> e inteligencia artificial en entornos públicos, privados, híbridos o de múltiples nubes. Amplia sin problemas las iniciativas de inteligencia artificial, convirtiendo proyectos piloto en implementaciones empresariales críticas para el negocio sin grandes inversiones iniciales.
Velocidad	Genere activos de AI para comercializar más rápido al agilizar el proceso de implementación y entrenamiento del modelo. <i>Watson Machine Learning</i> ; automatiza muchos aspectos de la capacitación de modelos, mientras que las optimizaciones de hardware multiplataforma aceleran los programas de capacitación al maximizar la utilización de recursos.
Sencillez	Reduce la curva de aprendizaje aprovechando una gran cantidad de modelos pre-entrenados y conjuntos de datos abiertos. Simplifique la administración del ciclo de vida con monitoreo de desempeño automatizado y realimentación continua, e interoperar fácilmente con otras herramientas de ciencia de datos con una arquitectura abierta y modular.

Tabla 4 Beneficios de Watson Machine Learning

Fuente: IBM Watson Machine Learning, Beneficios (2019).

IBM Watson Studio utiliza una versión web como se puede notar en la Figura 26 tomada del tutorial que publicó IBM (2018) en su sitio web , aunque posee una

versión de escritorio o local que permite crear y entrenar modelos de *machine learning* con tan solo unos “clicks”, en este ejemplo se analiza la información que posee el archivo “*customer_churn.csv*” para conocer ciertas características sobre los clientes.

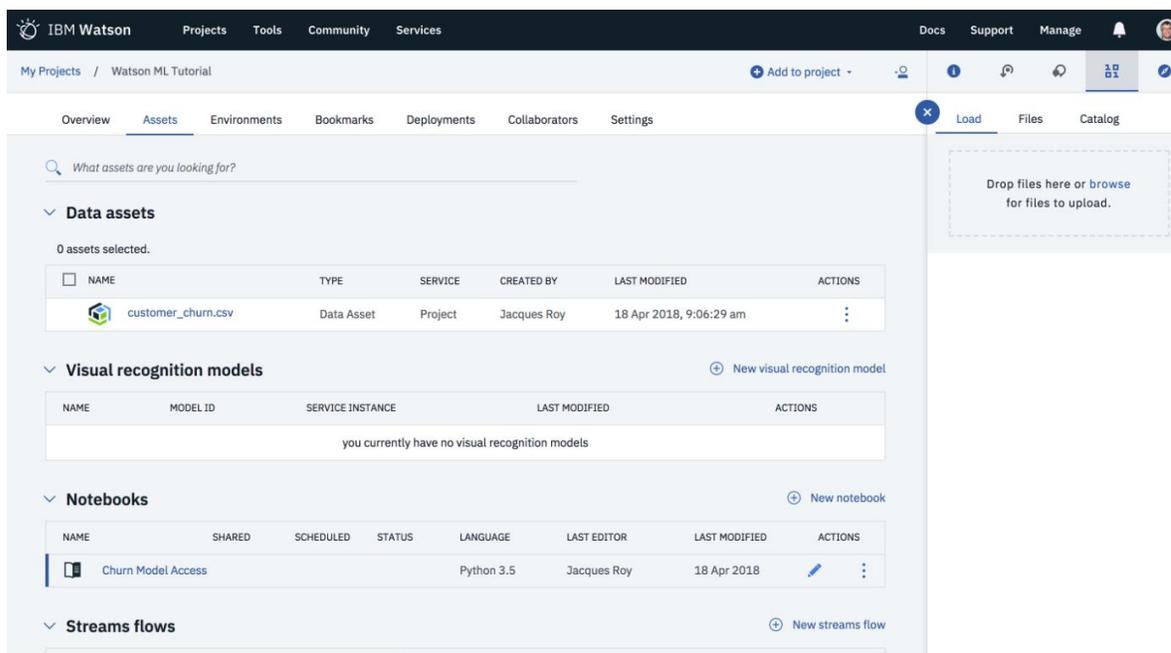


Figura 26 Interfaz de IBM Watson Machine Learning.

Fuente: Tutorial de IBM Watson Machine Learning, (2018).

Además, posee técnicas de *machine learning* predefinidas para realizar un entrenamiento de los modelos. En la Figura 27 se puede observar técnicas como clasificación binaria, clasificación multiclase y regresión.

Esta imagen fue tomada de un tutorial interactivo de *IBM Watson* (2018).

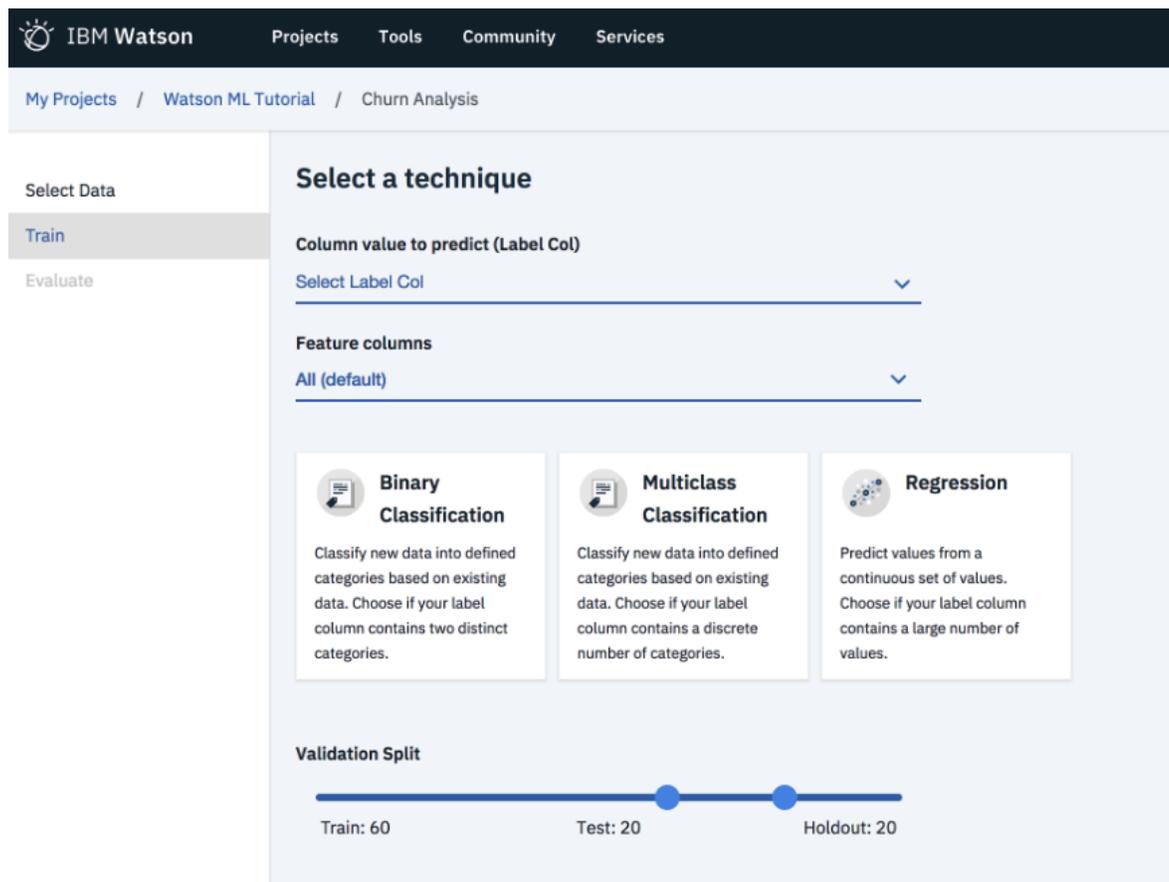


Figura 27 Técnicas de Machine Learning en Watson Machine Learning.

Fuente: Tutorial de IBM Watson Machine Learning, (2018).

2.6.1.3.5 Comparación de los *frameworks* de machine learning basados en la nube.

Para tener un mejor concepto de lo que ofrecen los *frameworks* basados en la nube se comparan los tipos de servicios esto, según el sitio web *AltexSoft*.

Podemos ver en la Figura 28, que de la comparativa uno de los *frameworks* con características más variadas, destaca *Microsoft Azure ML Studio*. Cubre la mayoría de las tareas relacionadas con el *machine learning*, proporciona dos productos distintos para construir modelos personalizados.

CLOUD MACHINE LEARNING SERVICES COMPARISON				
	Amazon	Microsoft	Google	IBM
Automated and semi-automated ML services				
	Amazon ML	Microsoft Azure ML Studio	Google Prediction API	IBM Watson ML Model Builder
Classification	✓	✓	deprecated	✓
Regression	✓	✓		✓
Clustering	✓	✓		✗
Anomaly detection	✗	✓		✗
Recommendation	✗	✓		✗
Ranking	✗	✓		✗
Platforms for custom modeling				
	Amazon SageMaker	Azure ML Services	Google ML Engine	IBM Watson ML Studio
Built-in algorithms	✓	✗	✗	✓
Supported frameworks	TensorFlow, MXNet, Keras, Gluon, PyTorch, Caffe2, Chainer, Torch	TensorFlow, scikit-learn, Microsoft Cognitive Toolkit, Spark ML	TensorFlow, scikit-learn, XGBoost, Keras	TensorFlow, Spark MLlib, scikit-learn, XGBoost, PyTorch, IBM SPSS, PMML

Figura 28 Comparación de frameworks basados en la nube.

Fuente: AltexSoft, Comparación del aprendizaje automático como un servicio: Amazon, Microsoft Azure, Google Cloud AI, IBM Watson (2018).

2.6.2 Estilos de machine learning.

Según, el problema establecido es necesario utilizar un estilo de *machine learning*, entre ellos se encuentran el estilo supervisado, no supervisado y por refuerzo.

2.6.2.1 Estilo supervisado

Según, un artículo de Javier Luna (2018) publicado en el sitio web *Medium*, explica el estilo supervisado de *machine learning* y el concepto de supervisado radica en que los datos se etiquetan y se ordenan para un mejor análisis:

“En los algoritmos de aprendizaje supervisado se genera un modelo predictivo, basado en datos de entrada y salida. La palabra clave “supervisado” viene de la idea de tener un conjunto de datos previamente etiquetado y clasificado, es decir, tener un conjunto de muestra el cual ya se sabe a qué grupo, valor o categoría pertenecen los ejemplos. Con este grupo de datos, el cual llamamos datos de entrenamiento, se realiza el ajuste al modelo inicial planteado. De esta forma es como el algoritmo va “aprendiendo” a clasificar las muestras de entrada comparando el resultado del modelo, y la etiqueta real de la muestra, realizando las compensaciones respectivas al modelo de acuerdo con cada error en la estimación del resultado. Por ejemplo, el aprendizaje supervisado ha sido utilizado para la programación de vehículos autónomos”.

Podemos ver con más claridad cómo funciona el estilo supervisado en la Figura 29, tomada del artículo de Javier Luna (2018). Las etiquetas ordenan los datos (textos de entrenamiento, documentos, imágenes) para que después de aplicar el algoritmo de *machine learning* tengan una salida etiquetada y fácil de entender.

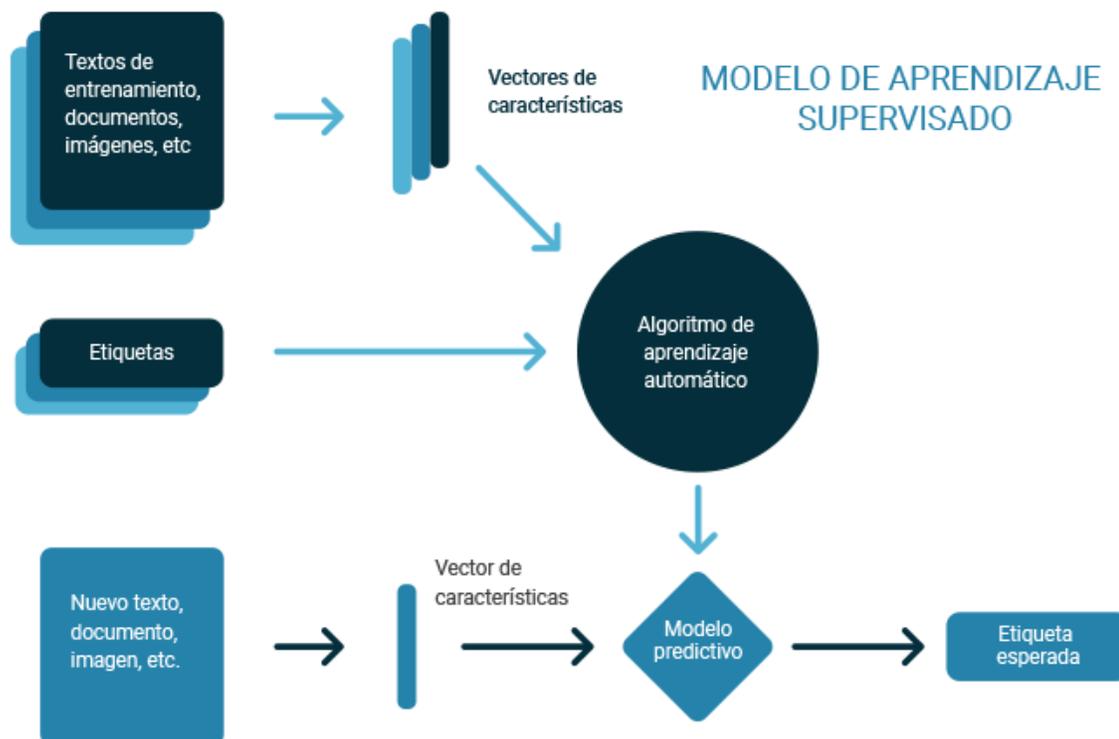


Figura 29 Diagrama de flujo del estilo supervisado.

Fuente: Tipos de aprendizaje automático, Javier Luna (2018).

2.6.2.2 Estilo no supervisado

En el artículo de Javier Luna (2018), explica que hay una similitud entre el estilo supervisado y no supervisado:

“Los algoritmos de aprendizaje no supervisado trabajan de forma muy similar a los supervisados, con la diferencia de que estos solo ajustan su modelo predictivo tomando en cuenta los datos de entrada, sin importar los de salida. Es decir, a diferencia del supervisado, los datos de entrada no están clasificados ni etiquetados, y no son necesarias estas características para

entrenar el modelo. Dentro de este tipo de algoritmos, el agrupamiento o *clustering* en inglés, es el más utilizado, pues particiona los datos en grupos que posean características similares entre sí. Una aplicación de estos métodos es la compresión de imágenes.

En la Figura 30 tomada del artículo de Javier Luna (2018), podemos notar que a los datos (Textos de entrenamiento, documentos, imágenes) no se les asigna una etiqueta.

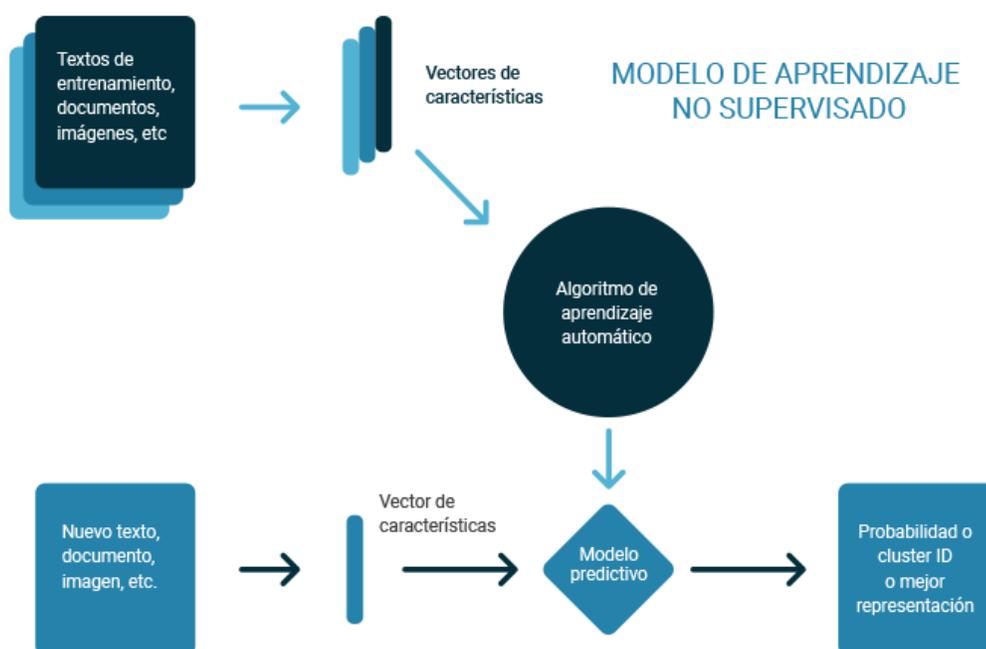


Figura 30 Diagrama de flujo del estilo no supervisado.

Fuente: Tipos de aprendizaje automático, Javier Luna (2018).

2.6.2.3 Estilo por refuerzo

En el artículo de Javier Luna (2018), explica cómo se define el estilo por refuerzo:

“Los algoritmos de aprendizaje por refuerzo definen modelos y funciones enfocadas en maximizar una medida de “recompensas”, basados en “acciones” y al ambiente en el que el agente inteligente se desempeñará. Este algoritmo es el más apegado a la psicología conductista de los humanos, pues es un modelo acción-recompensa, que busca que el algoritmo se ajuste a la mejor “recompensa” dada por el ambiente, y sus acciones por tomar están sujetas a estas recompensas. Este tipo de métodos pueden usarse para hacer que los *robots* aprendan a realizar diferentes tareas”.

En la Figura 31 tomada del artículo de Javier Luna (2018), se observa el ciclo que está compuesto por un ambiente y un agente seleccionador de algoritmo que permiten una mejora en las salidas por medio de las recompensas.

MODELO DE APRENDIZAJE POR REFUERZO



Figura 31 Diagrama de flujo del estilo por refuerzo.

Fuente: Tipos de aprendizaje automático, Javier Luna (2018).

2.6.3 Tipos de enfoques de *machine learning*.

Dentro de los estilos de *machine learning* se encuentran los tipos de enfoques según el problema, entre ellos se encuentran los siguientes:

- Clasificación
- Predicción

En los siguientes apartados se explicarán con más detalle los enfoques listados.

2.6.3.1 Clasificación

Según, Sidath Asiri, Licenciado en el Departamento de Ciencias de la Computación e Ingeniería de la Universidad de Moratuwa (2018), presenta una definición sobre los algoritmos de clasificación para *machine learning*:

“La clasificación es el proceso de predecir la clase que poseen datos dados. Las clases se llaman a veces como objetivos / etiquetas o categorías. El modelado predictivo de clasificación es la tarea de aproximar una función de mapeo (f) de las variables de entrada (X) a las variables de salida discretas (y)”.

Para aclarar la definición presenta el siguiente ejemplo, un correo electrónico de texto se puede clasificar como perteneciente a una de dos clases: "*spam*" y "*no spam*".

Según, Jason Brownlee, con un doctorado en inteligencia artificial (2017), menciona en su artículo “Diferencia entre clasificación y regresión en el aprendizaje automático”, las características de un problema de clasificación:

- Un problema de clasificación requiere que los ejemplos se clasifiquen en una de dos o más clases.
- Una clasificación puede tener variables de entrada de valores reales o discretas.

- Un problema con dos clases a menudo se llama un problema de clasificación binario o de dos clases.
- Un problema con más de dos clases a menudo se llama un problema de clasificación de múltiples clases.
- Un problema en el que se asignan varias clases a un ejemplo se denomina problema de clasificación de etiquetas múltiples.

2.6.3.2 Predicción

El modelo predictivo, también llamado modelo de regresión predictiva es uno de los tipos de enfoques para resolver problemas con *machine learning*.

Jason Brownlee (2017) explica cómo funciona el modelo de regresión:

“El modelo predictivo de regresión es la tarea de aproximar una función de mapeo (f) de las variables de entrada (X) a una variable de salida continua (y). Una variable de salida continua es un valor real, como un valor entero o de punto flotante. Estos son a menudo cantidades, tales como cantidades y tamaños”.

A modo ejemplo para entender la definición anterior: “...se puede predecir que una casa se venderá por un valor específico en dólares, quizás en el rango de \$ 100,000 a \$ 200,000”.

Para tener clara la diferencia entre los dos tipos de enfoques Alexandre Drouin (2017) en su artículo “*Microbiome Summer School 2017, Introduction to*

Machine Learning”, en la Figura 32, donde se puede observar un ejemplo en la ciencia médica se utiliza el enfoque clasificatorio para categorizar los pacientes (enfermo o saludable) y por medio del enfoque regresivo para predecir la cantidad de años que puede sobrevivir un paciente a una enfermedad.

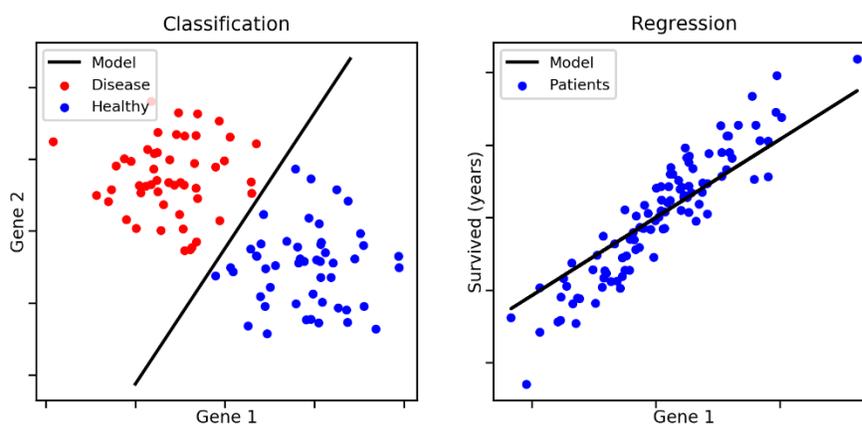


Figura 32 Diferencia entre el enfoque clasificatorio y regresivo en la ciencia médica.

Fuente: Microbiome Summer School 2017, Introduction to Machine Learning, *Alexandre Drouin (2017)*

Estos enfoques contienen métodos de *machine learning* que serán explicados en los párrafos posteriores para una mayor comprensión.

2.6.4 Métodos de machine learning.

Los métodos de *machine learning* son esencialmente algoritmos, los algoritmos, según Stone y Knuth (1981) de la Universidad de Stanford lo definen como: “conjunto de reglas que define con precisión una secuencia de operaciones de modo que cada regla sea efectiva y definida y la secuencia termine en un tiempo finito.” (pág. 8)

Los algoritmos de *machine learning* son en conjunto de métodos que permiten obtener un conocimiento o experiencia a partir de los datos, esto lo afirma Yedukondalu Golla (2018), los algoritmos de *machine learning* son: "...aquellos que pueden aprender de los datos y mejorar de la experiencia, sin intervención humana".

Después los algoritmos de *machine learning* se pueden agrupar, según la funcionalidad, Jason Brownlee (2013): "Los algoritmos a menudo se agrupan por similitud en términos de su función (cómo funcionan). Por ejemplo, los métodos basados en árboles y los métodos inspirados en redes neuronales."

Hay muchos algoritmos utilizados en *machine learning* que no serán parte de este marco teórico, para ello se utiliza el Anexo 8 Mapa de algoritmos de *machine learning* que brinda una lista más robusta.

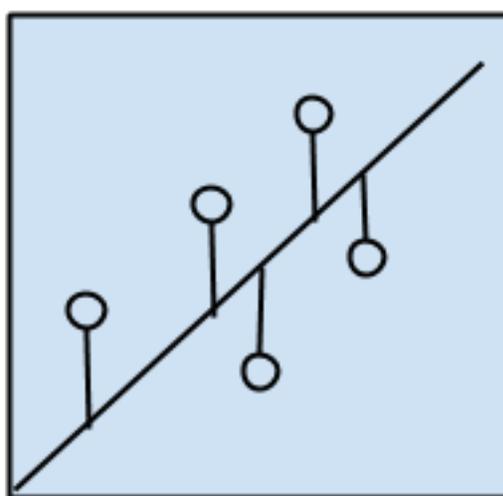
2.6.4.1 Algoritmos de regresión

Para definir qué es un algoritmo de regresión, Jason Brownlee (2013), explica que:

"La regresión tiene que ver con modelar la relación entre variables que se refina iterativamente utilizando una medida de error en las predicciones hechas por el modelo. Los métodos de regresión son un caballo de batalla de las estadísticas y se han incorporado al aprendizaje estadístico de máquinas. Esto puede ser confuso porque podemos usar la regresión para

referirnos a la clase de problema y la clase de algoritmo. Realmente, la regresión es un proceso”.

Como referencia la Figura 33, tomada del artículo “Un recorrido por los algoritmos de aprendizaje automático” de Jason Brownlee (2013), representa el funcionamiento de la regresión.



Regression Algorithms

Figura 33 Ilustración abstracta de cómo funciona un algoritmo de regresión.

Fuente: Jason Brownlee (2013) en su artículo “Un recorrido por los algoritmos de aprendizaje automático”

Los algoritmos de regresión más populares, según Jason Brownlee (2013) son :

- Regresión de mínimos cuadrados ordinarios (OLSR)
- Regresión lineal

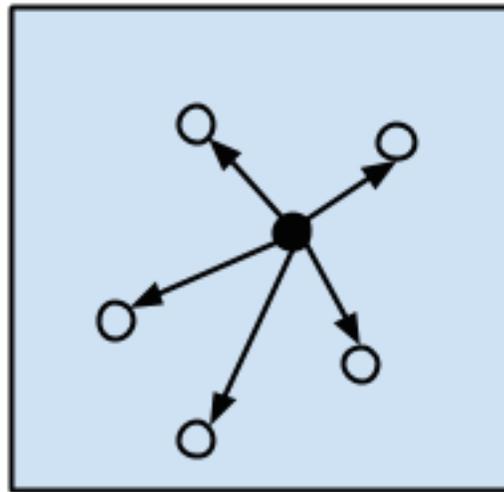
- Regresión logística
- Regresión escalonada
- *Splines* de regresión adaptativa multivariante (MARS)
- Suavizado de gráfico de dispersión estimado localmente (LOESS)

2.6.4.2 Algoritmos basados en instancias.

Los algoritmos basados en instancias, según Jason Brownlee (2013) en su artículo “Un recorrido por los algoritmos de aprendizaje automático” explica que:

“Los métodos generalmente construyen una base de datos de referencia y comparan datos nuevos con la base de datos utilizando una medida de similitud para encontrar la mejor coincidencia y hacer una predicción. Por esta razón, los métodos basados en instancias también se denominan métodos ganadores para todos y aprendizaje basado en la memoria. Se enfoca en la representación de las instancias almacenadas y las medidas de similitud utilizadas entre las instancias.

Según, la Figura 34, tomada del artículo “Un recorrido por los algoritmos de aprendizaje automático” de Jason Brownlee (2013), representa el funcionamiento de un algoritmo basado en instancias, donde se puede observar como el punto negro que representa a la base datos de referencia es comparada con otra base de datos.



Instance-based
Algorithms

Figura 34 Representación del funcionamiento de un algoritmo basado en instancias

Fuente: Jason Brownlee (2013) en su artículo "Un recorrido por los algoritmos de aprendizaje automático"

Los algoritmos basados en instancias más populares, según Jason Brownlee (2013) son:

- *k-Nearest Neighbor (kNN)*
- *Learning Vector Quantization (LVQ)*
- *Self-Organizing Map (SOM)*
- *Locally Weighted Learning (LWL)*

2.6.4.3 Algoritmos de árboles de decisión.

Uno de los algoritmos más populares es el de árbol de decisión, Rahul Saxena (2017), explica cómo funciona:

“El nivel de comprensión del algoritmo de árboles de decisión es muy fácil en comparación con otros algoritmos de clasificación. El algoritmo del árbol de decisión trata de resolver el problema utilizando la representación del árbol. Cada nodo interno del árbol corresponde a un atributo, y cada nodo hoja corresponde a una etiqueta de clase.

En la Figura 35, del artículo *“How decision tree algorithm Works”* de Rahul Saxena (2017), explica con esta representación gráfica cómo funciona un árbol de decisiones, donde se coloca el primer nodo como la raíz del árbol. En este caso el ejemplo es para decidir si se acepta un nuevo puesto de trabajo.

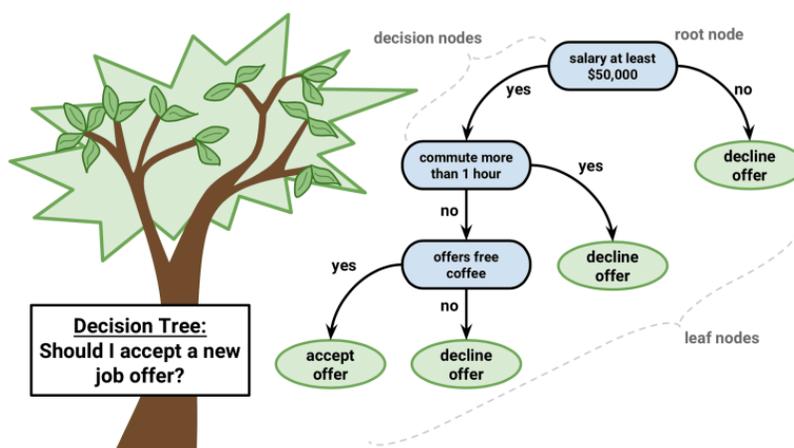


Figura 35 Representación de un algoritmo de árbol de decisiones.

Fuente: *How decision tree algorithm works* Rahul Saxena (2017)

Los algoritmos de árbol de decisiones más populares, según Jason Brownlee (2013) son:

- *Classification and Regression Tree (CART)*
- *Iterative Dichotomiser 3 (ID3)*
- *C4.5 and C5.0 (different versions of a powerful approach)*
- *Chi-squared Automatic Interaction Detection (CHAID)*
- *Decision Stump*
- *M5*
- *Conditional Decision Trees*

2.6.4.4 Algoritmos Bayesianos

El método Bayesiano es el que explícitamente aplica el teorema de Bayes para problemas de clasificación y regresión.

Según, el sitio Brilliant.org (2018) define el teorema de Bayes como:

“Una fórmula que describe cómo actualizar las probabilidades de las hipótesis cuando se presentan pruebas. Sigue simplemente de los axiomas de la probabilidad condicional, pero puede usarse para razonar poderosamente sobre una amplia gama de problemas que involucran actualizaciones de creencias”.

En la Figura 36 tomada del artículo *Bayes' Theorem and Conditional Probability*, Brilliant.org (2018), explica que:

Dada la hipótesis H y la evidencia E, el teorema de Bayes establece que la relación entre la probabilidad de la hipótesis antes de obtener la evidencia P (H) y la probabilidad de la hipótesis después de obtener la evidencia P (H | E).

$$P(H | E) = \frac{P(E | H)}{P(E)} P(H).$$

Figura 36 Fórmula del teorema de Bayes.

Fuente: *Bayes' Theorem and Conditional Probability, Brilliant.org (2018)*

Los algoritmos Bayesianos más populares, según Jason Brownlee (2013) son:

- *Naive Bayes*
- *Gaussian Naive Bayes*
- *Multinomial Naive Bayes*
- *Averaged One-Dependence Estimators (AODE)*
- *Bayesian Belief Network (BBN)*
- *Bayesian Network (BN)*

2.6.4.5 Algoritmos de agrupamiento

Los algoritmos de agrupamiento se pueden definir, según Jason Brownlee (2013) como:

“Los métodos de agrupación en clúster suelen organizarse, según los enfoques de modelado, como los basados en centroides y jerárquicos. Todos los métodos están relacionados con el uso de las estructuras inherentes en los datos para organizar mejor los datos en grupos de máxima similitud.

La

Figura 37 de Jason Brownlee (2013) en su artículo “Un recorrido por los algoritmos de aprendizaje automático”, representa el concepto de agrupamiento.

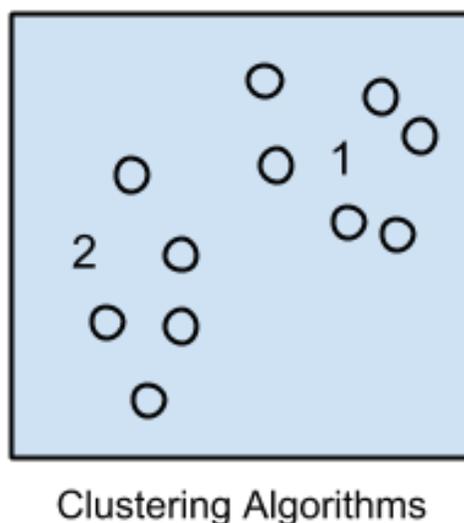


Figura 37 Representación del funcionamiento del algoritmo de agrupamiento.

Fuente: Jason Brownlee (2013) en su artículo “Un recorrido por los algoritmos de aprendizaje automático”

Los algoritmos Bayesianos más populares, según Jason Brownlee (2013) son:

- *k-Means*
- *k-Medians*
- *Expectation Maximisation (EM)*
- *Hierarchical Clustering*

2.6.4.6 Deep learning, Redes neuronales.

El *deep learning* una técnica de *machine learning* que enseña a los sistemas computacionales a hacer lo que resulta natural para las personas: aprender mediante ejemplos.

Para explicar cómo funciona se toma como referencia la empresa *MathWorks* (2018):

“La mayor parte de los métodos de aprendizaje emplean arquitecturas de redes neuronales, por lo que, a menudo, los modelos de *deep learning* se denominan redes neuronales profundas. El término “profundo” hace referencia al número de capas ocultas en la red neuronal. Las redes neuronales tradicionales solo contienen dos o tres capas ocultas, mientras que las redes profundas pueden tener hasta 150”.

Como se puede apreciar en la Figura 38, tomada del sitio MathWorks (2018) las capas ocultas son las que hace referencia al término “profundo” de una red neuronal y son las encargadas de combinarse para dar un resultado.

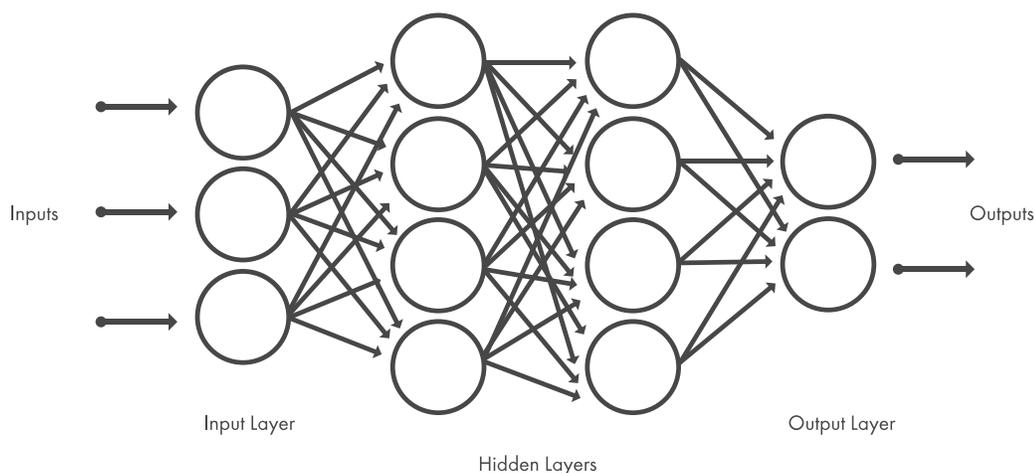


Figura 38 Redes neuronales organizadas en capas que constan de una serie de nodos interconectados. Las redes pueden tener decenas o cientos de capas ocultas.

Fuente: *Deep learning, MathWorks (2018)*

Con una definición más técnica, Daniel Graupe (2016) define *deep learning*:

Las redes neuronales de *deep learning* (DLNN) pueden ser definidas como arquitecturas de redes neuronales que pueden facilitar el *deep learning*, la recuperación y el análisis de datos que están profundamente guardados en la información de entrada y que no son fácilmente recuperables.

Su habilidad para profundizar en los datos de entrada es superior o más rápida a otros métodos computacionales debido a su eficiente integración de

varios métodos matemáticos, lógicos y computacionales, lineales, no lineales, analíticos o heurísticos, determinísticos o estocásticos a una tarea determinada. (p.1)

Para destacar su importancia se toma un periódico digital, El Economista (2017), en su publicación titulada, “El *deep learning* será la próxima revolución tecnológica”, *MBIT School*, único centro de formación español dedicado exclusivamente a *Business Intelligence* y *Big data*, ha celebrado una jornada sobre *Deep learning* en la que ha contado con la participación de Intel, se menciona:

“Se espera una gran demanda de este tipo de procesos de computación que se multiplicarán por 12 de ahora a 2020, según los expertos. En concreto, tres factores han permitido esta eclosión, que son la inundación de datos, la computación y capacidad de integración junto con la innovación, que ha pasado de las empresas al consumo con herramientas como el reconocimiento de voz o de imagen.

El *deep learning* es una tendencia creciente por el poder computacional, las grandes bases de datos y la gran cantidad de usos que se le puede dar para la solución a muchos problemas de la sociedad.

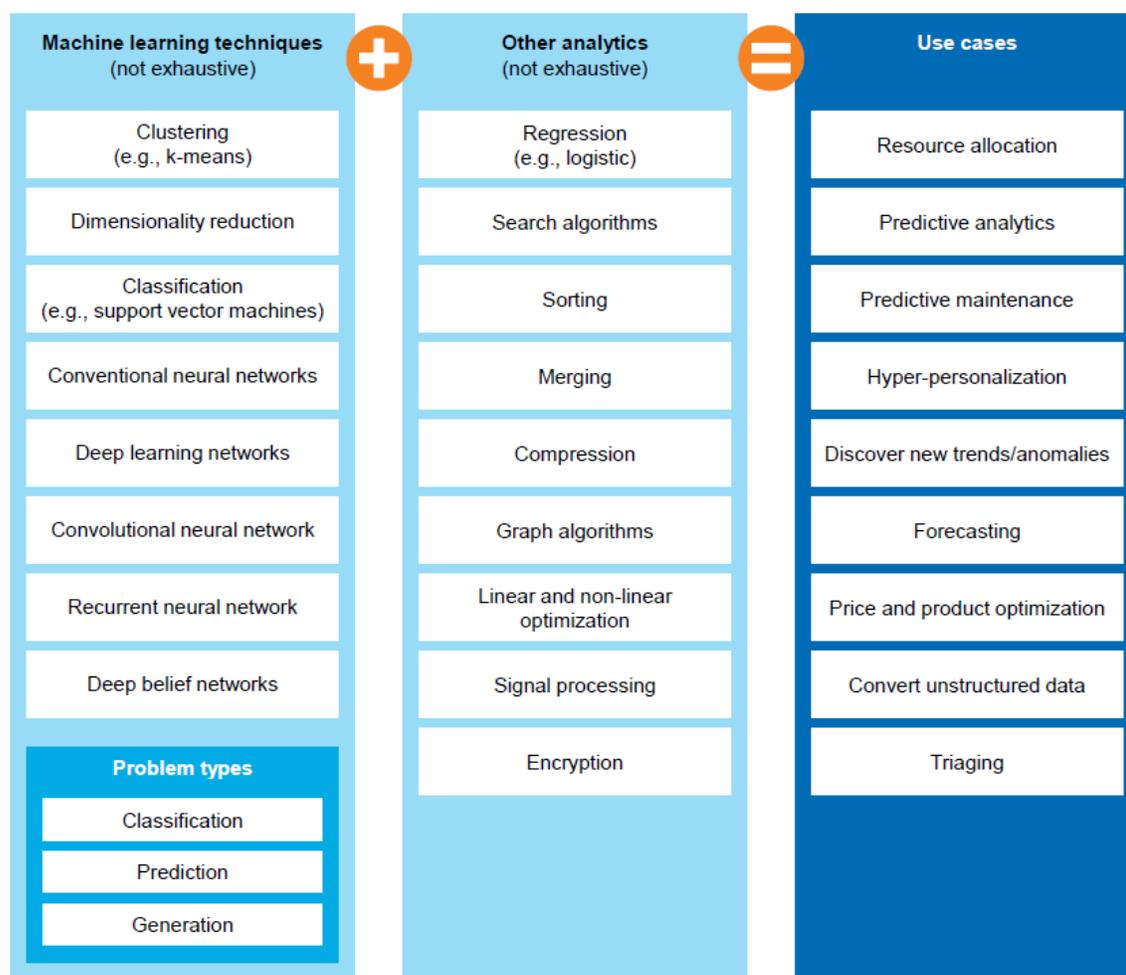
Los algoritmos de *machine learning* tienen múltiples aplicaciones por lo que en el siguiente apartado se explican diferentes casos usos en el nivel comercial.

2.6.5 Casos de uso específico de *machine learning*.

Los casos de usos específicos tienen que ver con la aplicación del *machine learning* y qué problemas resuelven.

La Figura 39 del reporte de *McKinsey Global Institute*, en su reporte "*The age of analytics*", muestra como la mezcla de métodos y técnicas generan un caso de uso como: Asignación de recursos, análisis predictivo, mantenimiento predictivo, hiper-personalización, descubrir nuevas tendencias o anomalías, pronósticos, optimización de precios y productos, convertir datos no estructurados y agrupamiento de pacientes, según enfermedad.

Machine learning can be combined with other types of analytics to solve a large swath of business problems



SOURCE: McKinsey Global Institute analysis

Figura 39 La mezcla de técnicas pueden resolver un caso de uso.

Fuente: MGI, *The age of analytics report (2016)*.

2.6.5.1 Chatbots una forma de conversar con sus clientes.

Uno de los casos de uso más importante hoy, son los *chatbots* utilizados en atención personalizada de clientes.

Según, Gilbert Mizrahi CEO de *Baloka* (2018): “Un *chatbot* es simplemente un *software* que permite comunicarte con tus clientes dentro de las aplicaciones de mensajería.”

Es decir un *chatbot* es un software desarrollado para automatizar la comunicación entre un sistema y el ser humano, para generar eficiencia en el flujo del servicio.

Para profundizar en el funcionamiento de un *chatbot* se toma de referencia el autor Fernández (2015), quien explica ciertos términos muy utilizados para el desarrollo de proyectos de este tipo.

Nombre	Definición
PNL, o el procesamiento del lenguaje natural.	PNL, o el procesamiento del lenguaje natural es un término general utilizado para describir la capacidad de una máquina para ingerir lo que se le dice, descomponerlo, comprender su significado, determinar la acción adecuada y responder en un idioma que el usuario pueda entender.
NLU, o <i>Natural Language Understanding</i> (Entendimiento de	NLU, o <i>Natural Language Understanding</i> es un subconjunto de

<p>lenguaje natural) es un subconjunto de NLP</p>	<p>NLP que trata con la faceta mucho más estrecha, pero igualmente importante, de cómo manejar mejor las entradas no estructuradas y convertirlas en una forma estructurada que una máquina puede entender y actuar. Mientras que los humanos son capaces de manejar sin esfuerzo las malas palabras, palabras intercambiadas, contracciones, coloquialismos y otras peculiaridades, las máquinas no son tan expertas en comprender y asimilara aquellos <i>inputs</i> con faltas de ortografía u otros supuestos.</p>
<p>NLG, o <i>Natural Language Generation</i> (Generación de lenguaje natural)</p>	<p>NLG, o <i>Natural Language Generation</i>, en pocas palabras, es lo que sucede cuando las computadoras escriben el lenguaje. Los procesos NLG convierten datos estructurados en texto.</p>

Tabla 5 Definición de NLP, NLU y NLG.

Fuente: Definición de NLP, NLU, NLG y cómo funcionan los Chatbots (Fernandes, 2015)

Para tener más claro el funcionamiento de un *chatbot* se plantea el siguiente ejemplo de elaboración propia:

“Imaginemos que un jefe solicita comprar el mejor computador de Costa Rica, surgen muchas preguntas, ¿Cuántos computadores necesita?, ¿Cuánta capacidad de memoria?, ¿De qué color? El primer paso es buscar distribuidores para saber dónde hay disponible, vuelves donde el jefe y se exponen los computadores encontrados y que valor tienen. Después de solicitar el dinero se compran los computadores seleccionados”.

Los *chatbots* funcionan de una manera similar, pero con diferente canal de comunicación. Los procesos de NLP convierten el texto en datos estructurados, la máquina convierte esta solicitud de texto sin formato en comandos codificados para sí mismo. El *chatbot* utiliza la información para tomar decisiones cumpliendo ciertos criterios para salir del ciclo de conversación.

Los procesos de *Natural Language Generation* (NLG) convierten los datos estructurados en texto con el fin de generar un lenguaje que un ser humano pueda entender.

Como se puede ver en la Figura 40, la capa de *machine learning* se compone de un bloque de NLP, NLU y un motor de decisiones, según Fernandes (2015). Esta

capa de *machine learning* es el traductor de lenguaje natural y permite al programa entender lo que un cliente solicita por medio de una capa de presentación como puede ser una web, Facebook o Slack.

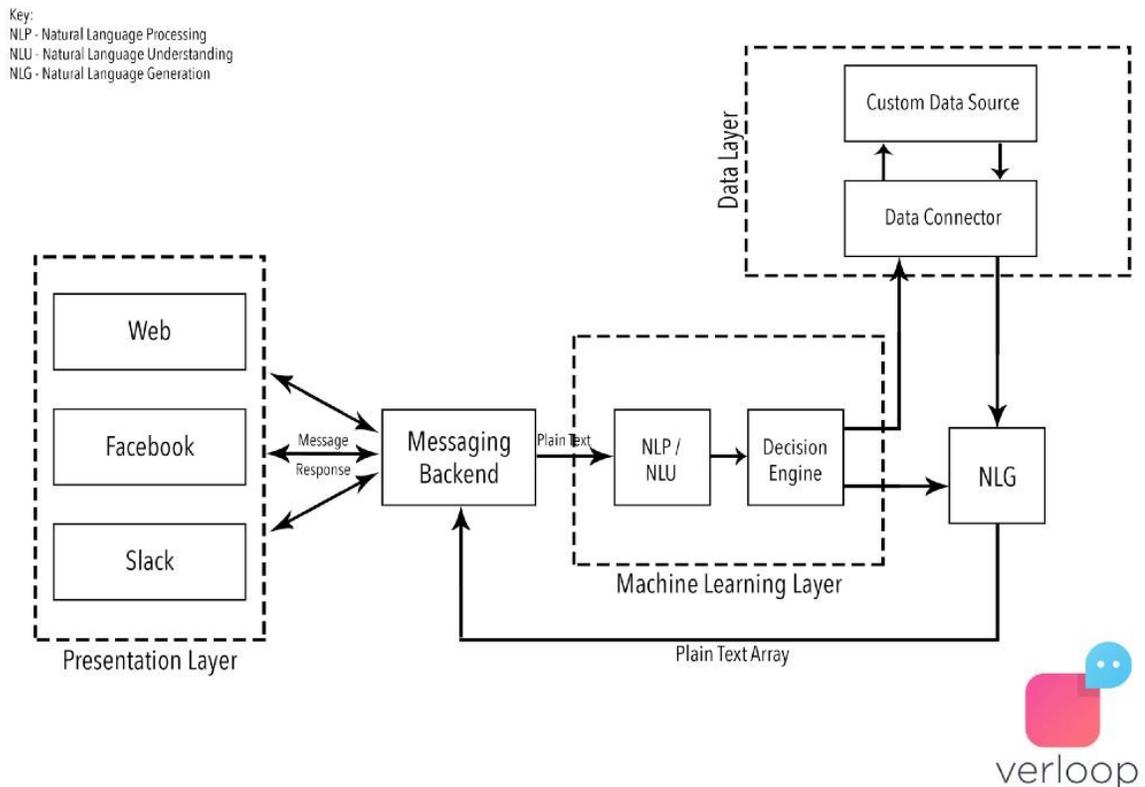


Figura 40 Diagrama de arquitectura de chatbots.

Fuente: Fuente: Definición de NLP, NLU, NLG y cómo funcionan los Chatbots (Fernandes, 2015)

2.6.5.1.1 Bot

Es una abreviación de robot o sistema robótico. Para una mejor definición se toma la publicación del El Economista:

“Un *bot* es un software de inteligencia artificial diseñado para realizar una serie de tareas por su cuenta y sin la ayuda del ser humano como hacer una reserva en un restaurante, marcar una fecha en el calendario o recoger y mostrar información a los usuarios”. (2016)

2.6.6 Retos de implementación de *machine learning*.

Una implementación se puede definir, según la Real Academia Española (2018) como: “Poner en funcionamiento o aplicar métodos, medidas, entre otros, para llevar algo a cabo.”

Estos retos van desde procesos administrativos, información, conocimientos, coordinación de una amplia gama de actividades, la supervisión del presupuesto, la incertidumbre, entre otras actividades.

En el artículo titulado “5 consejos para empezar con Machine Learning en la empresa”, por la empresa *CleverData* (2016) menciona que uno de los principales obstáculos a la hora de implementar *machine learning* en las empresas es debido a que:

“No tienen claro cómo empezar a usarlo y muchas veces tampoco entienden las ventajas que supone. Este último punto se despeja cuando mostramos ejemplos y explicamos lo más didácticamente posible la tecnología. En cambio, lo primero, cómo empezar a integrarlo en la empresa, es más complicado porque supone poner los pies en el suelo y empezar a trabajar

con una tecnología innovadora. La innovación tiene sus riesgos, pero estamos convencidos de que el *Machine Learning* ha llegado para quedarse y que va a revolucionar las sociedades, tanto como lo ha hecho el móvil”.

Es por ello por lo que un buen respaldo por medio de las juntas ejecutivas de las empresas genera un gran impacto a la hora de iniciar con una implementación con *machine learning*.

Según, el artículo titulado “Los desafíos para la implementación de *machine learning* en empresa” publicado por la empresa *GMP IT Consulting* menciona que:

“La incorporación de machine learning en las empresas gana cada vez más terreno en el campo de la innovación corporativa. Para comenzar un proyecto de este tipo los expertos aconsejan tres principios básicos: identificar un problema comercial específico, obtener apoyo ejecutivo e identificar beneficios comerciales mensurables a corto plazo”.

Dado lo anterior se define como retos de implementación como los impedimentos que se encuentran a la hora de poner en marcha la tecnología *machine learning*.

2.6.7 Tipos de consumo de *machine learning*.

El tipo de consumo que se quiere analizar en esta investigación se basa en la ubicación geográfica, para ello es bueno definir el concepto de “consumo” y se toma la referencia de Fernando Cuartas (2006) del diccionario financiero, se define el consumo como:

“La acción por la cual los diversos bienes y servicios son usados o aplicados a los fines a que están destinados, ya sea satisfaciendo las necesidades de los individuos o sirviendo los propósitos de la producción. La economía considera el consumo como el fin esencial de la actividad económica. Consumo es la utilización de bienes y servicios para nuestra satisfacción y la de otros.” (pág. 112)

Los tipos de consumo se clasifican de dos formas, local o en el extranjero. El término “local” se puede definir según *Oxford University* (2019): “Que es propio de un lugar o pertenece a él”. Según lo anterior se puede definir como consumo local, todo aquel producto o servicio que sea consumido en el territorio de Costa Rica.

El término “consumo en el extranjero,” según Dr. José Roberto Concha, director de Icecomex de la Universidad Icesi (2006), como: “el suministro de un servicio en un país, a un consumidor residente en otro país”.

2.7 Características organizativas de *machine learning*.

El desarrollo de *machine learning* supone una serie de características necesarias para que una empresa pueda organizarse para cumplir un objetivo o cualquier tipo de negocio. Matthew Schieltz (2013) explica en su artículo “Formas y características de una estructura organizativa”, la importancia de una buena organización:

Ya sean pequeñas o grandes, cada empresa debe considerar la forma en que se diseña su organización y estructura. Para operar con eficacia y eficiencia, una compañía necesita un sistema formal de comunicación, toma de decisiones y realización de tareas, que coincida con las necesidades de la organización.

2.7.1 Necesidades principales para desarrollar *machine learning*.

Las necesidades describen cuales son los elementos esenciales o prioritarios para la operación de desarrollo de productos o servicios con *machine learning*. En este caso tienen relevancia con los planes operativos que se pueden definir, según la referencia de la Cámara de Comercio de Santa Cruz de Tenerife (2019): “las operaciones de la empresa, es decir, el proceso fabricación del producto o de prestación del servicio, así como los recursos humanos, materiales y tecnológicos necesarios para el funcionamiento de dichas operaciones”.

2.7.2 Proyección del desarrollo tecnológico de *machine learning* en Costa Rica.

En los siguientes párrafos se recopilan opiniones de algunos autores sobre la proyección del desarrollo tecnológico desde la inteligencia artificial en general hasta el *machine learning* específicamente.

En la revista digital *IT Now* (2018), el artículo llamado “¿Cómo ha evolucionado la Inteligencia Artificial en Costa Rica?”, en el que hace referencia a un evento llamado “*DC Summit Digital Connection 2018*” en el que Óscar Ferraro

consultor de *GBM* habla sobre la realidad en Costa Rica sobre la inteligencia artificial:

“Ciertamente vivimos en un país cuyo desarrollo tecnológico e inclusive cultural, se encuentra desfasado en comparación con otros países, especialmente los de primer mundo. Sin embargo, en el territorio nacional cada vez estamos buscando innovar las empresas y avanzar hacia una Costa Rica más moderna”.

Además, en la publicación de *La Nación* (2018), “¿Qué hará la inteligencia artificial por nosotros?” entrevistan a Juan Luis Crespo, docente e investigador del TEC que menciona que: “...el futuro de la inteligencia artificial está en una interacción cada vez más natural entre máquinas y humanos, así como un crecimiento de flexibilidad en el desarrollo de la información”.

Desde la academia ven con mucho positivismo el desarrollo de la tecnología *machine learning* en Costa Rica, esto según el artículo “Simposio internacional mostró el presente y futuro en el innovador campo del *machine learning*”, por el Tecnológico de Costa Rica (2017), Saúl Calderón, coordinador del Simposio menciona:

“El sector tecnológico de Costa Rica está muy interesado en el tema, me parece que están incursionando en forma incipiente y necesitan del apoyo que les pueda dar la academia, porque tienen recursos limitados para hacer investigaciones. Hablamos de sinergias con la empresa privada para

investigaciones o prácticas. También nos interesa empezar a incluir esos temas en los currículos de las carreras involucradas, como computación, mecatrónica y computadores, con la finalidad de que nuestros egresados estén mejor preparados para el desarrollo de tecnologías que en el futuro serán fundamentales.

Lo cual demuestra una oportunidad de crecimiento de la rama de la inteligencia artificial, por lo que se espera conocer las opiniones de los entrevistados en esta investigación sobre el *machine learning* y su percepción a futuro.

2.7.3 Recurso humano

Se define recurso humano, según Barney y Clark (2007) como: “los recursos humanos incluyen la capacitación, la experiencia, el juicio, la inteligencia, las relaciones y el conocimiento de los gerentes y trabajadores individuales en las empresas” (pág. 24)

Un dato interesante en el artículo “Experto en *machine learning* será el perfil más demandado en España en el futuro” por la Revista *Cloud Computing* (2018) sobre un estudio que realizó la empresa *ServiceNow* menciona que:

“De cara a aprovechar todo el potencial que esta rama de la inteligencia artificial aporta; el estudio hace hincapié en la importancia de seleccionar empleados con el conocimiento y el talento necesarios para trabajar con procesos de aprendizaje automático y dispositivos inteligentes. De hecho, según datos de IDC, la inversión en aprendizaje automático casi se duplicará

antes de 2020 y los últimos análisis muestran que los especialistas en este campo son una de las figuras que más crecimiento están registrando en los departamentos informáticos.

La Figura 41 tomada del informe presentado por la empresa *ServiceNow* (2017) sobre la visión de los directores de sistemas informáticos en las empresas y menciona que el *machine learning* requiere talento y cambios en los procesos.

Con un 48% la respuesta fue “hacer cambios organizacionales para acomodarse a las necesidades del modelo de *machine learning*”, con 39% “redefinir la descripción de los puestos de trabajo para enfocarse en trabajos con máquinas”, con 27% “reclutar nuevos empleados con nuevas habilidades” y por último 18% “desarrollar políticas para asegurar la calidad de los datos”.

Fig. 5: Machine learning requires talent and process changes

Q: Which changes, if any, has your organization made? Respondents were asked to select all that apply. n=500

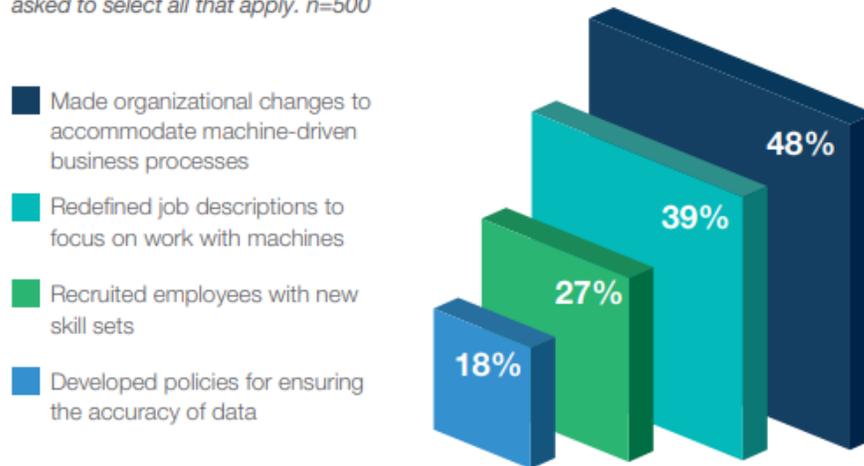


Figura 41 Estudio de ServiceNow sobre los cambios organizaciones para adoptar el machine learning.

Fuente: The Global CIO Point of View, ServiceNow (2017)

La Figura 41 demuestra que hay grandes esfuerzos por tener un modelo organizacional que beneficie el desarrollo del *machine learning* y además de encontrar recurso humano indicado para estos procesos.

Para esta investigación se pretende conocer cuáles son los puestos de trabajo encargados del desarrollo de *machine learning* en las empresas.

2.7.4 Capacitación de personal

Según, el sitio web *Emprende Pyme* (2010) define la capacitación como: “un conjunto de actividades didácticas orientadas a suplir las necesidades de la empresa y que se orientan hacia una ampliación de los conocimientos, habilidades

y aptitudes de los empleados, la cual les permitirá desarrollar sus actividades de manera eficiente”.

La capacitación de personal ha sido de interés para las empresas, el artículo “Capacitación y desarrollo personal en una empresa”, del sitio web Escuela Europea de Management (2017) menciona que: “El 84% de los directivos consideran como un factor muy importante la introducción de programas de capacitación y desarrollo personal en una empresa, según el informe *Global Human Capital Trends 2016*, de Deloitte”.

Existen diferentes estrategias para capacitar al personal de una empresa, entre ellos, cursos en línea, presenciales, también por medio de un especialista que comparte el conocimiento con el equipo de trabajo y las carreras o cursos universitarios.

Los cursos en línea tienen una modalidad interesante que permite ser matriculado de forma ilimitada, ser seguido en línea y de carácter abierto y gratuito. Según, el sitio MOOC.es (2019) define *MOOC* como:

El acrónimo en inglés de *Massive Online Open Courses* (o cursos online masivos y abiertos). Los cursos masivos no han sido otra cosa que la evolución de la educación abierta en internet. Todos los meses más de 700 universidades de todo el mundo, ofrecen cientos de cursos *online* gratuitos.

Las formas de capacitación para *machine learning* recopilados de una publicación titulada “¿Cómo capacitarse en machine learning?” del sitio web ReclUT (2018); estos cursos se pueden ver con más detalle en la Tabla 6.

Entre ellos hay plataformas *MOOC* o empresas grandes como *Google* que poseen cursos gratuitos para incrementar el conocimiento en *machine learning*.

Empresa	Curso
Coursera	Machine Learning Certification por Universidad de Stanford, uno de los cursos más populares, debido a que es impartido por Andrew Ng, ex jefe de Google Brain y Baidu AI Group. Entre los temas que se cubren están aprendizaje supervisado y no supervisado, mejores prácticas en aprendizaje automático y la estructura del programa se basará en estudios de casos múltiples y aplicaciones, para ayudarlo a aprender cómo aplicar algoritmos para construir robots inteligentes, comprensión de texto, informática médica, minería de bases de datos y otras áreas.
Coursera	Machine Learning Specialization. Curso de sobre especialización de ML en Python que consta de seis cursos. Se enfoca en la creación de aplicaciones basadas en ML, mediante el <i>Deep learning</i> . Un gran curso para las personas que quieren ir un paso

	<p>más allá de los métodos tradicionales de Machine Learning. Esta especialización completa es un poco costosa.</p>
Udacity	<p>Introducción al Machine Learning. Este curso familiariza al participante con los aspectos teóricos y prácticos del ML. Impartido por Sebastian Then, personaje clave de los autos sin conductor, este curso es un acercamiento más práctico para el profesional. Además de brindar un vistazo único con Python, un lenguaje utilizado para desarrollar en esta tecnología.</p>
Data Camp	<p>Un curso de certificación en ML, que va más enfocado a profesionales especializados en el entorno y lenguaje de programación R. Este curso se enfoca en brindar conocimiento útil sobre el uso de Machine Learning para modelos de capacitación de manera efectiva. El contenido del curso se entrega por medio de una combinación de videos y codificación de navegador web interactivo. Después de un par de módulos gratuitos, el curso completo es accesible a 25 dólares mensuales.</p>
Instituto Tecnológico de California	<p>Curso de Machine Learning por el Instituto Tecnológico de California. Un curso gratuito que está pensado para profesionales con cierto conocimiento en Ciencias de la Computación y que están pensando en desarrollar una carrera en ML. Los conceptos y técnicas fundamentales se explican en detalle. El foco de las</p>

	conferencias es la comprensión real, no solo el “conocimiento. Las conferencias usan visores incrementales (2853 en total) para simular el ritmo de enseñanza de la pizarra.
Google	Curso intensivo de aprendizaje automático de Google. Este curso destaca por ser uno de los pocos que se imparten en español, doblaje de voz que se realizó con herramientas de Machine Learning. Los participantes mejorarán sus conocimientos en álgebra y con el lenguaje de programación <i>Python</i> . El programa de estudio incluye conceptos como el estudio detallado del aprendizaje automático, reducción de la pérdida, primeros pasos con la API <i>TensorFlow</i> , validación y más.

Tabla 6 Cursos en línea para capacitación sobre machine learning.

Fuente: *¿Cómo capacitarse en Machine Learning?, ReclUT (2018).*

Con la capacitación de personal se hace referencia a las maneras cómo los trabajadores de una empresa mejoran sus conocimientos para adquirir una mayor efectividad en el desarrollo de *machine learning*.

Existen carreras enfocadas en *machine learning* como es el caso de la Universidad de *Carnegie Mellon* en Estados Unidos, pero tiene un costo muy elevado, como se puede ver en la Tabla 7 tomada del sitio web de la Universidad Carnegie Mellon (2019).

Tarifas para el 2019/2020	Por semestre	Por año*
Matrícula	\$ 24,500	\$ 49,000
Tarifa de Tecnología	\$ 215	\$ 430
Cuota de actividades	\$ 109	\$ 218
Cuota de transporte	\$ 112	\$ 224

Tabla 7 Matrícula para el Master en Ciencias de Machine learning de la Universidad de Carnegie Mellon.

Fuente: *Master of Science in Machine Learning, Financial Information, Carnegie Mellon University (2019).*

Para esta investigación se quiere conocer cuál es la forma más utilizada para capacitación de personal en las empresas entrevistadas.

2.7.5 Áreas económicas de vinculación impactadas con el machine learning.

En esta investigación se quieren conocer cuáles son las áreas o actividades económicas que están siendo impactadas por el *machine learning* por lo cual es importante conocer cómo se define una área o actividad económicas.

Según, Juan David Montoya, economista (2012), define actividad económica de la siguiente manera:

“Las actividades económicas o productivas son procesos que por medio del uso de factores de producción crean bienes y servicios para satisfacer las

necesidades de los consumidores en la economía. Estas incluyen actividades comerciales, pues el comercio también agrega valor a la economía.

El periódico El País (2017), publica el artículo “La inteligencia artificial obliga a redefinir la economía”, y menciona lo siguiente en cuanto al impacto de la inteligencia artificial en la economía:

“Bank of America Merrill Lynch prevé que en 2025 el impacto disruptivo de la inteligencia artificial podría alcanzar un rango de entre 14 billones y 33 billones de dólares, incluyendo nueve billones en ahorro de costes por la automatización de puestos de trabajo. Mckinsey Global Institute pone en perspectiva el momento que vivimos: “La contribución de la inteligencia artificial en la transformación de la sociedad será 3.000 veces superior a la revolución industrial”.

En la Figura 42 tomada del informe *The Age of analytics: competing in a data-driven world*, McKinsey Global Institute (2016), explica los principales indicadores de disrupción del *machine learning*. Entre los dominios más impactados se encuentran, empresas aseguradoras, sector salud, transporte y logística, ventas al por menor, sector bancario, capital humano y talento, entre otros.

Indicadores de potencial de disrupción:	Arquetipo de la disrupción	Dominios que podrían ser interrumpidos
Los activos están subutilizados debido a una señalización ineficiente.	Modelos de negocio habilitados por datos ortogonales.	Seguros Cuidados médicos Capital humano y talento
Desajuste oferta-demanda	Hiperescala, coincidencia en tiempo real	Transporte y logística Automovilismo Ciudades inteligentes e infraestructura
Dependencia de gran cantidad de datos personalizados.	Personalización radical	Cuidado médico Ventas al por menos Medios de comunicación Educación
Los datos están comprimidos o fragmentados	Capacidades masivas de integración de datos	Bancos Seguros Sector público Capital humano y talento
Gran valor en la combinación de datos de múltiples fuentes.	Descubrimiento basado en datos	Ciencias de la vida y farmacéuticas Ciencias materiales Tecnología
La I + D es fundamental para el modelo de negocio.	Toma de decisiones mejorada	Ciudades inteligentes Cuidado médico Seguros Capital humano y talento
La toma de decisiones está sujeta a sesgos humanos		
Velocidad de toma de decisiones limitada por limitaciones humanas.		
Gran valor asociado con la mejora de la precisión de la predicción.		

SOURCE: McKinsey Global Institute analysis

Figura 42 Dominios de disrupción del machine learning.

Fuente: *The Age of analytics: competing in a data-driven world*, McKinsey Global Institute (2016).

Dado lo anterior es importante conocer cuáles áreas económicas en Costa Rica más impactadas por el *machine learning*, según las empresas entrevistadas.

En el siguiente capítulo se desarrolla el proceso metodológico de esta investigación.

Capítulo III
Marco Metodológico

3.1 Tipo de investigación

Dado que la investigación fue dirigida a la determinación de las características del *machine learning* presentada en la oferta de servicios tecnológicos digitales en las empresas entrevistadas, esta abarcó el tipo de investigación explicativa y descriptiva.

Esta investigación se planteó de corte exploratoria, según Hernández (2014) define lo siguiente:

“Los estudios exploratorios se realizan cuando el objetivo es examinar un tema o problema de investigación poco estudiado, del cual se tienen muchas dudas o no se ha abordado antes. Es decir, cuando la revisión de la literatura reveló que tan solo hay guías no investigadas e ideas vagamente relacionadas con el problema de estudio o bien, si deseamos indagar sobre temas y áreas desde nuevas perspectivas”. (pág. 91)

Hernández Sampieri (2014) menciona sobre la investigación descriptiva lo siguiente:

“Con frecuencia, la meta del investigador consiste en describir fenómenos, situaciones, contextos y sucesos; esto es, detallar cómo son y se manifiestan. Con los estudios descriptivos se busca especificar las propiedades, las características y los perfiles de personas, grupos, comunidades, procesos, objetos o cualquier otro fenómeno que se someta a un análisis. Es decir, únicamente pretenden medir o recoger información de manera independiente

o conjunta sobre los conceptos o las variables a las que se refieren, esto es, su objetivo no es indicar cómo estas se relacionan. (pág. 92)

En relación con el criterio descriptivo y exploratorio dados, se planteó que esta investigación busca que sus resultados representen de forma acertada el análisis de las características del *machine learning*, las necesidades de las empresas que venden este servicio tecnológico, además de dar un panorama amplio sobre el ecosistema tecnológico y sus actores.

La investigación descriptiva se guía por las preguntas de investigación que formuló el investigador, no busca explicaciones o razones del porqué, de las situaciones sino solo representa los acontecimientos en un lugar y tiempo.

Por otra parte, este tipo de estudio permitió analizar el estado de desarrollo, la funcionalidad y los principales retos organizativos para el desarrollo de la tecnología *machine learning* de las empresas proveedoras que participan.

A su vez, permitió la descripción de los elementos principales, tales como: la realidad en la que se encuentra el país y el mercado que se quiere incursionar, así como conocer situaciones, costumbres y actitudes predominantes mediante la descripción de las actividades y los objetos presentes en el uso de esta tecnología.

3.2 Enfoque de investigación

Este proyecto contempló un enfoque mixto, pues en el caso del enfoque cualitativo se abarca un análisis de las características del *machine learning*, que pretende especificar los *Frameworks*, métodos, estilos, tipos de problemas que se solventan y áreas económicas que consumen tecnología, casos de uso de implementación de *machine learning*, ventajas, retos, consumo local o extranjero, necesidades, visión a futuro y capacitación de recurso humano.

Lo anterior se sustentan en la siguiente cita dado por el autor Hernández Sampieri, (2014) quien menciona:

“Los métodos mixtos representan un conjunto de procesos sistemáticos, empíricos y críticos de investigación e implican la recolección y el análisis de datos cuantitativos cualitativos, así como su integración y discusión conjunta, para realizar inferencias producto de toda la información recabada y lograr un mayor entendimiento del fenómeno bajo estudio”. (pág. 534)

Para el enfoque cuantitativo se logró cubrir cuántas son las empresas que utilizan cierto tipo de *machine learning* y cuantificar el uso de la tecnología en un área económica por medio de un informe de *estatus* de la tecnología en el país que brindará resultados sobre los métodos, *Frameworks*, área económica, casos de uso, visión a futuro de la tecnología en el país de las empresas proveedoras.

3.3 Población

En el caso específico de esta investigación, la población objetivo se delimitó por espacio, directamente en la GAM con el fin de contar con la mayor cantidad de empresas similares.

Lo anterior se sustenta en lo dado por Arias, Fidas G. (2012) define población o población objetivo a: “Un conjunto finito o infinito de elementos con características comunes para los cuales serán extensivas conclusiones de la investigación. Esta queda determinada por el problema y por los objetivos del estudio”. (p. 81)

Esta población dio la posibilidad de contar con un espectro mayor de empresas que cuenten con las características requeridas en la investigación y segundo se delimitó por medio de una encuesta, pues no existe un registro de empresas proveedoras con ofertas de servicios tecnológicos digitales con *machine learning* para definir una población de análisis.

Se mantuvo comunicación con diferentes entes estatales que tuvieran relación con el desarrollo tecnológico o de desarrollo económico empresarial, pero no se pudo encontrar un registro de empresas específico con la especialidad de *machine learning*, es por ello por lo que se justifica el trabajo de campo. Ver correos electrónicos de los entes consultados. Ver Apéndice A, Apéndice B, Apéndice C y Apéndice D que tienen que ver sobre las consultas realizadas.

Como parte de la definición, se decidió por un trabajo de campo personal, detallando por medio de una encuesta cuáles empresas ofertan de servicios tecnológicos digitales con *machine learning*.

Para Richard L. Sandhusen (2002), “las encuestas obtienen información sistemáticamente de los encuestados por medio de preguntas, ya sea personales, telefónicas o por correo”. (pág. 229).

Una gran fuente de información para recomendaciones de empresas que cuenten con servicios con *machine learning* fue por medio de un grupo llamado *Machine Learning Costa Rica* ubicado en la red social *Facebook*, que cuenta al momento de esta investigación con más de 1000 miembros, estos se dedican a compartir el conocimiento sobre proyectos, fuentes de empleo y hasta reuniones, algunas veces patrocinadas por empresas que dedican esfuerzos a este tipo de tecnología. De esta forma, se logró la identificación de empresas de forma más precisa, pues hay profesionales dentro de estas empresas y son contacto directo para la aplicación del instrumento de recolección de información.

Otra forma utilizada para la definición de la población fue por investigación por internet, por lo que muchas empresas publican en sus sitios los servicios con *machine learning*. También tomando en cuenta a empresas grandes de nivel internacional, para comprobar si poseen alguna actividad basada en *machine learning* en Costa Rica.

La siguiente tabla es información pública que puede ser encontrada en sitios web o por llamada telefónica a las empresas, por lo que su uso no tiene restricción. Estas empresas fueron encuestadas para constatar que hacen uso de la tecnología de *machine learning*, con el objetivo de filtrar a las empresas.

Subsector Tecnológico – Desarrollo de software, Ingeniería en software.

Empresas	Sector	Contacto
IBM Cognitiva	Tecnológico	+506 2105-9850 info@cognitiva.la dzuluaga@cognitiva.la
Xchematic	Tecnológico	ccerrato@xchematic.com
Escala24x7	Tecnológico	alejandro.planas@escala24x7.com
SimMachines	Tecnológico	julianrodriguez@simmachines.com
PrediSoft	Tecnológico	infocr@predisoft.com
Amazon Costa Rica	Tecnológico	2562 9111
Intel	Tecnológico	2298 6000
Grupo Prides	Tecnológico	info@grupoprides.com (506) 2528-5904 Byron Roman, broman@prides.net
Neustar	Tecnológico	Robin.Villalobos@neustar.biz Robin.Villalobos@team.neustar

TradeStation	Financiera tecnológica	4111 3100 Ptorre@tradestation.com
Crux Consultores	Tecnológico	rrodriguez@cruxconsultores.com
Equifax	Tecnológico	(506) 2509-0700
13Bots	Tecnológico	8947 9486 Ricardomurillo13@gmail.com
Singularities	Tecnológico	+506 8330-9787 Javier Porras javier.porras@singularities.com
LiderSoft	Tecnológico	<u>info@lidersoft.com</u> Giovanny Portuguez Gerente General
Avantica	Tecnológico	2283 9100 Luis.avila@avantica.net Cartago +5062283-9100 x 1410 Guillermo.zamora@avantica.net
Proximity Costa Rica	Tecnológico	info@proximitycr.com (506) 2288-1097 live@proximitycr.com
Softland	Tecnológico	2509-3800 800-softland
Kuiki	Tecnológico y Financiera	Oficinas Centrales: +506 4100-0770 servicioalcliente@kuikicredit.com

Mobilize	Tecnológico	+ 506.2519.1000
Latern Tech	Tecnológico	Moravia, Costa Rica Tel. +506 2535-6262 info@lantern.tech
Doer	Tecnológico	contacto@doer.cr
ARWEB	Tecnológico	info@arweb.com
Miweb.cr	Tecnológico	22538881 carlos@miweb.cr
AliveBox	Tecnológico	22402254 info@alivebox.com
GBM	Tecnológico	(506) 2284-3999 (506) 2284 3911 3117 Maria Jose
Logosoft	Tecnológico	(506) 2253-0800
Encore Capital	Tecnológico y Financiero	https://www.facebook.com/pg/encorecostarica Fabiola.alfaro@mcmcg.com
MediaLAB	Tecnológico y Publicidad	costarica@medialabla.com
Sapiens	Tecnológico	+506.4036.2700 sanjose@sapiensdev.com
Nova Comp	Tecnológico	2216-5800
Intertect	Tecnológico	2589 9200

Neocreativa	Tecnológico	4033-6718
Pwc	Tecnológico y Consultora	2224-1555

Tabla 8 Definición de la población con indicios de servicios basados en machine learning.

Fuente: Elaboración propia.

3.4 Fuentes de información

Respecto de las fuentes de información, Arias (2012), lo define como: “es todo lo que suministra datos o información. Según su naturaleza, las fuentes de información pueden ser documentales (proporcionan datos secundarios), y vivas (sujetos que aportan datos primarios)” (pág. 27).

En el caso de esta investigación se clasifica de la siguiente manera:

- Fuentes primarias:
 - Libros de Inteligencia artificial y *machine learning*, *Artificial Intelligence A Modern Approach (3rd Edition)*, *Machine Learning, an Algorithmic Perspective (2nd Edition)*, *Machine Learning por Tom. Mitchell*.
 - Reportes analíticos de *McKinsey Global Institute*, *The Age of analytics: competing in a data-driven world*,
 - Desarrolladores de software de empresas proveedoras que ofertan servicios tecnológicos digitales con *machine learning*, gerentes de empresas de desarrollo de software, científicos de datos y especialistas en *machine learning*.

- Fuentes secundarias: Se basarán en artículos de revistas y páginas de la web con información acerca las características del *machine learning*, tal es el caso de sas.com, Grupo Oficial en Facebook *Machine learning* Costa Rica, TECCOM 2017 donde se expone la alta tecnología en Costa Rica, Grupo Parma del Tecnológico de Costa Rica y PRISMA de la Universidad de Costa Rica.

3.5 Sujetos de información

En cuanto a los sujetos de información, Arias (2012) lo define como: “la persona que busca, obtiene o posee el conocimiento y el objeto como el hecho, fenómeno, tema o materia que el sujeto estudia.” (pág. 13)

En cuanto a los sujetos de información se pueden dividir en

- Departamentos de desarrollo de software: Colaboradores relacionados con el desarrollo de software, que tienen contacto directo con *frameworks*, métodos y proyectos que utilicen *machine learning*.
- Científicos de datos: Quienes crean modelos de *machine learning* para ser implementados en sistemas informáticos o sistemas especiales.
- Líderes de proyectos: Responsables por garantizar el éxito del proyecto, responsabilidad total del planeamiento y la ejecución acertadas del proyecto.

- Empresas proveedoras de servicios tecnológicos digitales con *machine learning*: El escenario principal de la investigación.

3.6 Planeación para el proceso de recolección de información.

3.6.1 Planteamiento de objetivos de la visita.

La respuesta obtenida de las entrevistas permitió la comprensión del estado de desarrollo del *machine learning*, las características y las áreas económicas impactadas por esta tecnología. Se realizó mediante visitas a las empresas con el espacio para explicar el propósito de la investigación y conocer lo que acontece con la tecnología de *machine learning*.

3.6.2 Población Objetivo

De acuerdo con la investigación, el enfoque fue delimitado al sector tecnológico digital que ha cumplido con la encuesta con la especificación de utilizar *machine learning* en sus servicios y que se encuentre en la gran área metropolitana.

3.6.3 Establecimiento del método de medición.

Como establecimiento del método de investigación, se realiza una presentación personal, también se envía mediante correo electrónico un volante informativo que explique con exactitud (ver Apéndice K), el proceso que de recolección y, por último, se envía mediante correo electrónico el cuestionario en

formato digital (Formulario de *Google*, ver Apéndice M) en caso de que la persona no cuente con tiempo suficiente para una entrevista personal.

Para el establecimiento del método de medición se utilizó un diagrama de flujo de elaboración propia que se encuentra en el Apéndice O que especifica como se inicia el procedimiento y como se aborda el filtrado de las empresas hasta la aplicación del instrumento de medición.

3.6.4 Instrumento de medición

En cuanto a los instrumentos de medición utilizados, se rescata lo dicho por Arias (2012) quien define técnica de recolección como: “Se entenderá por técnica de investigación, el procedimiento o forma particular de obtener datos o información.” (pág. 67)

Para la recolección de información de manera efectiva se plantea utilizar la entrevista, para ello se toma la definición de Arias (2012) entrevista como:

La entrevista, más que un simple interrogatorio, es una técnica basada en un diálogo o conversación “cara a cara”, entre el entrevistador y el entrevistado acerca de un tema previamente determinado, de tal manera que el entrevistador pueda obtener la información requerida. (pág. 73)

Además, se detalla que el tipo de entrevista es semi estructurada o formal, por lo que se cuenta con una guía de preguntas prediseñada que se hace llegar previamente al entrevistado para estar de acuerdo con su implementación.

Según Arias (2012):

“Aun cuando existe una guía de preguntas, el entrevistador puede realizar otras no contempladas inicialmente. Esto se debe a que una respuesta puede dar origen a una pregunta adicional o extraordinaria. Esta técnica se caracteriza por su flexibilidad. Además de sus instrumentos específicos, tanto la entrevista estructurada como la no estructurada pueden emplear instrumentos tales como el grabador y la cámara de video. (pág. 74)

En el caso de no poder realizar la entrevista personal, se contempló utilizar el cuestionario como técnica de recolección por medio de internet y que los entrevistados se sientan más cómodos de responder.

Para el caso específico Arias (2012), lo define como:

“Es la modalidad de encuesta que se realiza de forma escrita mediante un instrumento o formato en papel contentivo de una serie de preguntas. Se le denomina cuestionario auto administrado porque debe ser llenado por el encuestado, sin intervención del encuestador”. (pág. 74)

La guía de preguntas utilizada en la entrevista fue la misma utilizada en el cuestionario, con preguntas abiertas para la obtención de mayor cantidad de información, ver Apéndice L.

Los instrumentos de recolección cumplen los objetivos específicos, porque se aplican a la totalidad de la población identificada, según un objetivo específico de esta investigación y las herramientas que se aplican están claramente definidas y aprobadas por el tutor encargado de la investigación.

3.6.5 Preparación de los investigadores de campo.

Se planteó un protocolo para visita a las empresas proveedoras de servicios con *machine learning* con tal de presentar toda la formalidad y respeto en el momento de la entrevista a los participantes de la investigación.

A las empresas elegidas se les envió un volante informativo y las preguntas con anticipación vía correo electrónico para saber si están dispuestos a participar en la investigación, además de aclarar dudas. Ver Apéndice K.

Las visitas se agendaron con anticipación para no perjudicar las labores de los participantes, en caso de que no se pudo concretar la visita personal captó la información por medio del cuestionario alojado en un formulario digital de *Google*.

Para esto se realizó lo siguiente en cada visita:

- **Presentación:** Primero, se hizo una presentación del autor de la tesis, mostrando la carta de la Universidad Técnica Nacional que permitió mayor fiabilidad para la investigación y además se dio a conocer una

breve explicación del tema de investigación de la tesis y por qué su empresa ha sido elegida población de estudio.

- Reglas de la visita: Se explicó e informó sobre el proceso de la investigación a cada empresa y cómo sería el protocolo de recolección de información y la respectiva aclaración de dudas. A continuación, el protocolo:
 - Propósito general del estudio.
 - Motivaciones para el entrevistado (importancia de su participación).
 - Agradecimiento.
 - Tiempo aproximado de respuesta (un promedio o rango). Lo suficientemente abierto para no presionar al participante, pero tranquilizarlo.
 - Espacio para que firme o indique su consentimiento.
 - Identificación de quién o quiénes lo aplican.
 - Explicación breve del cómo se procesaría los cuestionarios y una cláusula de confidencialidad del manejo de la información individual.
 - Instrucciones iniciales claras y sencillas (cómo responder en general, con ejemplos si se requiere).

- Aceptación: En este punto se supo si alguien perteneciente a la empresa está dispuesto a proceder en ese momento, si podrá atender algún otro día, si participa por medio del formulario web o difiere participar en la investigación.
- Información de contactos: Se recolectaron los contactos para comunicarse con las empresas, ya sea por medio de correo electrónico, redes sociales, entre otros; cuyo objetivo es hacerles llegar un recordatorio y la dirección web de la encuesta que deben completar o cualquier otra consulta.
- Cierre: En este último punto se abre un espacio para que se pueda hacer cualquier tipo de preguntas y responder dudas, por otra parte, se agradece la atención tomada y el tiempo disponible.

3.6.6 Prueba piloto

La prueba piloto es crucial, pues permitió probar en el trabajo de campo cada uno de los pasos de la visita. Los resultados de la prueba piloto usualmente sugieren modificaciones antes de realizar las visitas.

Esta prueba consistió en seleccionar una empresa que forma parte de la población, y aplicar todo el proceso de visitas que se describió anteriormente, y por último, después de la presentación, aclarar cualquier duda y preguntar a la(s) persona(s) que atiende(n) en ese momento, como consideró la presentación realizada, si observa algún punto de mejora.

3.6.7 Organización y análisis de la información recolectada.

Para el proceso de organización y análisis de la información es necesario un ordenamiento y filtrado ya que la preguntas son abiertas en la entrevista, así como en el cuestionario.

Para esto se crearon cuatro (4) secciones:

- Información del entrevistado
- *Frameworks*
- Métodos
- Áreas económicas, Industria impactada y Recurso humano

Las secciones 2,3 y 4 se analizaron para la extracción de coincidencias cualitativas para la cuantificación del porcentaje de respuesta. Después se creó una matriz que permite determinar las relaciones entre las secciones.

Después de ordenar, filtrar, analizar y categorizar las respuestas se procedió a crear un resumen estadístico y hallazgos de toda la información recolectada.

3.7 Temporalidad

En cuanto a la temporalidad de la investigación, Hernández Sampieri (2010) dice: “El investigador contextualiza las experiencias en términos de su temporalidad (tiempo en que sucedieron), espacio (lugar en el cual ocurrieron), corporalidad (las

personas físicas que la vivieron) y el contexto relacional (los lazos que se generaron durante las experiencias).” (pág. 516).

En este caso, la investigación se ejecutó en los intervalos de mayo 2017 a junio 2018 para contar el tiempo suficiente para la identificación de las empresas, después para la aplicación de los instrumentos de recolección y finalmente el análisis la información recolectada. Además, para cumplir con tiempo disponible, según el Reglamento de Trabajos Finales de Graduación de la Universidad Técnica Nacional.

3.8 Matriz metodológica

A continuación, se muestra la matriz metodológica de la investigación, se divide por cada objetivo específico:

Establecer la oferta de servicios tecnológicos digitales basados en *machine learning* existentes en la Gran Área Metropolitana de Costa Rica, para lograr la sectorización de las empresas proveedoras de servicios mediante un planteamiento de características técnicas, funcionales y organizativas.

Variable	Definición conceptual	Dimensiones	Indicadores	Instrumento
Oferta de servicios tecnológicos digitales basados <i>machine learning</i> .	Conjunto de elementos tecnológicos digitales, metodológicos, técnicos y características	Conjunto de características tecnológicas digitales. Conjuntos de características	Tipos Tecnologías digitales y áreas tecnológicas.	Documentación fidedigna en internet. Documentación de autores de libros de

	de <i>machine learning</i> que cumplen las necesidades y deseos de los consumidores.	técnicas del <i>machine learning</i> Conjunto de servicios aplicados	Características técnicas del <i>machine learning</i> (Cantidad Tipos, Cantidad de métodos, Cantidad <i>frameworks</i>). Tipos de empresa (área económica particular). Servicios (Tipo de servicios aplicados, Tipos de áreas económicas).	relevancia en el tema. Documentación de recomendaciones de especialistas del tema. Estudio de mercado costarricense y la relevancia del tema.
--	--	---	--	---

Tabla 9: Objetivo 1 de la Matriz Metodológica

Fuente: Elaboración propia.

Identificar las empresas proveedoras que ofertan servicios tecnológicos digitales basados en *machine learning*, por medio de recopilación de información, que permita la definición total de la población al periodo estudiado.

Variable	Definición conceptual	Dimensiones	Indicadores	Instrumento
Empresas proveedoras que ofertan servicios	Unidad económico-social, integrada por elementos	Encuesta a una selección de empresas que tengan indicios de utilizar	Número de empresas.	Documentación fidedigna en internet.

tecnológicos digitales con <i>machine learning</i> .	humanos, materiales y técnicos, que tiene el objetivo de obtener utilidades por medio de su participación en el mercado de bienes y servicios. Para esto, hace uso de los factores productivos por medio de tecnologías digitales con <i>machine learning</i>	tecnologías como <i>machine learning</i> .	Sector económico al que pertenece.	Consulta con expertos o profesionales en el área de <i>machine learning</i> . Encuesta vía llamada de telefónica. Encuesta vía correo electrónico.
--	---	--	------------------------------------	--

Tabla 10 Objetivo 2 de la Matriz Metodológica

Fuente: Elaboración propia.

Analizar las características técnicas, funcionales y organizativas del *machine learning* que se encuentran en la oferta de servicios tecnológicos digitales de las empresas identificadas para la evaluación del desarrollo tecnológico.

Variable	Definición conceptual	Dimensiones	Indicadores	Instrumento
Características técnicas, funcionales y	Conjunto de elementos metodológicos, técnicos, económicos,	Conjunto de tipos de <i>machine learning</i> .	Cantidad tipos de <i>machine learning</i> .	Entrevistas. Cuestionarios.

<p>organizativas del <i>machine learning</i>.</p>	<p>laborales de relevancia con <i>machine learning</i>.</p>	<p>Conjuntos de métodos de <i>machine learning</i>.</p> <p>Conjunto de técnicas de <i>machine learning</i>.</p> <p>Conjunto de claves que generen ventajas competitivas.</p> <p>Conjunto de aplicaciones y casos de uso.</p> <p>Conjunto de elementos económicos.</p> <p>Conjunto de puestos laborales.</p>	<p>Cantidad de métodos.</p> <p>Cantidad de <i>frameworks</i> utilizados en proyectos.</p> <p>Cantidad de áreas económicas que están destinados los servicios de <i>machine learning</i>.</p> <p>Tipos de casos de uso (tipo de área económica, problema a solucionar).</p> <p>Tipos de ventajas (claves, desempeño).</p> <p>Tipos de retos (Tiempo, económicos, aprendizaje, percepción).</p>	
---	---	---	---	--

			<p>Tipo de consumo (local o exterior del país).</p> <p>Tipos de requisitos para desarrollar <i>machine learning</i> (Económicos, tiempo, conocimiento).</p> <p>Tipo de proyección (Positiva o negativa).</p> <p>Tipos de puestos laborales.</p> <p>Tipos de capacitaciones (Local, virtual, formal, informal).</p>	
--	--	--	--	--

Tabla 11 Objetivo 3 de la Matriz Metodológica

Fuente: Elaboración propia.

Generar un informe de resultados que permita la comprensión de las características técnicas, funcionales y organizativas de la oferta de servicios tecnológicos digitales basados en *machine learning*, para la vinculación con las áreas económicas impactadas por la tecnología.

Variable	Definición conceptual	Dimensiones	Indicadores	Instrumento
Informe para comprender el vínculo que existe entre las características de <i>machine learning</i> (<i>frameworks</i> , métodos, tipos) dentro de la oferta de servicios tecnológicos digitales y las áreas económicas impactadas.	Características de <i>machine learning</i> (<i>frameworks</i> , métodos, tipos) dentro de la oferta de servicios tecnológicos digitales analizadas para determinar su vinculación con las áreas económicas impactadas.	<p>Casos de uso de los <i>framework</i> específicos de <i>machine learning</i>.</p> <p>Casos de uso de los métodos específicos de <i>machine learning</i>.</p> <p>Casos de uso de los tipos de <i>machine learning</i>.</p>	<p>Cantidad tipos de <i>machine learning</i>.</p> <p>Cantidad de Métodos.</p> <p>Cantidad de <i>frameworks</i> utilizados en proyectos.</p> <p>Cantidad de áreas económicas que están destinadas los servicios de <i>machine learning</i>.</p>	<p>Entrevistas.</p> <p>Cuestionarios.</p> <p>Análisis de la población.</p> <p>Análisis de las características de <i>machine learning</i>.</p> <p>Análisis de las áreas económicas destinadas.</p> <p>Análisis y conclusiones de la relación que existe con</p>

			<p>Tipos de casos de uso (tipo de área económica, problema a solucionar).</p> <p>Tipos de ventajas (claves, desempeño).</p> <p>Tipos de retos (Tiempo, económicos, aprendizaje, percepción).</p> <p>Tipo de consumo (local o exterior del país).</p> <p>Tipos de requisitos para desarrollar <i>machine learning</i> (Económicos, tiempo, conocimiento).</p> <p>Tipo de proyección</p>	<p>las características técnicas y las áreas económicas.</p> <p>Documentación fidedigna en internet.</p> <p>Documentación de autores de libros de relevancia en el tema.</p>
--	--	--	--	---

			<p>(Positiva o negativa).</p> <p>Tipos de puestos laborales.</p> <p>Tipos de capacitaciones (Local, virtual, formal, informal).</p> <p>Tipos de empresa (área económica particular).</p> <p>Servicios (Tipo de servicios aplicados, Tipos de áreas económicas).</p>	
--	--	--	---	--

Tabla 12 Objetivo 4 de la Matriz Metodológica

Fuente: *Elaboración propia.*

Capítulo IV
Análisis situacional

4.1 Oferta de servicios tecnológicos digitales basados en machine learning, área metropolitana de CR.

En este apartado se analiza la oferta de servicios tecnológicos digitales basados en *machine learning*, específicamente en el área metropolitana de Costa Rica, que permiten la sectorización y análisis de las características técnicas, funcionales y organizativas encontradas en la oferta.

4.1.1 Resumen y análisis de la información recolectada.

4.1.1.1 Análisis de la población

Se realizó la encuesta a 33 empresas del área tecnologías digitales y que tuvieran alguna relevancia con desarrollo de software, análisis de datos, marketing digital e inteligencia de negocio.

De las empresas encuestadas se pudo constatar que 16 utilizan *machine learning* en sus productos o servicios, pero es importante mencionar que después de hacer esfuerzos por tener contacto por diferentes medios, sin embargo, en algunos casos no hubo respuesta por lo que la información por utilizar es de las 16 empresas que se logró contactar y corresponden al 100% de la población.

4.1.1.1.1 Empresas proveedoras que ofrecen servicios de *machine learning*

De acuerdo con las consultas realizadas a las empresas que conforman la población de interés, se obtuvo los siguientes resultados:

Población encuestada

	Empresas	Categoría
1	Grupo Prides	Utilizan tecnología <i>Machine learning</i> en sus productos o servicios.
2	IBM Cognitiva	Utilizan tecnología <i>Machine learning</i> en sus productos o servicios.
3	Escala24x7	Utilizan tecnología <i>Machine learning</i> en sus productos o servicios.
4	miweb.cr	Utilizan tecnología <i>Machine learning</i> en sus productos o servicios.
5	Singularities	Utilizan tecnología <i>Machine learning</i> en sus productos o servicios.
6	13Bots	Utilizan tecnología <i>Machine learning</i> en sus productos o servicios.
7	Neustar	Utilizan tecnología <i>Machine learning</i> en sus productos o servicios.
8	TradeStation	Utilizan tecnología <i>Machine learning</i> en sus productos o servicios.
9	Xchematic	Utilizan tecnología <i>Machine learning</i> en sus productos o servicios.
10	SimMachines	Utilizan tecnología <i>Machine learning</i> en sus productos o servicios, esto por medio de la página web de simMachines y además de su respuesta por correo electrónico.

11	Crux	Utilizan tecnología <i>Machine learning</i> en sus productos o servicios, esto por medio de la página web de Consultores Crux.
12	Lantern Tech	No se pudo concretar una comunicación con el departamento encargado, pero sí se confirma por medios publicitarios que brindan servicios de ese tipo.
13	Predisoft	No se pudo concretar una comunicación con el departamento encargado, pero sí se confirma por medios publicitarios que brindan servicios de ese tipo.
14	Intel	No se pudo concretar una comunicación con el departamento encargado, pero sí se confirma por medios publicitarios que brindan servicios de ese tipo.
15	Avantica	No se pudo concretar una comunicación con el departamento encargado, pero sí se confirma por medios publicitarios que brindan servicios de ese tipo.
16	Proximity	No se pudo concretar una comunicación con el departamento encargado, pero sí se confirma por medios publicitarios que brindan servicios de ese tipo.
17	Kuiki	No se pudo concretar una comunicación con el departamento encargado.

18	Equifax	No se pudo concretar una comunicación con el departamento encargado.
19	Doer	No se pudo concretar una comunicación con el departamento encargado.
20	ARWEB	No se pudo concretar una comunicación con el departamento encargado.
21	Alivebox	No se pudo concretar una comunicación con el departamento encargado.
22	GBM	No se pudo concretar una comunicación con el departamento encargado.
23	Logosoft	No se pudo concretar una comunicación con el departamento encargado.
24	NovaComp	No se pudo concretar una comunicación con el departamento encargado.
25	Intertec	No se pudo concretar una comunicación con el departamento encargado.
26	Neocreativa	No se pudo concretar una comunicación con el departamento encargado.
27	Mobilize (ArtinSoft)	No se pudo concretar una comunicación con el departamento encargado.
28	Amazon Costa Rica	No se pudo concretar una comunicación con el departamento encargado.
29	Lidersoft	No desarrollan productos con <i>Machine learning</i> .

30	Softland	No desarrollan productos con <i>Machine learning</i> .
31	Encore capital	No utilizan tecnología <i>Machine learning</i> en sus productos o servicios.
32	MediaLAB	No utilizan tecnología <i>Machine learning</i> en sus productos o servicios.
33	Sapiens	No utilizan tecnología <i>Machine learning</i> en sus productos o servicios.

Tabla 13 Cantidad de empresas encuestadas.

Fuente: Elaboración propia.

Para mayor entendimiento se presenta el siguiente gráfico del resultado de la encuesta aplicada a las empresas listadas en la Tabla 13.

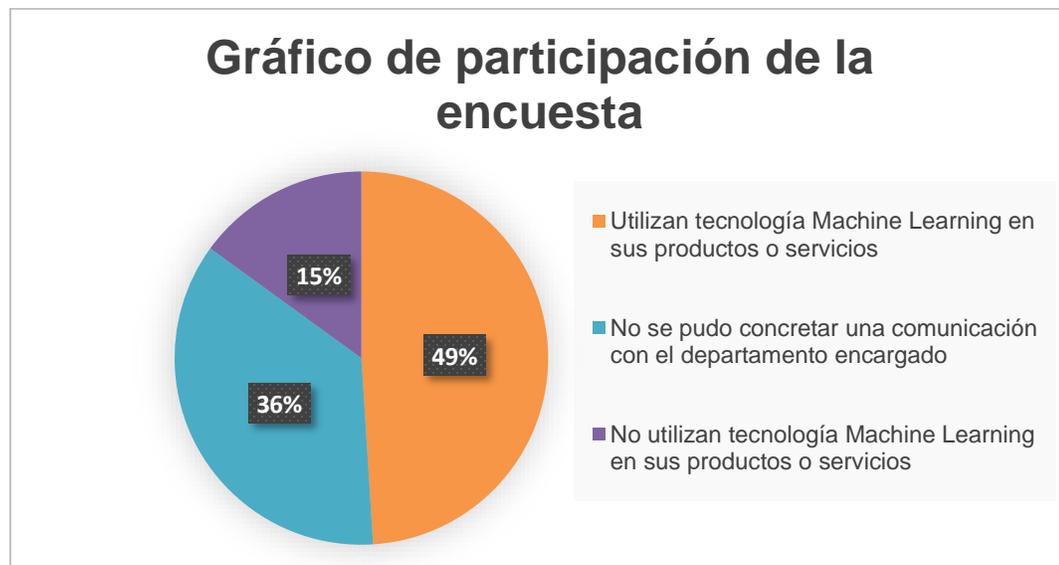


Gráfico 1 Gráfico de participación de la encuesta.

Fuente: Elaboración propia.

Después de censar una lista de 33 empresas con posibilidades de tener productos o servicios con *machine learning* se obtienen los siguientes resultados:

Según, el gráfico anterior con 36 por ciento de las empresas encuestadas después de varios esfuerzos por hacer no se logró concretar una comunicación con los departamentos encargados. Cabe mencionar que se utilizaron diferentes medios ya sea telefónico, correo electrónico y referencias, sin embargo, en algunos casos no hubo respuesta por lo que se procedió a trabajar con la información de las empresas que, si se lograron contactar, que corresponden al 49 por ciento.

Con un 49 por ciento se confirmó por medio de correo electrónico, llamada telefónica o referencia profesional, el desarrollo de *machine learning* en productos y servicios. Los casos en que no se pudo tener comunicación se validó su información por medio de los sitios web o publicaciones en periódicos digitales. (Ver apéndice Apéndice F y Apéndice J.)

Un 15 por ciento de las empresas con las cuales se estableció comunicación, negaron el uso de *machine learning* en productos o servicios. (Ver Apéndice E.)

4.1.1.2 Análisis de las características técnicas, funcionales y organizativas de la oferta de servicios.

En esta sección se analizan los datos recolectados para determinar las características técnicas, funcionales y organizativas del *machine learning* que se encuentran en la oferta de servicios tecnológicos digitales de las empresas identificadas.

4.1.1.2.1 Frameworks

El objetivo de esta sección es conocer cuáles son los *frameworks* más utilizados para desarrollar productos o servicios con *machine learning*.

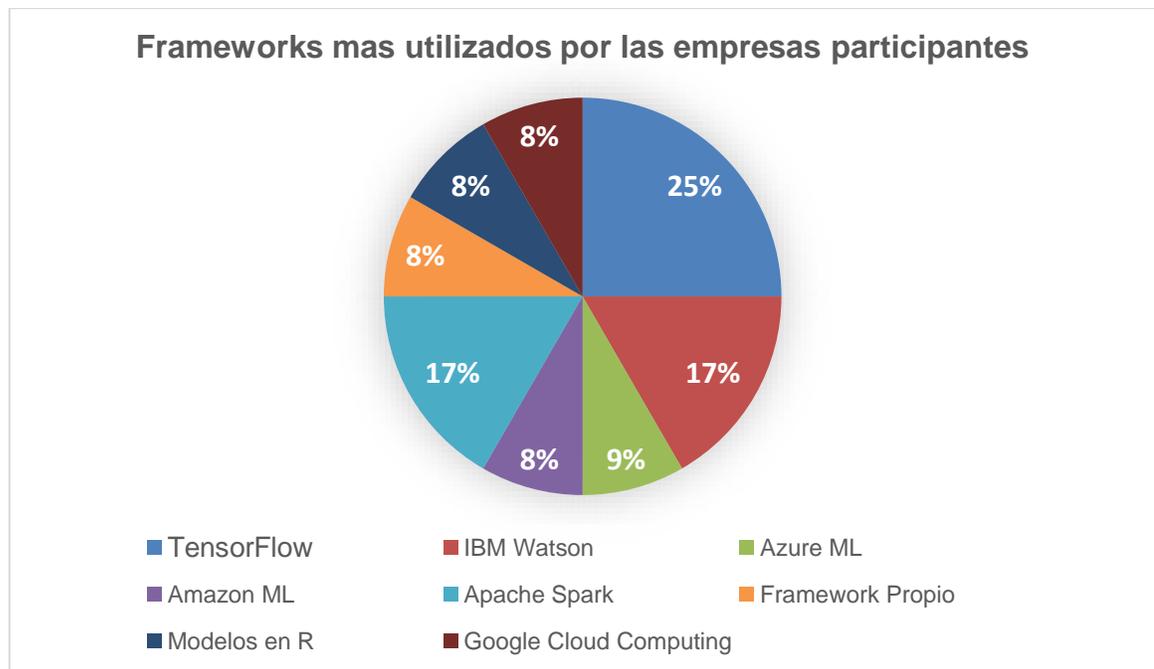


Gráfico 2 Frameworks más utilizados por las empresas participantes

Fuente: *Elaboración propia.*

Dentro de los *frameworks* más utilizados, según la investigación destaca *Tensor Flow* con un 25 por ciento, esto se debe a la flexibilidad que tiene, al buen desempeño que tiene en reconocimiento del habla e imágenes y también por ser de código abierto; el 17 por ciento señaló *Apache Spark*, el 17 por ciento *IBM Watson*, con 8 por ciento, *Amazon ML*, *Azure ML*, *Google Cloud Computing* con su herramienta de *machine learning*.

Los entrevistados mencionaron que tenían poca flexibilidad para el desarrollo de ciertas implementaciones por lo que optaban por otros *frameworks* con mayor elasticidad, pues los *frameworks* basados en un sistema en la nube contienen un limitado grupo de herramientas. También con 8 por ciento dijeron que utilizan modelos desarrollados en lenguaje R, y por último un 8 por ciento como *framework* propio de la empresa para desarrollar *machine learning*.

También expresaron que preferían un *framework* que les brinde más flexibilidad, de código abierto y bien documentado, pues a la hora de utilizarlo se necesita de tiempo de investigación y pruebas por lo que *TensorFlow* facilita mucho los desarrollos.

Jessica Califano (2018), Gerente de *Marketing* de Producto en Temboo menciona que:

“El fuerte soporte de *TensorFlow* para el aprendizaje profundo y algoritmos de aprendizaje automático es solo una de las muchas ventajas del *framework*. El soporte que le ha dado *Google* ha ocasionado que la plataforma en una de las usadas a nivel mundial”.

A continuación, se muestra una clasificación de 23 bibliotecas de *deep learning* de código abierto, según una investigación de Michael Li (2017):

“Basadas en la actividad de *Github* y *Stack Overflow*, así como en los resultados de búsqueda de *Google*. La tabla muestra los puntajes estandarizados, donde un valor de 1 significa una desviación estándar por

encima del promedio (promedio = puntaje de 0). Por ejemplo, *Caffe* es una desviación estándar por encima del promedio en la actividad de *Github*, mientras que *deeplearning4j* está cerca del promedio.

Ver a continuación los métodos:

Library	Rank	Overall	Github	Stack Overflow	Google Results
tensorflow	1	10.87	4.25	4.37	2.24
keras	2	1.93	0.61	0.83	0.48
caffe	3	1.86	1.00	0.30	0.55
theano	4	0.76	-0.16	0.36	0.55
pytorch	5	0.48	-0.20	-0.30	0.98
sonnet	6	0.43	-0.33	-0.36	1.12
mxnet	7	0.10	0.12	-0.31	0.28
torch	8	0.01	-0.15	-0.01	0.17
cntk	9	-0.02	0.10	-0.28	0.17
dlib	10	-0.60	-0.40	-0.22	0.02
caffe2	11	-0.67	-0.27	-0.36	-0.04
chainer	12	-0.70	-0.40	-0.23	-0.07
paddlepaddle	13	-0.83	-0.27	-0.37	-0.20
deeplearning4j	14	-0.89	-0.06	-0.32	-0.51
lasagne	15	-1.11	-0.38	-0.29	-0.44
bigdl	16	-1.13	-0.46	-0.37	-0.30
dynet	17	-1.25	-0.47	-0.37	-0.42
apache singa	18	-1.34	-0.50	-0.37	-0.47
nvidia digits	19	-1.39	-0.41	-0.35	-0.64
matconvnet	20	-1.41	-0.49	-0.35	-0.58
tflearn	21	-1.45	-0.23	-0.28	-0.94
nervana neon	22	-1.65	-0.39	-0.37	-0.89
opennn	23	-1.97	-0.53	-0.37	-1.07

Figura 43 Frameworks más utilizados según GitHub, StackOverflow y búsquedas de Google.

Fuente: Michael Li, *The Data Incubator* (2017).

4.1.1.2 Análisis de métodos de *Machine learning*

El objetivo de esta sección es conocer cuáles son los métodos más utilizados para desarrollar productos o servicios con *machine learning*.

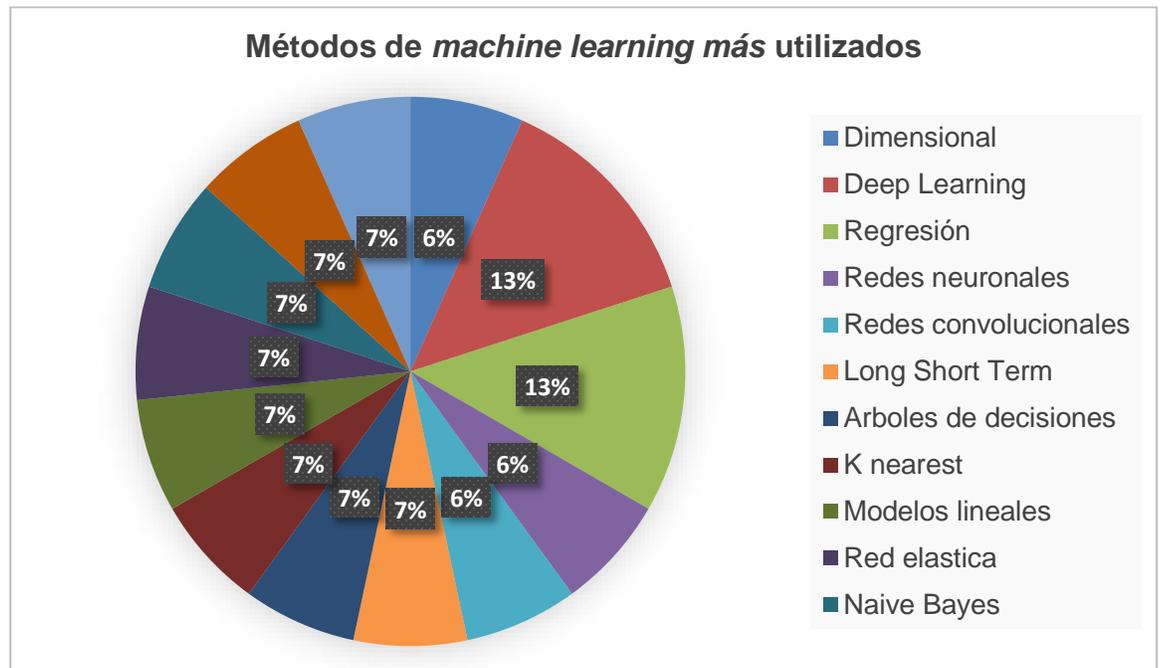


Gráfico 3 Métodos de *machine learning* más utilizados

Fuente: *Elaboración propia.*

Un 13% utilizan métodos de regresión, un 13% emplean métodos especializados en *deep learning*, por último, un 7% utilizan métodos dimensionales, redes neuronales, redes convolucionales, *Long Short Term*, Árboles de decisiones, *K nearest*, Modelos lineales, Red elástica, *Naive Bayes* para desarrollar productos o servicios con *machine learning*.

Según, *MathWorks*, (2018), desarrolladores del software *MATLAB* y *Simulink*, mencionan: “Las técnicas de regresión predicen respuestas continuas, por ejemplo, cambios en la temperatura o fluctuaciones en la demanda de energía. Las aplicaciones típicas incluyen previsión de carga de electricidad y negociación algorítmica.”

También *MathWorks* (2018), menciona que: “Los algoritmos de regresión comunes incluyen el modelo lineal, el modelo no lineal, la regularización, la regresión por pasos, los árboles de decisión reforzados y empaquetados, las redes neuronales y el aprendizaje neuro-difuso adaptativo.”

Esto se respalda en el nivel mundial, según la encuesta de *KDNuggets* (2017) del año 2017, que se puede muestra a continuación:

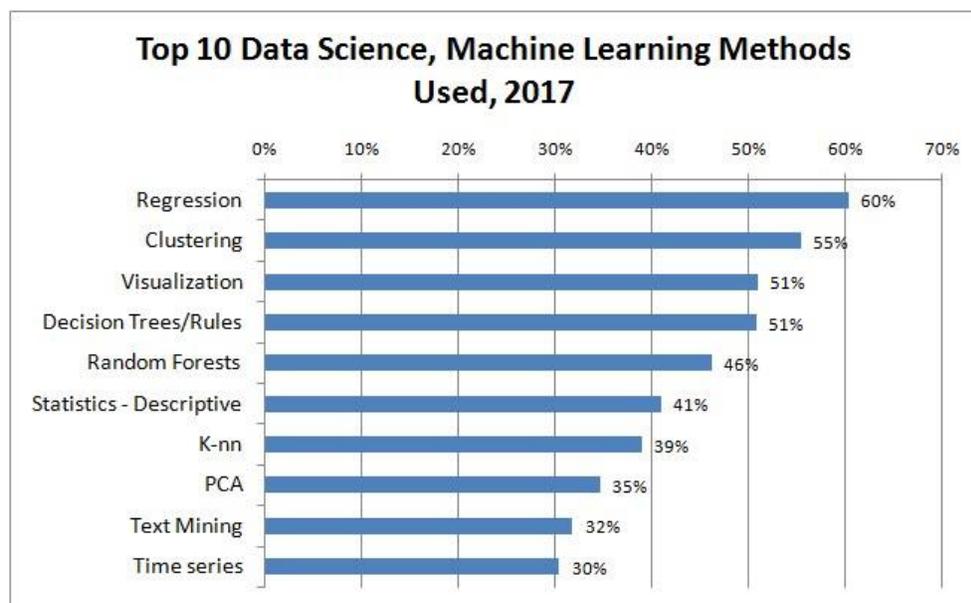


Figura 44 Top 10 Métodos de Machine learning más usados por Científicos de Datos, 2017

Fuente: *KDNuggets*, 2017.

4.1.1.2.3 Análisis de estilos de *Machine learning*.

En esta sección se busca encontrar cuáles son los estilos de *machine learning* más utilizados por las empresas.

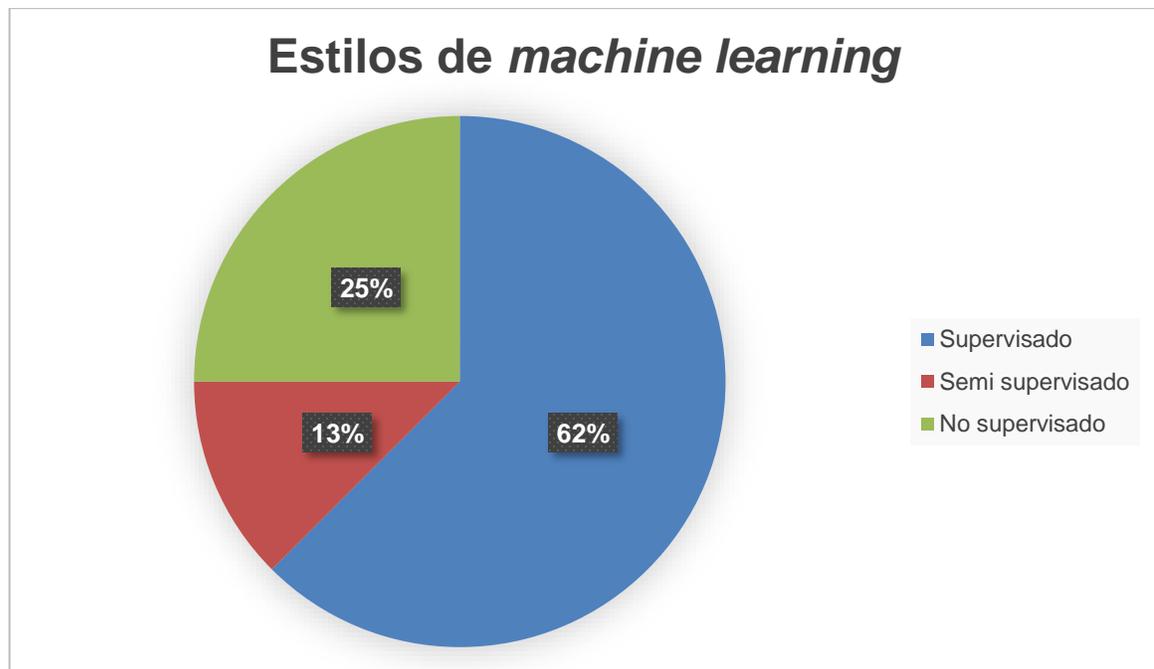


Gráfico 4 Estilos de *machine learning*

Fuente: *Elaboración propia.*

El estilo de aprendizaje más utilizado en *machine learning* con un 62%, es el estilo Supervisado, seguido con un 25% el estilo No supervisado y por último un 13% para el estilo Semi Supervisado. Mucho se debe a que los estilos permiten cierta flexibilidad y a veces hasta mejores resultados, según el tipo de problema. Los profesionales entrevistados comentaron que mucho se debe al proceso de experimentación para conocer cuál es el estilo que más se adapta a las necesidades empresariales.

4.1.1.2.4 Análisis de los tipos de enfoque de *Machine learning*.

En esta sección se tuvo como objetivo conocer los tipos de enfoques existentes de *machine learning* en las distintas empresas.

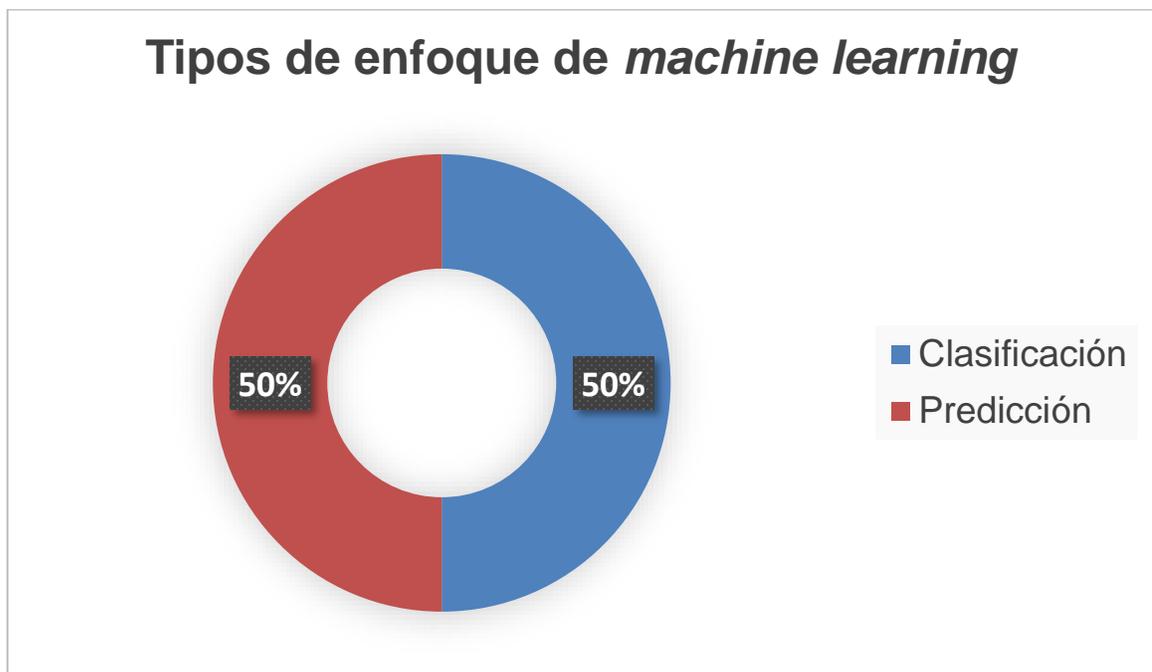


Gráfico 5 Tipos de enfoque de *machine learning*

Fuente: *Elaboración propia.*

En el caso de los tipos de enfoque, resultó una preferencia de 50% por clasificación y predicción, aunque las respuestas de esta pregunta no fueron claras o correctas; se le brindaban algunos tipos dentro de las preguntas, pero los entrevistados o encuestados no manejaban un esquema de cómo se subdivide los tipos de enfoque.

4.1.1.2.5 Áreas económicas impactadas

En esta sección se tuvo como objetivo evaluar las áreas económicas que más implementan productos o servicios con *machine learning* en las distintas empresas.

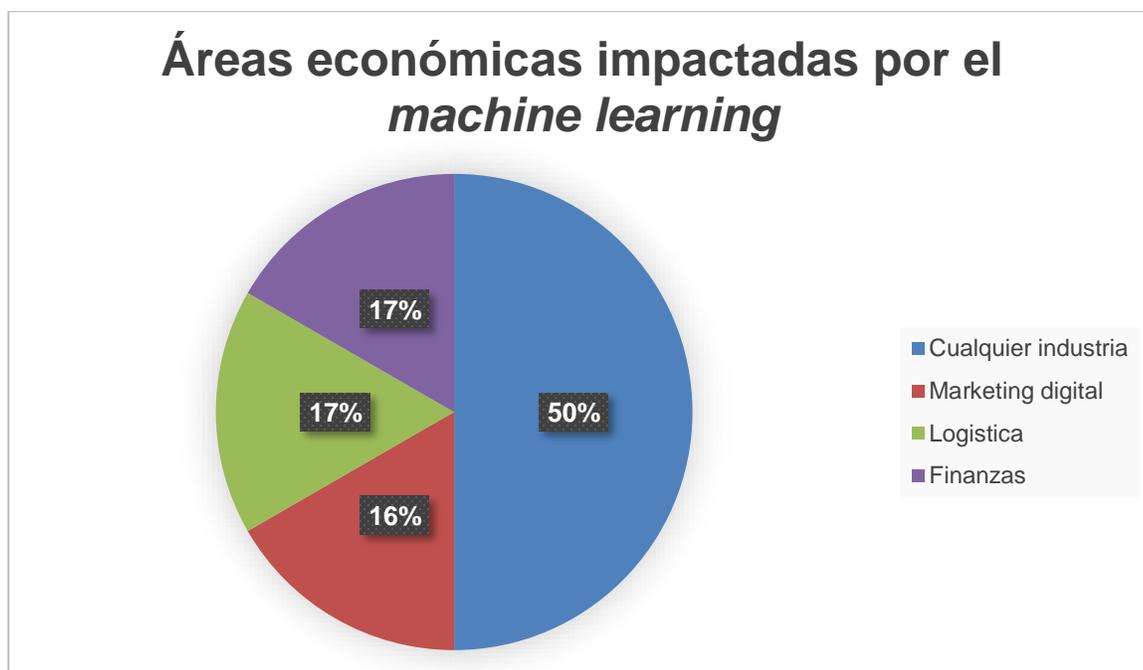


Gráfico 6 Áreas económicas impactadas por el *machine learning*

Fuente: *Elaboración propia.*

Se destaca que el 50% de las empresas mencionan que en cualquier tipo de industria o área económica puede implementarse un producto o servicio de *machine learning*, dado a que se puede personalizar, según sea la necesidad, seguido con un 14% las áreas de Logística, Finanzas y Marketing Digital.

Según McKinsey Global Institute (2016):

“La investigación anterior de MGI sobre automatización encontró que el 45 por ciento de todas las actividades laborales, asociadas con \$ 14.6 billones

de salarios en todo el mundo, tienen el potencial de ser automatizadas al adaptar la tecnología actualmente demostrada. Alrededor del 80 por ciento de eso podría implementarse mediante el uso de las capacidades existentes del *machine learning*. (pág. 15)

Esto se refuerza con las respuestas de los entrevistados que mencionan que en cualquier industria se puede automatizar procesos e incluso sacar provecho a las grandes cantidades de datos y reducción de costos.

4.1.1.2.6 Análisis de casos de uso específico.

En esta sección se realizó la encuesta para aspectos de casos de uso, con el objetivo de conocer específicamente el uso que posee en cada empresa, también categorizar los casos de uso para mayor comprensión y nichos de negocio de los casos de uso.

Según, las respuestas de los entrevistados, el reconocimiento de intenciones para el seguimiento de clientes predomina con un 17% en casos de uso para las decisiones y estrategias en marketing, mientras los siguientes casos con 8% cada uno:

- *Chatbots*.
- Dictado a texto.
- Clasificación de intenciones por medio de texto.
- Reconocimiento de intenciones.
- Proyección de ventas.

- Agentes virtuales con lenguaje natural.
- Clasificación de imágenes.
- Optimización de logística.
- Reducción de costos en finanzas.
- Predicción de demanda.
- Análisis de riesgos de créditos.

Para mayor entendimiento se puede ver los casos de usos específicos de *machine learning* en el siguiente gráfico.

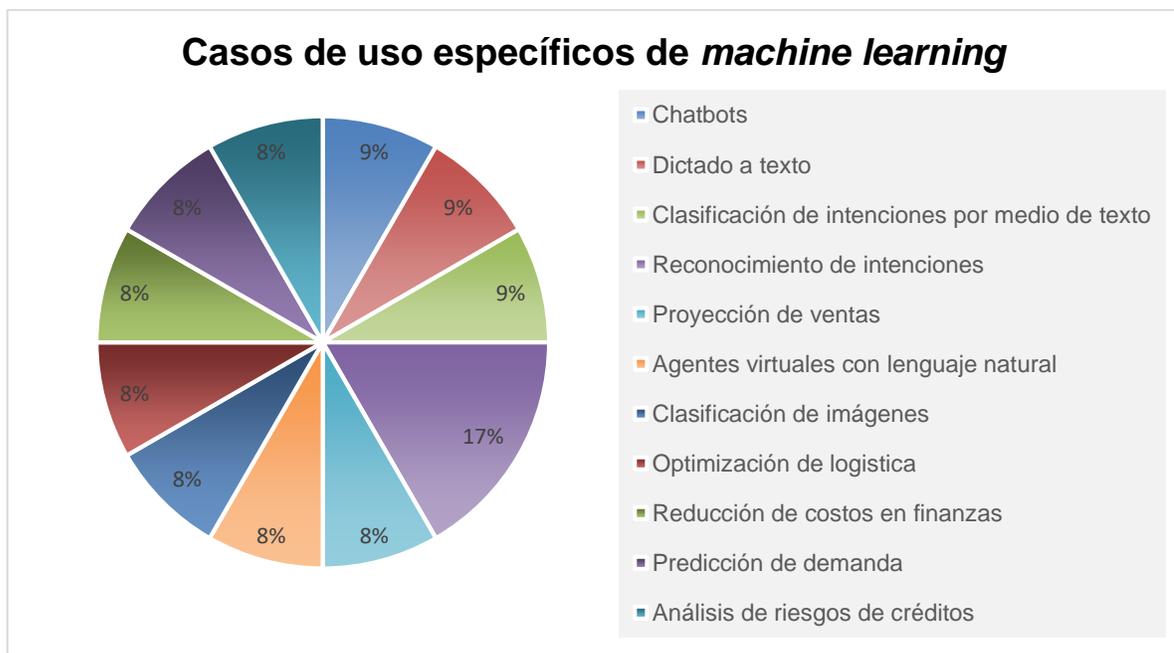


Gráfico 7 Casos de uso específicos de *machine learning*.

Fuente: *Elaboración propia.*

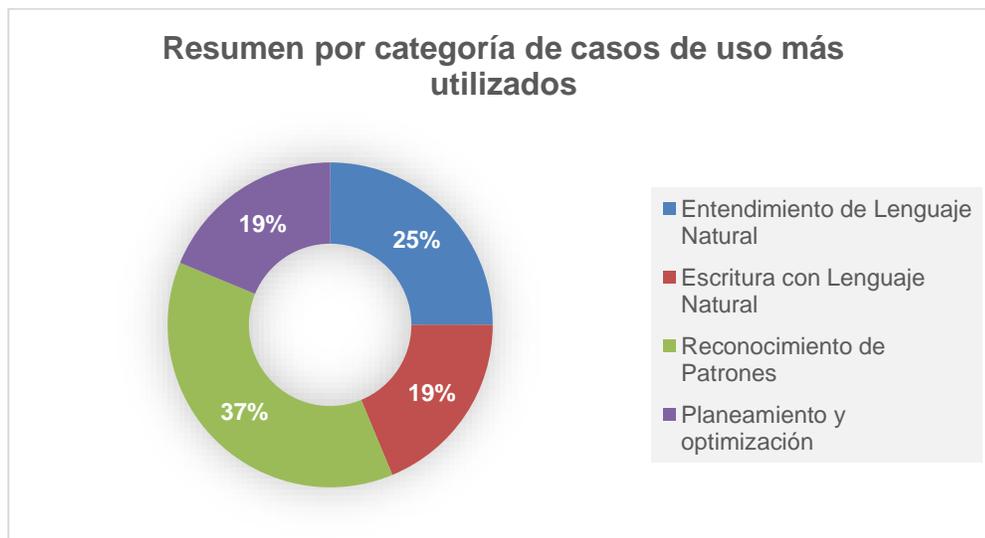


Gráfico 8 Resumen por categoría de casos de uso más utilizados

Fuente: *Elaboración propia.*

Las categorías que resumen los casos de uso para tener un mejor entendimiento de las capacidades más utilizadas del *machine learning*.

Con un 37% reconocimiento de patrones, con 25% entendimiento de lenguaje natural, con 19% escritura con lenguaje natural.



Gráfico 9 Caso de uso en que se utiliza Machine learning por nicho de negocio.

Fuente: Elaboración propia.

Con un 23 por ciento el sector económico de *marketing* utilizado en toma de decisiones estratégicas para publicidad, con 22 por ciento Inteligencia de negocios, 22 por ciento analítica *web* para el mejoramiento de interacciones y control de flujo de los usuarios, 11 por ciento comercio electrónico, 11 por ciento decodificación y clasificación de texto, igual con un 11 por ciento los *chatbots* que se han vuelto populares para suplir al ser humano en los departamento de servicio al cliente y reducir costos.

4.1.1.2.7 Ventajas de Machine learning.

El objetivo de esta sección es mostrar resultados relacionados con las ventajas expuestas por los entrevistados.

Una ventaja determinada en los resultados de *Áreas económicas* impactadas (ver 4.1.1.2.5) del cuestionario, se debe a que el *machine learning* se

puede implementar en cualquier industria y esto abre las puertas a que muchas empresas puedan optar por un modelo de negocio donde se pueda explotar los bancos de información o mejorar sus procesos.

Las ventajas de tener servicios o productos con *machine learning*, según las respuestas, resalta la optimización de procesos y la reducción de costos. También un mejoramiento de la experiencia para cliente, mejora en los tiempos de respuesta en tiempo real. Esto tiene que ver mucho con la explotación de los datos almacenados.

En algunos casos los datos se ingresan manualmente por los empleados, por lo que algunas empresas decidieron optar por un método más automático y que clasificara la información de manera inmediata.

Otra ventaja es la escalabilidad que cuenta el implementar productos o servicios con *machine learning*, por lo que los métodos y enfoques van mejorando con el tiempo, siempre y cuando el objetivo se cumpla.

- Incremento de la experiencia y comercio conversacional.
- Mejorar la eficiencia y efectividad de los procesos.
- Reducción de costos, mejora en tiempos de respuesta en procesos.
- Resolver problemas que previamente no podía abordar, explotando su data.

- Eliminar la digitación manual, eliminar la presencia humana, eliminar trabajo manual, automatizar proceso.
- Tener una ventaja competitiva al explotar sus fuentes.
- Procesos más eficientes. Encontrar potenciales clientes.

4.1.1.2.8 Retos de implementación de *Machine learning*.

El objetivo de esta sección es conocer los principales retos al implementar productos o servicios con *machine learning*, pues es importante a la hora de tomar una decisión empresarial de implementar esta tecnología en sus procesos o servicios y el balance que se tiene que cumplir para la obtención de beneficios.

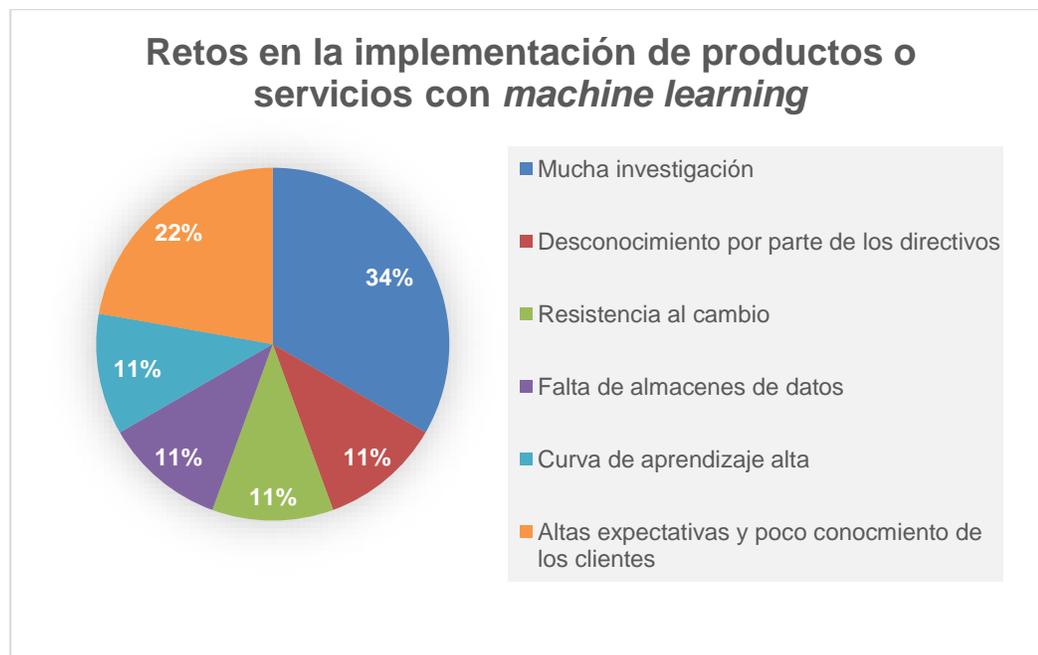


Gráfico 10 Retos en la implementación de productos o servicios con *machine learning*.

Fuente: *Elaboración propia.*

Con un 34%, el tema de la investigación fue mencionada, por lo que requiere de tiempo y pruebas para que se pueda lograr alguno de los objetivos de reducción y optimización de procesos.

Con un 22% las altas expectativas y poco conocimiento de los clientes de la tecnología *machine learning*, los entrevistados mencionaron que los clientes tienen altas expectativas respecto de los resultados y no son conscientes que a veces no se arrojan buenos resultados o requiere de más investigación para tener una mejora real.

Por esto se requiere una educación y comunicación con los clientes que pagan por este tipo de implementaciones y que muchas veces hacen grandes inversiones con la mentalidad de recibir un reingreso inmediato, incluso de manera mágica.

Con un 11 por ciento sobresale el desconocimiento de los directivos, los entrevistados dicen que: muchas de las implementaciones surgen a raíz de la innovación del departamento de desarrollo de software, pero con el impedimento o poco patrocinio de los directivos de las empresas que desconocen el potencial que existe al invertir tiempo y recursos en implementación con *machine learning*.

Con un 11 por ciento, la resistencia al cambio, esto se vincula a lo mencionado donde los directivos tienen el desconocimiento de cómo mejorar sus procesos o productos por medio de la tecnología *machine learning*, pues existe resistencia a incursionar en una implementación de ese tipo y como se menciona

en el nivel global, los primeros en iniciar un cambio e invertir en nuevas tecnologías tendrán mejores beneficios por la ola o revolución como muchos la llaman del *machine learning*.

Como menciona Enrique Martín experto en tecnología de Oracle Ibérica (2017):” S existe un riesgo con el *machine learning* es ignorarlo. 2018 debería ser el año en que las empresas se comprometan, si aún no lo han hecho, a explorar y desbloquear el valor del *machine learning*.”

El 11 por ciento indica la falta de almacenes de datos que proporcionen el combustible necesario para analizar y sacar provecho por medio de esta tecnología. Es por eso por lo que recomiendan que los datos se almacenen de forma ordenada y que estos datos a futuro puedan servir como una oportunidad para la toma de decisiones.

Otro 11 por ciento, se refirió a la curva de aprendizaje para llevar a cabo proyectos con *machine learning*, por lo que el personal tendrá que capacitarse y requerirá de tiempo, aunque todo depende de las bases académicas que tenga el recurso humano.

4.1.1.2.9 Tipo de consumo de *machine learning*.

En esta sección se busca conocer cuál es el tipo de consumo, ya sea de manera local o en el extranjero, pues muchas de las ventajas del *machine learning* puede provocar una mejora en los servicios o procesos.

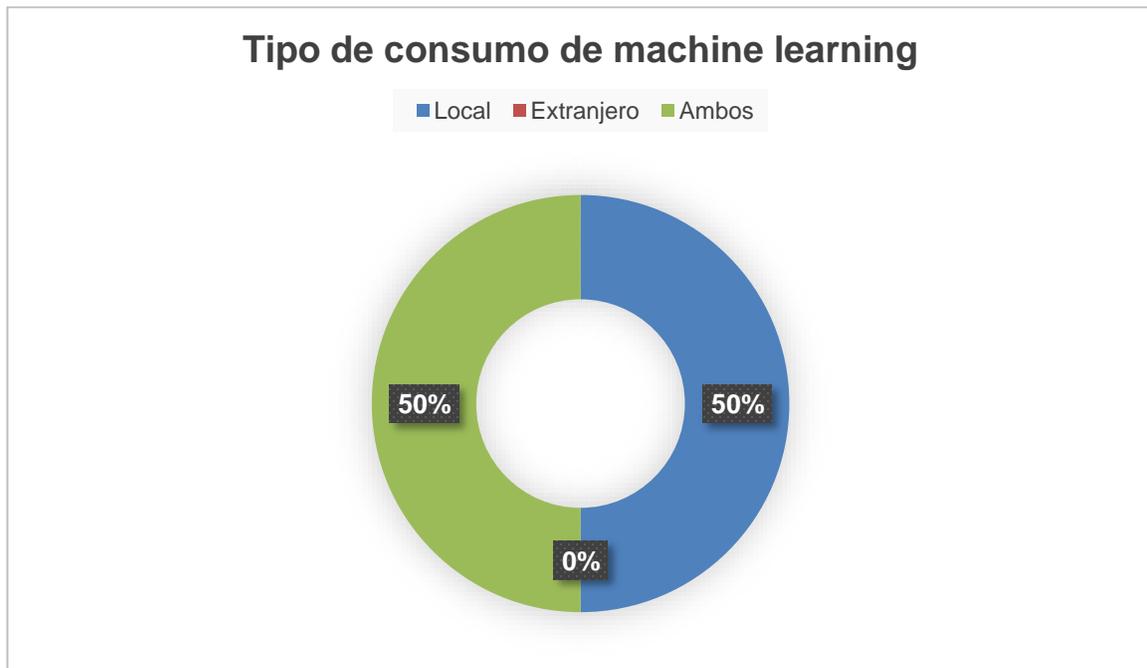


Gráfico 11 Tipo de consumo de machine learning

Fuente: *Elaboración propia.*

Según los resultados el tipo de consumo es de 50% de manera local, cabe especificar que uno de los entrevistados menciona que el consumo es a lo interno de las empresas y están enfocados en la mejora en procesos, la reducción de costos y su posicionamiento en el mercado.

Luego con un 50% los entrevistados mencionaron que se da en ambos casos de manera local y en el extranjero, pues las plataformas que desarrollan se venden en el nivel nacional y en el extranjero.

4.1.1.2.10 Necesidades o retos principales para desarrollar *Machine learning*.

En esta sección se analiza cuáles son las principales necesidades o retos para desarrollar *machine learning*. Esto es de suma importancia pues si una

empresa desea implementar procesos con esta tecnología tendrá que lidiar con ciertas necesidades o impedimentos que cruzar para tener un buen despliegue.

Una de las necesidades encontradas es el deseo de innovar que impulsa a cambiar la percepción del modelo de desarrollo tecnológico, donde existe una gran ola tecnológica que no hay que dejar pasar.

Otra es la falta de personal capacitado en especialidades relacionadas con *machine learning*, sumado a esto la zona de *comfort* que existe en los desarrolladores por no querer ingresar en el área de automatización y análisis de datos.

El patrocinio de los altos mandos en las empresas es sumamente necesario para lograr un avance en cómo se desarrollan productos o servicios, es mencionado por los entrevistados, pues muchas de las innovaciones surgieron por parte del departamento de ingeniería en *software* y no por una visión superior en las empresas. Los prototipos deben tener la aprobación de los superiores representa un impedimento al avance de esta tecnología por lo que las decisiones son difíciles de tomar, tampoco hay que decir que *machine learning* es un acto de magia, pero si es necesario un conocimiento supremo, investigación y mucha práctica.

Los costos de especialistas en esta rama resultan altos por lo que la tendencia por abaratar costos es contratar ingenieros en *software* con algún conocimiento en *machine learning* y luego que estos generen conocimiento por sí solos o de manera personal. Esto también se concatena con la falta de producción

de profesionales en esa área por parte de la academia, este tema se ahonda más en la sección 4.3 de análisis de la academia.

4.1.1.2.11 Proyección del desarrollo tecnológico de *Machine learning*.

En esta sección se quiere conocer el punto de vista de los entrevistados en cuanto al desarrollo de la tecnología *machine learning* en los siguientes años y de qué manera se adoptaría este modelo tecnológico en el país.

Uno de los entrevistados mencionó lo siguiente: se espera el despliegue de soluciones más sencillas; el conocimiento y la experiencia está en crecimiento por lo que los modelos de *machine learning* serán cada vez más sencillos y atinados, según los requerimientos.

Esto sumado a la mayor madurez del mercado para entender las ventajas de proyectos con *machine learning* provocando un incremento de la adopción de nuevas tecnologías.

La comunidad está creciendo y es posible que surjan muchos emprendimientos que provoquen un mayor desarrollo y despliegue de actividades económicas enfocadas en *machine learning*.

También se menciona que se espera una mayor diversificación en otras líneas de negocio como el reconocimiento de imágenes y continuar con el procesamiento de texto a un mejor nivel.

Por otro lado, uno de los entrevistados mencionó que el panorama es difícil, no se conoce mucho del medio, mucha gente no sabe del tema, además esperará ver más *startups* brindando servicios en el nivel nacional o en el extranjero y esperará que las empresas generen departamentos de científicos de datos como tal.

4.1.1.2.12 Recurso humano

En esta sección se tiene como objetivo conocer cuáles son los puestos de trabajo que contratan para producir productos o servicios con *machine learning*.



Gráfico 12 Tipos de profesiones contratados para desarrollo de Machine learning

Fuente: *Elaboración propia.*

El principal recurso humano que contratan para generar equipos para desarrollo e implementación de tecnología de *machine learning* son los ingenieros

en software con un 55%; esto porque hay gran cantidad de profesionales que proporcionan las universidades.

Con un 18%, especialistas en *machine learning*, que cuentan con experiencia y aumentan la velocidad de desarrollo de los proyectos. Con 9%, ingenieros eléctricos que cuentan con énfasis en computación también son buscados para formar equipos. Con 9% se menciona gerencia en ventas, esto debido a que hay que saber vender estos productos, por su novedad muchos clientes lo desconocen. Con otro 9%, los físicos, este grupo es pequeño, pero se logra conocer que conforman un equipo con los ingenieros en software.

Los equipos para desarrollo de *machine learning* por lo general se ha visto que son en su totalidad de ingeniería en software, pero en algunos casos estos trabajan en conjunto con algún otro tipo de profesional, como ingenieros eléctricos, ingenieros en estadística y físicos, para producir un modelo que cumpla los requerimientos establecidos. Esto denota lo diversificado que es el campo del *machine learning* y que es una situación importante, pues se puede implementar en muchas áreas.

4.1.1.2.13 Capacitación de personal

En esta sección se busca conocer cuáles son los métodos de capacitación más usados por el recurso humano encargado en desarrollar productos o servicios con *machine learning*.

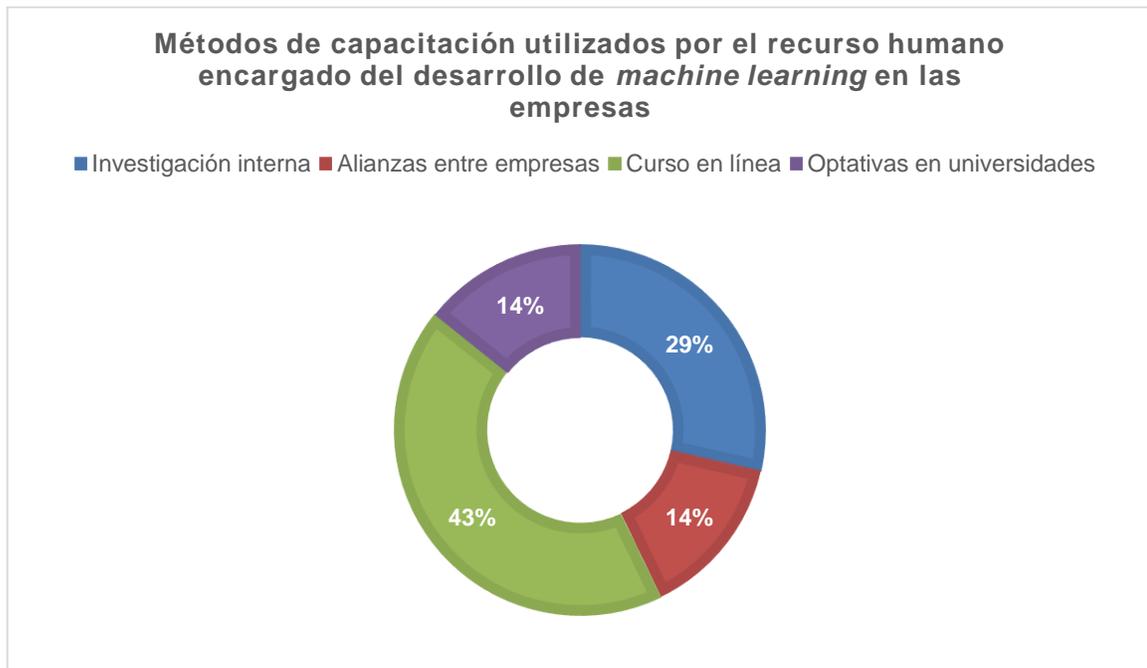


Gráfico 13 Métodos de capacitación utilizados por el recurso humano encargado del desarrollo de machine learning en las empresas

Fuente: *Elaboración propia.*

Con un 43%, los cursos en línea dominan la forma de capacitarse en *machine learning*, de los mencionados destaca *Coursera*, *Udacity*, *PluralSight*, muchos de estos sitios respaldados por prestigiosas universidades de nivel mundial y el contenido es apropiado para el mejoramiento de las habilidades en *machine learning*.

El 29%, opina que con investigación interna genera conocimiento para el futuro, así como recursos de publicaciones científicas con *machine learning*. El 14%, destaca alianzas entre empresas proporcionando todo el conocimiento para capacitarse, este es el caso que ofrece IBM con su sección especializada de IBM Watson.

El 14%, comenta la importancia de las materias optativas en las universidades, este es el caso en la Universidad Nacional de Costa Rica, en la carrera de Ingeniería en Sistemas de Información, donde imparten minería de datos y que forma parte del multidisciplinario campo del *machine learning*.

4.2 Informe sobre la relevancia entre las características técnicas, funcionales y organizativa respecto de las áreas económicas.

En esta sección se detallan los resultados expuestos y se analiza las características del *machine learning* y su impacto en las áreas económicas o industrias.

Según lo expuesto, se puede establecer que la cantidad de *frameworks*, métodos, estilos y tipos de enfoques, permiten una serie de combinaciones que facilita el impacto o su adición en la oferta de servicios o para uso interno de las empresas en una variedad de áreas económicas o industrias. Además, cabe destacar que el tipo de consumo evidencia que es un servicio exportable.

En el nivel organizativo los servicios tecnológicos digitales basados en *machine learning* encontrados en la oferta presentan una serie de impedimentos o retos que disminuyen la capacidad de desarrollo en el área económica, como, por ejemplo, la falta de respaldo de altos mandos o ejecutivos, resistencia al cambio y una gran cantidad de tiempo de investigación para lograr una solución que se ajuste a las necesidades.

En la Tabla 14 se detalla la respuesta de la empresa Neustar y como combinan características técnicas del *machine learning* para solventar una necesidad de *marketing digital*.

En este caso en particular la utilización de redes neuronales permite el análisis de las palabras y lenguaje natural para clasificar el contenido y con la ventaja que mencionó la empresa Neustar en la encuesta: “Eliminar digitación manual, eliminar presencia humana, automatizar procesos humanos y captura de datos”. La eliminación humana es un factor que puede perjudicar futuros empleos y como el *machine learning* impacta la economía.

Marketing Digital				
Framework	Métodos	Estilos	Tipo de enfoque	Área económica
<i>Tensor Flow, Amazon ML, Apache spark</i>	Redes neuronales, <i>deep learning</i> , convoluciones, <i>long short term</i>	Supervisado	Clasificación de entidades y lenguaje natural	Mercadeo

Tabla 14 Respuesta de Neustar con impacto en el área económica de mercadeo o marketing digital.

Fuente: Elaboración propia.

En la siguiente tabla se expone otro caso como el de la empresa Singularities que presenta diferentes tipos de *frameworks*, una gran cantidad de métodos, dos tipos de estilos y un enfoque a la clasificación o predicción. Respecto de las áreas económicas que impacta el servicio digital que ofertan se indica que puede aplicar a múltiples áreas, pero detallan en áreas como la financiera y de logística.

Finanzas y Logística				
Framework	Métodos	Estilos	Tipo de enfoque	Área económica
<i>Apache Spark ML, Framework propio de la empresa, modelos en R, TensorFlow.</i>	<i>Dimensional, Regression, Instance-based algorithms, deep learning, K-nearest Árboles de decisiones, k nearest, regression logística, Ordinary least squares regression. Regresión lineal, modelos lineales generalizados con distribuciones, lasso, elastic net, naive bayes, support vector machines, métodos de árboles. Dimnesionality reduction no se usa mucho por el poder computacional. Bootstrap aggregating, ensemble averaging, random forest, Gradient boosted, decision tree y Association rule learning.</i>	<i>Supervisado, No supervisado.</i>	<i>Clasificación y predicción. Métodos predictivos.</i>	<i>En múltiples áreas, logística y finanzas.</i>

Tabla 15 Respuesta de la empresa Singularities y su impacto en las áreas económicas como la financiera y logística.

Fuente: *Elaboración propia.*

Además, según la información que proporcionó la empresa Singularities, los casos de uso van desde la optimización en tiempo real hasta la analítica para la toma de decisiones. Respecto del área de logística, se utiliza para la optimización

de procesos para reducir costos y en el área de finanzas funciona para el análisis de tendencias de clientes y generar análisis de riesgos de créditos.

Estos resultados permiten comprender como el *machine learning* impacta en las áreas económicas y como se presentan efectos positivos como negativos que hay que tomar en cuenta.

4.3 Análisis de la academia.

Esta sección tiene como objetivo el análisis de la oferta y los esfuerzos de la academia para producir profesionales que puedan crear productos o servicios con *machine learning*.

La academia tiene un alto grado de responsabilidad de producir profesionales para la innovación y el área tecnológica, por ello esta ola de sistema cognitivos tiene que abordarse de una manera muy específica.

Se analizan los planes de estudio de las carreras o centros de investigación de las universidades estatales de Costa Rica que impulsen el desarrollo de la tecnología de *machine learning*.

Las facultades de interés para su análisis se tomaron en cuenta las ingenierías y ciencias básicas que tienen una alta relación con el campo de aplicación del *machine learning*.

4.3.1 Universidades estatales de Costa Rica.

Las universidades estatales de Costa Rica son: Universidad Técnica Nacional, Universidad de Costa Rica, Universidad Nacional, Universidad Nacional, Instituto Tecnológico de Costa Rica y la Universidad Estatal a Distancia.

4.3.1.1 Universidad de Costa Rica

En cuanto a la Universidad de Costa Rica se analizaron los planes de estudio de las facultades de Ingeniería, Ciencias básicas, Salud, Ciencias agroalimentarias para conocer alguna relevancia con *machine learning*.

4.3.1.1.1 Escuela de Ingeniería eléctrica

En la escuela de Ingeniería eléctrica de la Universidad de Costa Rica se encontró en su plan de estudio una materia optativa llamada Visión por computadora.

Visión por computadora es según la *British Machine Vision Association* y *Society for Pattern Recognition*:

“La visión por computadora tiene que ver con la extracción automática, el análisis y la comprensión de información útil de una sola imagen o una secuencia de imágenes. Implica el desarrollo de una base teórica y algorítmica para lograr una comprensión visual automática. (2018).

La visión por computador se asocia al *machine learning*, por lo que su enseñanza forma una base para este tipo de aplicación.

Curso	Código
Visión por computadora	IE0449

Tabla 16 Cursos optativos Escuela de Ingeniería Eléctrica, UCR.

Fuente: Planes de estudio de la Facultad de Ingeniería (2018), ver Anexo 1.

4.3.1.1.1 Programa de Posgrado en Ingeniería Eléctrica (PPIE)

La Escuela de Ingeniería Eléctrica cuenta con un programa de posgrado detalla:

“Desea desarrollar en los estudiantes habilidades de investigación, en donde este aprenda y ponga en práctica el método científico usado en Ingeniería Eléctrica, en pos de mejorar sustancialmente su capacidad de generar conocimiento, resolver problemas, plasmar sus hallazgos en artículos científicos y comunicarlos a la sociedad. Para lograr este fin, se ofrece la Maestría Académica en Ingeniería Eléctrica, con apoyo de la Escuela de Ingeniería Eléctrica. (2018)

Las maestrías se desarrollan en laboratorios de investigación con tecnología de punta.

Laboratorio	Descripción
LIIB: Laboratorio de investigación en ingeniería biomédica.	Trabaja en aplicar soluciones de ingeniería a problemas relacionados con la medicina y la salud. El proyecto principal actual es utilizar una “Interfaz Cerebro-Computador” para leer señales eléctricas del cerebro, transmitir las a

	<p>un computador, y utilizarlas para controlar algún dispositivo. Se pretende, en efecto, controlar algún dispositivo solamente por medio del pensamiento.</p>
<p>IPCV-LAB: Laboratorio de investigación en procesamiento digital de imágenes y visión por computador.</p>	<p>Realiza investigación básica en formación, percepción, captura, digitalización y modelado de señales de video, para el desarrollo de nuevos algoritmos capaces de:</p> <p>Estimar el movimiento tridimensional de robots a partir del análisis de una o más señales de video provenientes de cámaras atadas rígidamente a su estructura (odometría visual monocular y estereoscópica), para la navegación autónoma de robots móviles terrestres o voladores de exploración planetaria, la estabilización digital de las señales de video capturadas por los <i>robots</i> y la generación de panoramas a partir de esas señales</p> <p>Estimar el movimiento tridimensional de una persona a partir del análisis de una o más señales de video provenientes de cámaras colocadas frente a esta, para la teleoperación de <i>robots</i> espaciales</p>

	<p>antropomórficos, el control de avatares en juegos electrónicos y la transmisión de señales de video a muy bajos bit-rates</p> <p>Analizar imágenes biomédicas capturadas por microscopios para el monitoreo y diagnóstico de enfermedades.</p> <p>Analizar imágenes de productos industriales capturadas por cámaras colocadas sobre las líneas de producción para la detección temprana de errores en los productos.</p>
<p>ARCOS-Lab: Laboratorio de investigación en robots autónomos y sistemas cognitivo.</p>	<p>Trabaja en la creación y mejoramiento de sistemas cognitivos y <i>robots</i> autónomos que permitan ayudarle a los humanos en tareas tediosas, peligrosas o sumamente complicadas que requieren de capacidades avanzadas de inteligencia cognitiva. Algunos ejemplos de este tipo de tareas son:</p> <p>Instalación y mantenimiento de equipo peligroso o en áreas peligrosas.</p> <p>Tareas repetitivas difíciles que causan enfermedades en las personas.</p> <p>Búsqueda y rescate</p>

	Evaluación de situaciones peligrosas Monitoreo
PRIS-Lab: Laboratorio de investigación en reconocimiento de patrones y sistemas inteligentes.	Su principal objetivo es la investigación de nuevos algoritmos de procesamiento de la información para el modelado de procesos de interés en diversas áreas del conocimiento. Sus principales intereses son: reconocimiento de patrones, sistemas inteligentes, aprendizaje automático, minado de datos, procesamiento digital de señales, procesamiento digital de imágenes, rastreo de objetos, estructuras abstractas de datos, algoritmos y programación.

Tabla 17 Laboratorios del Programa de Posgrado en Ingeniería Eléctrica (PPIE).

Fuente: Laboratorios de investigación, Programa de posgrado en ingeniería eléctrica (2018).

Cabe destacar que la duración es de dos años y se requiere de dedicación de tiempo completo por parte del estudiantado.

4.3.1.1.1.1 Pris-Lab

El *machine learning* es aplicable en muchas áreas y se descubrió que el Laboratorio de Reconocimiento de Patrones y Sistemas Inteligentes (PRIS-Lab) de la Escuela de Ingeniería Eléctrica (EIE) de la Universidad de Costa Rica (UCR), cuenta con distintos enfoques que tienden a mezclarse para lograr proyectos de innovación y mejorar los conocimientos en *machine learning*.

En su sitio web <https://pris.eie.ucr.ac.cr/>, explican cómo nació esta propuesta:

El Laboratorio de Reconocimiento de Patrones y Sistemas Inteligentes es un laboratorio transdisciplinar que forma parte de la Escuela de Ingeniería Eléctrica de la Universidad de Costa Rica. Fue fundado en 2012 por el Dr. rer. nat. Francisco Siles Canales como un núcleo de formación entre distintas disciplinas. Busca la generación de investigación científica e innovación tecnológica, en pos de la integración entre la Comunidad Académica, el Gobierno, el Sector Productivo, la Sociedad Civil y el Medio Ambiente, para el mejoramiento de la calidad de vida de las personas. (2018)

El Pris-Lab cuenta con 83 colaboradores de 18 carreras en áreas como Ingeniería, Artes, Letras, Ciencias Sociales, Ciencias Básicas. (Alfani, 2017).

Dentro de los equipos de investigación se encuentran:

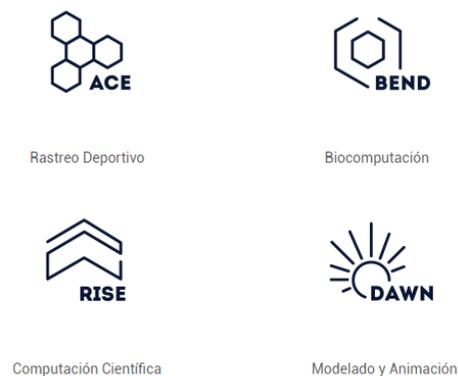


Figura 45. Equipos de investigación Pris-LAB EIE UCR.

Fuente: Sitio Oficial Pris-Lab (2018).



Figura 46. Equipos de investigación Pris-LAB EIE UCR. (UCR E. d., 2018)

Fuente: Sitio Oficial Pris-Lab (2018).

Dentro de los grupos donde se nota el uso de *machine learning* destaca el grupo RISE de Computación Científica:

“La computación científica diseña e implementa nuevos algoritmos eficientes para el soporte de operaciones de reconocimiento de patrones y sistemas inteligentes aplicados a problemas científicos. En RISE, solucionamos problemas computacionales complejos, con la ayuda de métodos numéricos, métodos algorítmicos, computación distribuida y paralela, así como la visualización científica para la exploración de datos”. (2018)

También se menciona que dentro de los proyectos establecidos por el equipo RISE (2018) está: “el análisis de video de alta resolución, simulación de modelos no lineales y la evaluación de métodos de aprendizaje automático.”

Dentro del equipo RISE hay estudiantes de Ingeniería eléctrica, Ingeniería política y Biología tropical. (2018)

4.3.1.1.2 Escuela de Ingeniería ciencias de la computación e informática.

En la Escuela de Ingeniería Ciencias de la Computación e Informática, se analizó el plan de estudio y se encontró que en el énfasis en Ciencias de la computación un curso que tiene como título Inteligencia artificial, aunque no se logra constatar la profundidad de este curso, se puede contemplar como base introductoria para el aprendizaje de *machine learning*.

Énfasis	Nivel	Nombre y código
CIENCIAS DE LA COMPUTACIÓN.	Bachillerato	INTELIGENCIA ARTIFICIAL CI0129

Tabla 18 Cursos del énfasis de ciencias de la computación, UCR.

Fuente: Planes de estudio de la Escuela de Ciencias de la Computación e Informática con varios énfasis. (2018), ver Anexo 2.

También se identifica algunos cursos optativos en los planes de Bachillerato y Licenciatura. Cursos como concretamente aprendizaje automático (*machine learning*), esto es bueno recalcar que existe una iniciativa por ir a la vanguardia. (Ver Anexo 3 y Anexo 4).

Curso	Código
Aprendizaje automático	CI2603
Sistemas expertos	CI2606
Inteligencia artificial	CI2651
Redes neuronales	CI2655

Reconocimiento, análisis y clasificación de formas	CI2654
--	--------

Tabla 19 Cursos optativos en planes de Bachillerato y Licenciatura con referencia a machine learning.

Fuente: Planes de estudio de la Facultad de Ingeniería, Ciencias de la Computación e Informática, Bachillerato y Licenciatura, (2018), ver Anexo 3 y Anexo 4.

4.3.1.2 Universidad Nacional

En la Universidad Nacional, se investigaron los planes de las facultades de Tecnología e Ingeniería que tenían relevancia con *machine learning*.

4.3.1.2.1 Escuela de Informática, Ingeniería en Sistemas de Información (Bachillerato y Licenciatura)

En la Escuela de Informática, Ingeniería en Sistemas de Información, en su plan de estudio se encuentran algunos cursos relevantes con *machine learning*. Es importante mencionar que uno de los entrevistados en esta investigación menciona que estos cursos optativos le fueron de mucha ayuda a la hora de incursionar en *machine learning*. (Ver Anexo 5 y Anexo 6).

Curso	Código
Inteligencia artificial	EIF420O
Minería de datos I	EIF434O

Minería de datos II	EIF435O
---------------------	---------

Tabla 20 Cursos optativos en planes de Bachillerato y Licenciatura con referencia a machine learning.

Fuente: Plan de Estudios Ingeniería en Sistemas con Grado Bachillerato y Salida Lateral de Diplomado. (2018) Ver Anexo 5 y Anexo 6.

4.3.1.3 Universidad Técnica Nacional

Se analizaron las carreras de ingenierías y tecnologías, pero dentro de los planes académicos desde diplomado, bachillerato y licenciatura no se halló ningún curso que contenga relevancia con *machine learning*. Además, no existe un grupo de investigación que está desarrollando con algún elemento de *machine learning* o inteligencia artificial.

No se hizo hallazgo de relevancia con *machine learning* en los planes de estudio. El análisis se basó en los planes de estudio que están publicados en el sitio oficial de la Universidad Técnica Nacional: <http://www.utn.ac.cr/carreras>.

4.3.1.4 Instituto Tecnológico de Costa Rica

Para el análisis de planes de estudio se tomaron en cuenta carreras tecnológicas e ingenierías.

El plan de Bachillerato en Ingeniería en computación se encontró un curso en inteligencia artificial que cabe mencionar no se profundizó si dicho curso abarca concretamente *machine learning*, pero sirve de base para la introducción a un campo tan multidisciplinario como lo es *machine learning*. (Ver Anexo 7).

Nombre del curso	Código
INTELIGENCIA ARTIFICIAL	IC6200

Tabla 21 El plan de Bachillerato en ingeniería en computación, relevancia de materia con *machine learning*.

Fuente: El plan de Bachillerato en ingeniería en computación, TEC (2018), (ver Anexo 7).

4.3.1.4.1 Parma Group

Pattern Recognition and Machine Learning Group es un grupo de investigación con sede en el Tecnológico de Costa Rica orientado a trabajar en el área de *machine learning* compuesto por profesionales de distintas áreas de trabajo que busca la generación de investigación científica, innovación tecnológica e integración con los distintos sectores productivos del país.

Este grupo genera proyectos de investigación utilizando técnicas de reconocimiento de patrones y *machine learning*, tales como:

Organización de simposios internacionales sobre aplicaciones del aprendizaje automático cuyo objetivo es (2017): “Estimular la investigación y desarrollo en el país en temas afines al aprendizaje automático (*machine learning*), vinculando los sectores académico, gubernamental y empresarial.”

También la creación de escuelas de verano para aprender y mejorar conocimientos en reconocimiento de patrones y *machine learning*.

El grupo PARMA trabaja en oportunidades para el *machine learning* en el combate del cáncer de mama con el Dr. José Luis Quirós, Patólogo quien trabaja

en el hospital Max Peralta, y además es profesor/investigador de la Universidad de Costa Rica.

Otro de muchos proyectos está en el análisis y mejoramiento de lectura de radiografías:

En Costa Rica, especialistas del grupo de Reconocimiento de patrones y aprendizaje automatizado (PARMA, por sus siglas en inglés) del Tecnológico de Costa Rica (TEC) trabajan en un proyecto que analiza radiografías para determinar la edad de infantes y desarrollan una aplicación que le permitirá a médicos diagnosticar trastornos de crecimiento (2018).

4.3.1.5 Universidad Estatal a Distancia

Los objetivos de análisis fueron los planes de estudio de carreras de ingeniería y relacionadas con la tecnología en el nivel de diplomado, bachillerato, licenciatura, maestría y doctorado. para conocer si hay alguna iniciativa por impulsar la enseñanza de *machine learning*.

No se hizo hallazgo de relevancia con *machine learning* en los planes de estudio. Además, no existe un grupo de investigación que tenga desarrollos en *machine learning* o inteligencia artificial como sucede en universidades como la Universidad de Costa Rica o el Tecnológico de Costa Rica.

El análisis se basó en los planes de estudio que están publicado en el sitio oficial de la Universidad Estatal a Distancia: <https://www.uned.ac.cr/carreras-uned>.

4.3.2 Centros privados de enseñanza en correlación con *machine learning*.

Se toman en cuenta centros de enseñanza privada que los profesionales utilizan para adquirir bases de conocimiento de *machine learning*.

4.3.2.1 Programa Iberoamericano de Formación en minería de datos. (Promidat)

El Programa Iberoamericano de formación en minería de datos es (2018):
“Una empresa que se dedica a la capacitación y formación en minería de datos, tanto virtual como presencial, con especial énfasis en la capacitación del personal de las empresas y en los procesos de análisis de datos de estas.”

Los cursos que ofrecen pueden tomarse en diferente orden y lo certifican como Experto en minería de datos al completar los 10 cursos.

Estos cursos están orientados, según Promidat:

“Se orienta especialmente a analistas de riesgo, *marketing* y mercadeo, personal involucrado en proyectos de pronóstico y predicción o en proyectos de “Business Intelligence”, administradores de bases de datos, analistas de sistemas, profesionales en estadística o economía y en general a personas que, basados en los datos de la organización, colaboran en la toma de decisiones. (2018)

La minería de datos tiene cierta similitud con *machine learning*, pero existen ciertas diferencias, *CleverData* (2014) menciona estas diferencias en un artículo:

“Desde nuestro punto de vista, la principal diferencia radica en el objetivo que tiene cada una de las disciplinas. Mientras que la minería de datos descubre patrones anteriormente desconocidos, el *machine learning* se usa para reproducir patrones conocidos y hacer predicciones basadas en los patrones.

En pocas palabras se podría decir que la minería de datos tiene una función exploratoria mientras que el *machine learning* se focaliza la predicción.

4.3.2.2 Cenfotec

Este centro de estudios superior se enfoca en la tecnología de la información y comunicación y cuenta con un posgrado y un certificado en Analítica de datos y *big data* (2018): “Este programa en Ciencia y analítica de datos masivos (*Data Analytics/Big data Certificate*) que ofrece la Universidad Cenfotec junto con XTOL”.

Una entrevista al señor Adrián Monge, integrante del grupo de *Machine Learning* Costa Rica en la red social *Facebook*, quien cursó el posgrado de analítica de datos en Cenfotec relató que el posgrado abarca varias ramas como *big data* (programación e infraestructura), *machine learning*, gestión de datos e incluso gestión de proyectos en *big data* y *analytics*.

“En sí la experiencia ha sido muy buena, pues los cursos son 100% virtuales, hay una sesión semanal de aproximadamente de 1 hora, pero las clases no son presenciales, sino que sirven para evacuar dudas y comentar sobre los proyectos o el ámbito en general, el aprendizaje depende totalmente de uno mismo y el empeño que uno le pueda poner determina su éxito. “

También comenta que ha sido de gran utilidad, pues es un área en la que quiere desempeñar como carrera profesional. Actualmente, trabaja en *business intelligence*, muy relacionado con el tema de *big data*, pero un volumen menor y con un enfoque más que todo a reportería.

El curso tuvo mucha relación con *machine learning* y se vieron temas tales como: Encontrar patrones o predecir basado en los datos y en el posgrado se le hizo mucho énfasis, sobre todo en el lenguaje R con algoritmos y tratamiento de datos en crudo.

Un comentario importante que hizo el señor Monge en cuanto a la oferta de *machine learning* y la necesidad de mejorar la educación, según la demanda fue el siguiente:

“Claro, evidentemente ocupamos renovar la oferta porque cuesta mucho encontrar especializaciones nuevas que en otras partes del mundo ya tienen un gran crecimiento, lamentablemente estamos muy atrás”.

Capítulo V

Conclusiones y Recomendaciones

5.1 Conclusiones

Esta investigación tuvo como objetivo elaborar un estudio técnico de la oferta de servicios tecnológicos digitales con *machine learning* por medio de las empresas proveedoras que permita el análisis de sus características principales en la Gran Área Metropolitana de Costa Rica durante el periodo 2017-2018.

Se concluye que el estado de desarrollo de la tecnología *machine learning* está en crecimiento y apoyado por algunas empresas por lo que estos primeros esfuerzos son importantes para tener una base sólida de conocimiento por experiencia e investigación.

Las características principales encontradas revelan que las empresas están trabajando de manera similar en nivel mundial y no se halló diferencia en el mercado costarricense.

Se puede concluir que el estado del *machine learning* se encuentra en una fase inicial y de crecimiento, sin llegar a su completo desarrollo, según las características encontradas en el nivel técnico, funcional y organizacional.

5.1.1 Definición de la oferta de servicios tecnológicos con *machine learning* e identificación de las empresas.

Respecto de la definición de la oferta de servicios con *machine learning*, es difícil sectorizar este grupo de empresas debido a la fácil adaptación de esta tecnología a cualquier área económica, aunque la mayoría son del sector tecnológico de desarrollo de software, pues existe un gran nicho de mercado.

Estas empresas están madurando y experimentando con el *machine learning*, por lo que se presentará mayor integración de este tipo servicios en la vida cotidiana del costarricense con mayor intensidad.

5.1.2 Características técnicas

Se concluye que el *framework* más utilizado por las empresas entrevistadas, es *TensorFlow* por su gran flexibilidad, código abierto y el gran apoyo que tiene de la comunidad. El tiempo de desarrollo de prototipos es importante a la hora de elegir un *framework*, *TensorFlow* tiene la ventaja de contar con una velocidad de producción y respaldo bien documentado.

Esta elección es compartida en el nivel mundial por lo que se está tomando un camino seguro para el desarrollo servicios o productos.

En cuanto a los métodos de *machine learning* se comprobó que sigue la misma línea que en el nivel mundial con los métodos de regresión y *deep learning*, aunque este último tiene un campo más grande y se puede considerar como un subcampo dentro del *machine learning* especializado en procesamiento de imágenes y lectura.

El *deep learning* es la siguiente ola tecnológica, pues es más potente debido al sistemas de capas ocultas que permiten una generación de resultados inmensa y variable, en otras palabras, una similitud a cómo funciona el cerebro humano.

Respecto de los estilos de *machine learning*, se concluye que el estilo más utilizado es el supervisado, pues permite un seguimiento y control de los resultados que el modelo arroje y así lograr el objetivo que la empresa estipule.

5.1.3 Características funcionales

Se concluye que la tecnología de *machine learning* se puede aplicar en cualquier tipo de área económica siempre y cuando se cuente con una fuente de datos almacenada, ya sea en la nube o servidores grandes de la cual se pueda obtener modelos que permitan un beneficio sustancial.

También el *machine learning* trabaja áreas como *marketing* digital, logística y finanzas específicamente, se espera que más áreas económicas implementen este tipo de tecnología debido a sus grandes ventajas.

Sobre los usos, se concluye que uno de los casos destacados son los *chatbots*, pues mejoran la interacción con cliente, reducen costos en cuanto a contratación de *call centers* y maximizan las ventas de productos. Es decir, que la revolución del *machine learning* vendrá a reducir los empleos automatizables, pero generando otras áreas de desarrollo económico, el *machine learning* es altamente adaptable a las áreas económicas.

5.1.4 Características organizativas

Se concluye, además, que un gran reto es la inversión de tiempo en la investigación y pruebas para lograr objetivos planteados inicialmente, como la eficiencia y optimización de procesos.

Un departamento a cargo de un proyecto puede tardar mucho tiempo para determinar cuál método, estilos y tipos se requieren para la generación de un modelo que cumpla con dichos objetivos. Esto sumado a las altas expectativas de los altos mandos que no tienen el conocimiento de que es realmente *machine learning* y los casos de éxito encienden ese interés por las ventajas que puede brindar esta tecnología.

También muchas veces los resultados de un proyecto con *machine learning* no son del todo el mejor u óptimo y esto lleva a la frustración de los altos mandos, pero sumamente necesario es el tiempo de maduración de los modelos y generación de los algoritmos más efectivos.

Por otra parte, hay altos mandos que no ven una ventaja en la incursión en *machine learning*. Muchas veces la raíz de la innovación surge del mismo departamento de tecnología que debe convencer a sus superiores por medio de prototipos que igualmente consumen tiempo y que cuesta mucho trabajo. Por lo que conseguir patrocinio para proyectos es sumamente desgastante.

La capacitación personal es sumamente importante para el desarrollo de proyectos con *machine learning* y el tiempo de capacitación en un espacio de pruebas para ir generando experiencia.

Se encontró que el *machine learning* se consume de manera local y también se exporta el conocimiento y servicios a otros países.

A la hora de implementar servicios o productos de *machine learning* surgen necesidades. El deseo de innovar impulsa a cambiar la forma de brindar servicios y esta ola tecnológica debe ser tomada con cierta capacidad.

La falta de personal capacitado en especialidades en *machine learning* es otra necesidad para las empresas y sumando a la zona de *confort* que existe en los ingenieros en *software* y las universidades en general por tener un plan estratégico para la apertura de nuevas especialidades.

La percepción a futuro de los entrevistados sobre *machine learning* es promisorio y se está aceptando en más empresas. La cantidad de despliegues de soluciones con inteligencia artificial aumentarán y los modelos de *machine learning* serán cada vez más sencillos y atinados, según los requerimientos.

El mercado y la comunidad de *machine learning* está en constante maduración, por lo que seguidamente provocará que surjan más emprendimientos, en consecuencia, desarrollo y actividades económicas enfocadas en *machine learning*.

Los equipos para desarrollo de *machine learning* por lo general se ha visto son en su totalidad de ingeniería en *software*, pero en algunos casos estos trabajan en conjunto con algún otro tipo de profesional, como ingenieros eléctricos, ingenieros en estadística y físicos, para producir un modelo que cumpla los requerimientos establecidos. Esto denota lo diversificado que es el campo del

machine learning y que es una situación importante, pues se puede implementar en muchas áreas.

Además, en cuanto a capacitación del recurso humano, son mínimas las instituciones que enseñan algún tipo de rama que esté relacionado con *machine learning*. En consecuencia, los profesionales deben capacitarse de manera personal por medio de cursos en línea que en algunos casos cuentan con un certificado de aprobación para respaldar el esfuerzo realizado y por medio de investigación científica en las universidades.

Por otra parte, la investigación interna de las empresas que se realizan en el tiempo y respaldo de investigaciones científicas mejoran el conocimiento de los departamentos.

5.1.5 Academia

Aunque no forma parte de los objetivos de esta investigación fue oportuno investigar cómo se están capacitando los profesionales y si hay un plan para que estas empresas investigadas cuenten con recurso humano apto para sus funciones.

Se puede concluir que, en las universidades estatales de Costa Rica, se detectó que muchas están enfocadas a la investigación y publicaciones científicas, además de formar alianzas con universidades en el extranjero, así como la colaboración en eventos que impulsen la enseñanza del *machine learning* e incentivar a estudiantes a vincularse con esta especialidad. Los planes de estudios

no están enfocados en algún tipo de énfasis en la ciencia de datos o similar a *machine learning*.

En el nivel de centros privados de enseñanza existen posgrados y cursos de actualización en el área de minería de datos y *big data*, que permiten que profesionales y estudiantes se capaciten para tener las habilidades suficientes para desempeñar un puesto que requiera estas capacidades de resolución de problemas y presentar propuestas con sistemas cognitivos que permitan un mejor desempeño de tareas.

5.2 Recomendaciones

5.2.1 Definición de la oferta de servicios tecnológicos con *machine learning* e identificación de las empresas.

Se recomienda una mejor sectorización de esta población por medio de la fundación de un ente respaldado ya sea por universidades del estado o del gobierno de Costa Rica que promuevan el interés en este tipo de tecnología, además de investigar si existen más empresas que trabajen con la tecnología *machine learning* en diferentes áreas económicas. Con esta base de información de empresas que deseen participar se puede lograr convenios, patrocinios y un canal de empleabilidad.

5.2.2 Características técnicas

Se recomienda crear un estándar de las características que engloba el *machine learning* que permita una mejor comprensión de los métodos, estilos y

tipos. También un análisis a fondo sobre cada una de las características técnicas que pueden ser objeto de investigación.

Se recomienda utilizar *machine learning* supervisado para una primera incursión, pues las ventajas principales de la técnica supervisada frente a otras técnicas por su fácil entendimiento, responde a preguntas concretas y cuenta con métodos para medir la calidad de los algoritmos antes de ponerlos en producción.

Se recomienda tener bases de datos optimizadas, filtradas y ordenadas, que permitan desarrollar los procesos de implementación de *machine learning* de manera eficiente. Previendo un gasto de tiempo en el ordenamiento y filtrado de los datos. Como recomendación general, es de importancia tener datos de calidad que una gran cantidad de datos almacenados.

5.2.3 Características funcionales

Se recomienda investigar e incursionar en servicios o productos con *machine learning* cuando exista una instancia donde se pueda automatizar algún proceso en diferente área económica, por medio de proyectos que ayuden por un lado al negocio a entender las posibilidades que se les abren y por otro que las áreas tecnológicas empiecen a valorar cómo integrarlo en sus sistemas.

Se recomienda tener planes de reubicación de personal a otras áreas no automatizables, pues los puestos manuales y cíclicos realizados por trabajadores son altamente automatizables y corren el riesgo de eliminarse, por lo que es un

deber social tener un control estratégico que este incluido en la implementación de *machine learning*.

Esto depende de los intereses que cada empresa; para algunas no será rentable mantener trabajadores, en cambio resultará factible la implementación *machine learning* por la reducción de costos o mejoras en los servicios o productos.

Se recomienda a los ministerios, instituciones autónomas de Costa Rica valorar la incursión en el mejoramiento de procesos por medio de *machine learning* para el bien de la población.

5.2.4 Características organizativas

En cuanto a las empresas que deseen incursionar en utilizar servicios o productos con *machine learning* se recomienda tener las condiciones apropiadas como patrocinio, espacios dedicados a *machine learning* y oportunidad de contratar profesionales que tengan conocimientos o especialización en *machine learning*. Se recomienda realizar una investigación para conocer cuántos de estos proyectos logran tener éxito.

Además, que las empresas mantengan el acercamiento con las universidades para tener en cuenta los estudiantes que deseen hacer práctica profesional o trabajos de investigación. Los centros de investigación de *machine learning* en las universidades pueden ser la forma de conocer los tipos de publicaciones científicas y las especialidades que manejan cada profesional.

5.2.5 Academia

Las universidades estatales y fuentes de enseñanza tienen gran responsabilidad sobre el desarrollo tecnológico del país, por lo que se hace un llamado a mejorar los planes de estudio y crear nuevas carreras especializadas en la ciencia de datos y *machine learning*.

Esta ola tecnológica viene muy rápido y es sumamente necesario estar preparados para que no se tome por sorpresa. La pérdida de empleos a causa de esta revolución es de mucho cuidado. Hay que salir de la zona de *comfort* y promover nichos para la innovación.

Los planes de estudio tienen que revalorarse para crear ya sea más cursos dentro los planes generales de las carreras de ingeniería y tecnología, así como interdisciplinaridad de proyectos de investigación que fortalecen los lazos entre las escuelas de cada universidad.

Promover la investigación y generación de espacios para que los estudiantes puedan ir adquiriendo conocimiento en tecnologías del futuro.

Las universidades que han tomado la decisión innovar en el campo de *machine learning* y otras subramas son el Tecnológico de Costa Rica y la Universidad de Costa Rica con sus centros de investigación, incentivan al estudiante a capacitarse en estas tecnologías y dar a conocer sus trabajos de investigación en el nivel internacional.

Es oportuno mencionar que el Tecnológico de Costa Rica está valorando a mediano plazo crear algún tipo de especialidad en *machine learning*.

Es necesario mencionar que una fuerte base en las matemáticas mejora la comprensión en cuanto a la generación de modelos de *machine learning*, por lo que se hace la recomendación de mejorar el sistema de enseñanza primaria y secundaria de Costa Rica para que este problema no se magnifique cuando los estudiantes lleguen a las universidades. Es decir, es una cadena de principio a fin y debemos estar preparados como sociedad.

Las oportunidades están afuera, la documentación sobre *machine learning* crece día a día y las soluciones son cada vez más eficientes. Pero el paso debe darse y no quedarse rezagado, es algo que debe estudiarse a fondo.

Por ahora existe una comunidad en la red social Facebook llamada *Machine learning Costa Rica* que busca apoyar el conocimiento en esta rama tecnológica, compartir noticias, puestos de trabajo relacionados y reuniones para compartir experiencias. Sería recomendable que las universidades del Estado formaran un convenio o alianza que permita impulsar la enseñanza del *machine learning*.

Los profesionales tienen una gran oportunidad de diversificar en esta rama y que mucha de la información está en cursos en línea y documentos científicos, que pueden ser de gran ayuda a la hora de formar las bases.

Se recomienda hacer un estudio de nivel latinoamericano acerca del desarrollo de *machine learning*, esto en conjunto con los gobiernos de la región.

En cuanto a los entes gubernamentales de Costa Rica, se recomienda generar planes estratégicos de evolución tecnológica para crear opciones para que

el estudiante de universidades estatales puedan retribuir de manera de proyecto o práctica profesional con la necesidad de mejorar los sistemas tecnológicos estatales.

Capítulo VI

Guía introductoria al *machine learning* en empresas

Esta guía es una base o punto de referencia que pueden utilizar las empresas para tomar en consideración a la hora de incursionar en el campo del *machine learning*.

- Fase exploratoria.
- Fase diseño, experimentación e implementación.

6.1 Fase exploratoria

En esta fase es necesario investigar toda información relevante sobre *machine learning*, como análisis de mercado que permita conocer lo realizado por otras empresas, y así visualizar la posibilidad de la implementación en *machine learning* sin realizar una inversión considerable.

6.1.1 Determinación de los objetivos de negocio.

Es recomendable investigar cuáles son los objetivos que busca la empresa con un servicio o producto, por lo que es importante tener claro cuáles son las áreas para optimizar. Cabe destacar que es mejor enfocar el esfuerzo en una solución pequeña que consuma recursos en forma limitada, pero que permita la medición de desempeño sin tener que incurrir en grandes pérdidas o propuestas sin sentido.

Se debe tomar en cuenta los siguientes puntos:

- Generar un mapa de procesos internos de la empresa.
- Crear un mapa de los servicios ofertados y los procesos de inicio a fin.
- Determinar cuáles son los factores para optimizar en los procesos internos o servicios por ofertar.
- Seleccionar un objetivo por desarrollar como plan piloto.

6.1.2 Análisis de la competencia, una buena estrategia.

Como primer paso es necesario investigar qué hacen otras empresas para determinar si existe una solución similar, y de esta manera no invertir esfuerzos para generar una solución u optimización ya existente. Además, esto ayuda en la generación de propuestas diferenciadas y aumenta la competitividad.

Se pueden tomar los siguientes puntos para recolección de la información:

- Buscar información en fuentes de datos públicas.
- Seleccionar empresas del sector económico o que cuenten con servicios con *machine learning*.
- Analizar las soluciones que brindan la competencia.
- Determinación del factor de diferenciación.
- Ventajas de los servicios.
- Investigar tendencias de *machine learning* en otros nichos de mercado.

6.1.3 Recomendaciones sobre infraestructura para el desarrollo de *machine learning*.

Hay métodos de *machine learning* que demandan alto poder computacional. Por ejemplo, métodos como el *deep learning* demandan memoria gráfica, por lo que es necesario invertir en unidades de procesamiento gráfico (GPU) de alta capacidad.

Cabe mencionar que con un computador de mediana capacidad² con solo el uso de la unidad central de procesamiento (CPU) se puede iniciar operaciones o procesos de *machine learning*. Por lo que no es necesario una gran inversión para investigar el potencial del *machine learning*.

La infraestructura basada en la nube brinda cualidades convenientes como un alto poder computacional como CPU, GPU y almacenamiento elástico, que permite incluir más información en las bases de datos, sin necesidad de hacer cambios de hardware, pues las grandes granjas de servidores mantenidas por grandes empresas como Google, Amazon o Microsoft generan un gran respaldo, además de que se evita el mantenimiento de equipo tecnológico.

² Computador de mediana capacidad: Un computador con un procesador de mediana categoría, una cantidad mínima de 4GB de memoria RAM y una tarjeta de video de bajo perfil.

Los costos de utilizar estos servicios se basan en el tiempo de uso, así que cuántos más recursos computacionales y horas utilizadas mayor será el costo.

Para conocer el costo de la renta de estos servicios, tanto como Google, Amazon y Microsoft cuentan con calculadoras especializadas en dar un aproximado de los costos por incurrir.

Es necesario considerar la seguridad en caso de contar con información sensible, esta información estará alojada en servidores administrador por terceros por lo que sí es prioridad la seguridad se debe considerar un cuarto de servidores en las instalaciones de cada empresa.

6.1.4 Recomendaciones sobre recurso humano por contratar.

Es importante la contratación de especialistas en *machine learning* para un desarrollo sólido de las implementaciones. También se puede contar con Ingenieros en Software o Ciencias de la Computación con alguna experiencia en analítica, minería de datos o *machine learning*.

Es importante también buscar un canal de comunicación entre las empresas que deseen implementar *machine learning* con las universidades estatales para hacer alianzas que permitan direccionar los estudiantes que necesiten realizar proyectos o tesis de graduación en empresas. Esto permite que las empresas logren generar conocimiento y los estudiantes experiencia de *machine learning*.

6.2 Fase de diseño, experimentación e implementación.

6.2.1 Etapas base para desarrollar *machine learning*.

En esta sección se plantean las etapas para realizar el diseño o la estrategia para abordar el desarrollo de *machine learning*.

Las etapas son las siguientes:

- Clasificación de los problemas
- Identificar la información necesaria
- Preparación de la información
- Determinar las técnicas de *machine learning*.
- Ejecutar y validar
- Desplegar

6.2.1.1 Clasificación de los problemas.

Es necesario determinar cuáles son los problemas para resolver, como recomendación se pueden dividir en problemas exploratorios o predictivos. Cuando se necesite determinar patrones en la información o grupos, es cuando se debe utilizar modelos exploratorios. Si se necesita determinar cuándo será la siguiente ocurrencia de un proceso se deben utilizar procesos predictivos.

6.2.1.2 Identificar la información necesaria.

La información almacenada se puede adquirir por medio de sistemas computacionales como los ERP (sistemas de planificación de recursos empresariales), dispositivos IoT (Internet de las Cosa), *chats*, redes sociales, correos electrónicos y bases de datos. Esta información puede poseer cualidades estructuradas o no estructurada.

6.2.1.3 Preparación de la información.

Respecto de la preparación de la información, se recomienda técnicas como transformación, normalización y depuración para la selección de grupos para la aplicación de los modelos de *machine learning*.

6.2.1.4 Determinar las técnicas de *machine learning*.

Determinar cuáles son las técnicas de *machine learning* óptimas para resolver depende del tipo de problema o tipo de enfoque. La combinación de las técnicas puede variar, incluso explorar nuevos métodos que permitan resolver el problema planteado.

Marketing Digital				
Framework	Métodos	Estilos	Tipo de enfoque	Área económica
<i>Tensor Flow, Amazon ML, Apache spark</i>	Redes neuronales, <i>deep learning</i> , convoluciones, <i>long short term</i>	Supervisado	Clasificación de entidades y lenguaje natural	Mercadeo

Tabla 22 Respuesta de Neustar con impacto en el área económica de mercadeo o marketing digital.

Fuente: *Elaboración propia.*

A modo de ejemplo, según la información analizada en esta investigación se encontró que para un problema como la “clasificación de entidades y lenguaje natural” se puede resolver con la combinación de *TensorFlow, Amazon ML, Apache Spark* como *frameworks*, con métodos de redes neuronales, *deep learning*, convoluciones y *long short term*.

El estilo supervisado permite un control sobre las entradas y salidas del modelo, además de que el estilo busca la exploración de patrones o grupos que permitan determinar cierto problema.

Se puede plantear, además, diferentes combinaciones de técnicas para validar el desempeño, pues todo depende del objetivo por resolver.

6.2.1.5 Ejecutar y validar

Después de determinar las técnicas por utilizar es necesario validar su rendimiento. Para ello se debe plantear una lista de pruebas con diferentes técnicas que contemplen los siguientes puntos como guía:

- Estilo de *machine learning*.
- Métodos utilizados.
- Enfoque.
- Selección de datos de entrada y salida.
- Duración de la prueba.
- Rendimiento de hardware.
- Analizar los resultados.
- Cumplimiento del objetivo.

Es recomendable crear un espacio de pruebas con una copia como muestra de la información seleccionada para evitar pérdida de los datos.

6.2.1.6 Desplegar

El último paso es desplegar el servicio o producto que genere un valor de negocio. Este valor de negocio es para la toma de decisiones, mejoramiento de sistemas informáticos o almacenamiento de los resultados de la experimentación para futuros análisis.

También en esta fase se puede suplir el funcionamiento de sistemas existentes, procesos manuales o aplicaciones, es decir, eliminar procesos que son menos rentables. Por lo que se recomienda establecer un plan estratégico que permita decidir cuándo y cómo implementar esta transformación del modelo de negocio.

Referencias Bibliográficas

- (PEN), P. E. (2014). *Estado de la Ciencia, la Tecnología y la Innovación*. San José: Programa Estado (PEN).
- Alfani, F. B. (24 de 11 del 2017). *M.Sc. Denise Dajles Kellerman buscará impulsar la investigación en el PRIS-Lab UCR*. Obtenido de <https://www.ucr.ac.cr/noticias/2017/11/24/m-sc-denise-dajles-kellerman-buscara-impulsar-la-investigacion-en-el-pris-lab-ucr.html>
- AltexSoft. (19 de julio de 2018). *Comparación del aprendizaje automático como un servicio: Amazon, Microsoft Azure, Google Cloud AI, IBM Watson*. Obtenido de Altexsoft: <https://www.altexsoft.com/blog/datascience/comparing-machine-learning-as-a-service-amazon-microsoft-azure-google-cloud-ai-ibm-watson/>
- Amster, A. (10 de 03 de 2016). *qubole.com*. Obtenido de Los Mejores Casos De Uso De Apache Spark: <https://www.qubole.com/blog/apache-spark-use-cases/>
- Andrade Espinoza, S. (2006). *Diccionario de economía*. Lima: Editorial Andrade. Recuperado el 14 de 06 de 2017
- Arias, Fidas G. (2012). *El proyecto de investigación*. Caracas, República Bolivariana de Venezuela: Editorial Episteme. Recuperado el 06 de agosto de 2017

- Asiri, S. (11 de junio de 2018). *Machine Learning Classifiers*. Obtenido de towardsdatascience: <https://towardsdatascience.com/machine-learning-classifiers-a5cc4e1b0623>
- Atilim Gunes Baydin, B. A. (2018). Automatic Differentiation. *Journal of Machine Learning Research*, 1-43.
- Azure, M. (15 de 01 de 2019). *Microsoft Azure*. Obtenido de Azure Machine Learning Services: <https://azure.microsoft.com/en-us/services/machine-learning-service/>
- Barney, J. B. (2007). *Resource-based theory: Creating and sustaining competitive advantage*. Oxford: Oxford University Press.
- Bishop, C. (2006). *Pattern Recognition and Machine Learning*. Berlin: Springer-Verlag GmbH; Edición: 1st ed. 2006. Corr. 2nd printing 2011 .
- Blaisdell, R. (27 de junio de 2017). *rickscLOUD.com*. Obtenido de Meeting the most popular deep learning framework – Tensorflow: <https://rickscLOUD.com/meeting-the-most-popular-deep-learning-framework-tensorflow/>
- Brennan, B. (12 de 11 de 2017). *Medium.com*. Obtenido de Libraries vs. Frameworks: <https://medium.com/datafire-io/libraries-vs-frameworks-626cdde799a7>

- brilliant.org. (12 de diciembre de 2018). *Bayes' Theorem and Conditional Probability*. Obtenido de brilliant.org: <https://brilliant.org/wiki/bayes-theorem/>
- Brownlee, J. (25 de noviembre de 2013). *Un recorrido por los algoritmos de aprendizaje automático*. Obtenido de machinelearningmastery.com: <https://machinelearningmastery.com/a-tour-of-machine-learning-algorithms/>
- Brownlee, J. (11 de diciembre de 2017). *Diferencia entre clasificación y regresión en el aprendizaje automático*. Obtenido de machinelearningmastery.com: <https://machinelearningmastery.com/classification-versus-regression-in-machine-learning/>
- Business, F. R. (15 de 01 de 2019). *FinancesOnline Reviews for Business*. Obtenido de Revisión de Azure Machine Learning Studio: <https://reviews.financesonline.com/p/azure-machine-learning-studio/>
- Califano, J. (14 de mayo de 2018). *¿Por qué es Tensorflow tan popular?* Obtenido de Quora: <https://www.quora.com/Why-is-Tensorflow-so-popular>
- Chacón, K. (15 de diciembre de 2016). *Crédito se alía con la inteligencia artificial*. Recuperado el 26 de Octubre de 2017, de La Nación: <https://www.nacion.com/economia/banca/credito-se-alia-con-la-inteligencia-artificial/ZFCP6ZBAMNAWPDFTXJ3TOG2I4/story/>
- Chacón, K. (21 de 02 de 2018). *Nuevo chatbot puede analizar su información para gestionarle créditos*. Obtenido de El Financiero:

<https://www.elfinancierocr.com/tecnologia/nuevo-ichatboti-puede-analizar-su-informacion/RV7M7N6VOJA73I42HGYYW27GLU/story/>

CHRISTOPHER LOVELOCK, J. W. (2009). *MARKETING DE SERVICIOS personal, tecnología y estrategia*. México: Pearson Educación.

Cloud, G. (31 de 01 de 2019). *Google Cloud*. Obtenido de Cloud AutoML:

<https://cloud.google.com/automl/>

Cloud, G. (30 de 01 de 2019). *Google Cloud*. Obtenido de AutoML Natural

Language Beginner's guide: <https://cloud.google.com/natural-language/automl/docs/beginners-guide>

Cobo, H. (27 de abril de 2017). *Machine Learning, Una expresión de la inteligencia artificial*. Recuperado el 12 de octubre de 2017, de Mundo Ejecutivo México:

<http://mundoejecutivo.com.mx/management/2017/04/24/machine-learning-expresi-n-inteligencia-artificial>

Computing, R. C. (03 de 01 de 2018). *Experto En Aprendizaje Automático Será El*

Perfil Más Demandado En España En El Futuro. Obtenido de Sitio Web

Revista Cloud Computing:

<https://www.revistacloudcomputing.com/2018/01/experto-en-aprendizaje-automatico-sera-el-perfil-mas-demandado-en-espana-en-el-futuro/>

Concha, D. J. (01 de 09 de 2006). *EXPORTACIONES DE SERVICIOS*. Obtenido de Universidad Icesi:

<http://www.icesi.edu.co/blogs/icecomex/2006/09/01/exportaciones-de-servicios/>

Cordero, C. (2016 de 08 de 24). *Costa Rica explora atraer inversiones globales en centros de datos*. Recuperado el 2018 de 03 de 14, de La Nación Costa Rica: <https://www.nacion.com/economia/negocios/costa-rica-explora-atraer-inversiones-globales-en-centros-de-datos/DZQ3NUXHVNFAFKSF3IO2A4DHE/story/>

Datos, P. I. (21 de 08 de 2018). *¿Quiénes somos?* Obtenido de Promidat: https://promidat.com/%C2%BFqui%C3%A9nes_somos

Debitoor. (16 de 10 de 2018). *¿Qué es un proveedor?* Obtenido de debitoor.es: <https://debitoor.es/glosario/definicion-proveedor>

Dorantes, R. (09 de 11 de 2017). *www.entrepreneur.com*. Obtenido de Qué es una startup: <https://www.entrepreneur.com/article/304376>

Drouin, A. (24 de junio de 2017). *Microbiome Summer School 2017, Introduction to Machine Learning*. Obtenido de aldro61.github.io: <https://aldro61.github.io/microbiome-summer-school-2017/sections/basics/>

E. Knuth, D. (1981). *Algorithms in Modern Mathematics and Computer Science*. Stanford, CA, USA : Springer.

Económicas, A. (2011). *Actividades Económicas*. Obtenido de Sector terciario: <http://actividadeseconomicasgeografia.blogspot.com/p/sector-terciario.html>

Economist, T. (20 de Agosto de 2016). *Machine learning, Of prediction and policy*.

Recuperado el 06 de noviembre de 2017, de The Economist:

<https://www.economist.com/finance-and-economics/2016/08/20/of-prediction-and-policy>

eleconomista.es. (13 de 04 de 2016). *eleconomista.es*. Obtenido de ¿Qué son exactamente los chatbots y para qué sirven?:

<http://www.eleconomista.es/tecnologia/noticias/7488529/04/16/Que-son-exactamente-los-chatbots-y-para-que-sirven.html>

EIEconomista.es. (06 de abril de 2017). *EI DEEP LEARNING será la próxima revolución tecnológica*. Recuperado el 10 de octubre de 2017, de

Eleconomista.es:

<http://www.eleconomista.es/campus/noticias/8278033/04/17/EI-DEEP-LEARNING-sera-la-proxima-revolucion-tecnologica.html>

elétrica, P. d. (21 de 08 de 2018). *Investigación Laboratorios*. Obtenido de Universidad de Costa Rica:

<http://posgrado.eie.ucr.ac.cr/index.php/investigacion/>

elétrica, P. d. (21 de 08 de 2018). *Universidad de Costa Rica*. Obtenido de Escuela de Ingeniería Eléctrica:

<http://posgrado.eie.ucr.ac.cr/index.php/admision/>

EmprendePyme.net. (21 de 10 de 2010). *Emprende Pyme*. Obtenido de ¿Qué es la capacitación?: <https://www.emprendepyme.net/que-es-la-capacitacion.html>

España, N. (16 de Agosto de 2016). *Las bases de la Inteligencia Artificial*. Obtenido de NetworkWorld España From IDG: <http://www.networkworld.es/mundo-profesional/las-bases-de-la-inteligencia-artificial>

Española, R. A. (12 de diciembre de 2018). *Real Academia Española* . Obtenido de rae.es: <http://dle.rae.es/srv/search?m=30&w=implementar>

Exastax. (03 de 02 de 2017). *Exastax.com*. Obtenido de TOP FIVE USE CASES OF TENSORFLOW: <https://www.bigdataguys.com/top-five-use-cases-of-tensorflow-deep-learning/>

Faggella, D. (29 de octubre de 2018). *What is Machine Learning?* Obtenido de techemergence.com: <https://www.techemergence.com/what-is-machine-learning/>

Fernandes, A. (01 de 08 de 2015). *Definición de NLP, NLU, NLG y cómo funcionan los Chatbots*. Obtenido de planetachatbot.com: <https://planetachatbot.com/definicion-nlp-nlu-nlg-y-c0mo-funcionan-los-chatbots-fcf4b22e065a>

FERNÁNDEZ, D. (13 de Mayo de 2017). *La inteligencia artificial obliga a redefinir la economía*. Obtenido de Sitio Web El País:

https://elpais.com/economia/2017/05/12/actualidad/1494601971_737485.html

Fernandez, S. (07 de 10 de 2016). Machine Learning y conocimientos principales. *Machine Learning*. Madrid, Madrid, España.

Financiero, E. (22 de 01 de 2017). *Nuevas tecnologías impulsarán empleo de informáticas ticas en 2017*. (C. C. Pérez, Editor) Obtenido de El Financiero: <https://www.elfinancierocr.com/tecnologia/nuevas-tecnologias-impulsaran-empleo-de-informaticas-ticas-en-2017/JHOMIA2EYZBYTIJPECLQJCL7I4/story/>

Flach, P. (2012). *Machine Learning: The Art and Science of Algorithms that Make Sense of Data*. Cambridge: Cambridge University Press; 1 edition.

Foundation, A. S. (27 de 02 de 2014). *The Apache Software Foundation Blog*. Obtenido de The Apache Software Foundation Announces Apache™ Spark™ as a Top-Level Project: https://blogs.apache.org/foundation/entry/the_apache_software_foundation_announces50

G, Y. (11 de 06 del 2018). *Towards Data Science*. Obtenido de [towardsdatascience.com: https://towardsdatascience.com/getting-started-with-automl-vision-alpha-ba769121235c](https://towardsdatascience.com/getting-started-with-automl-vision-alpha-ba769121235c)

George F. Luger, W. A. (1993). *Artificial intelligence: structures and strategies for complex problem solving*. Michigan: Benjamin/Cummings Pub. Co.

Golla, Y. (12 de septiembre del 2018). *What is machine learning algorithm?*

Obtenido de [www.quora.com: https://www.quora.com/What-is-machine-learning-algorithm](https://www.quora.com/What-is-machine-learning-algorithm)

GONZÁLEZ, A. (01 de 07 del 2014). *¿Qué es Machine Learning?* Recuperado el

02 de 11 del 2016, de CleverData: <https://cleverdata.io/que-es-machine-learning-big-data/>

González, A. (30 de julio del 2014). *Conceptos básicos de Machine Learning.*

Obtenido de CleverData: <https://cleverdata.io/conceptos-basicos-machine-learning/>

González, A. (07 de 11 del 2016). *5 consejos para empezar con Machine Learning*

en la empresa. Obtenido de Sitio Web Clever Data: Uno de los principales obstáculos que estamos viendo en la implementación de Machine Learning en las empresas es, precisamente, que no tienen claro cómo empezar a usarlo y muchas veces tampoco entienden las ventajas que supone.

Gonzalez, J. L. (08 de febrero del 2018). *Tipos de aprendizaje automático.*

Obtenido de Medium.com: <https://medium.com/soldai/tipos-de-aprendizaje-autom%C3%A1tico-6413e3c615e2>

Gonzalez, V. (8 de noviembre del 2016). *Una breve historia del Machine Learning.*

Recuperado el 15 de 08 del 2017, de SynergicPartners:

<http://www.synergicpartners.com/una-breve-historia-del-machine-learning/>

Google. (08 de 11 del 2018). *CLOUD MACHINE LEARNING ENGINE*. Obtenido de Google: <https://cloud.google.com/ml-engine/>

Google. (20 de 01 del 2019). *Google Cloud*. Obtenido de Cloud AI Products: <https://cloud.google.com/products/ai/>

Google. (s.f.). *Aprendizaje automático y coincidencia de patrones*. Obtenido de Google Cloud: <https://cloud.google.com/prediction/>

Graupe, D. (2016). *Deep Learning Neural Networks*. Chicago: World Scientific Publishing Co. Pte. Ltd.

Hastie, T. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Second Edition*. Berlin: Springer-Verlag GmbH; Edición: 2nd ed. 2009, Corr. 9th printing 2017.

Hipertextual.com. (29 de 06 del 2016). *Las ventajas de utilizar inteligencia artificial en el sector financiero*. Recuperado el 28 de 05 de 2017, de HiperTextual.com: <https://hipertextual.com/presentado-por/caixabank/ibm-watson-supercomputer>

IBM. (27 de junio del 2018). *Machine learning & AI in Watson Studio*. Obtenido de IBM Watson Machine Learning: <https://www.ibm.com/cloud/garage/dte/producttour/product-tour-ibm-watson-machine-learning-watson-studio>

- IBM. (30 de 01 del 2019). *IBM Watson* . Obtenido de IBM Watson Machine Learning: <https://www.ibm.com/cloud/machine-learning>
- Institute, M. G. (2016). *The Age of Analytics competing in a data-driven world*. Global: McKinsey & Company.
- IT, G. (05 de 01 del 2019). *Los desafíos para la implementación de machine learning en empresas*. Obtenido de GMP IT Consulting: <https://gmpitconsulting.com.ar/los-desafios-para-la-implementacion-de-machine-learning-en-empresas/>
- IT, R. (28 de 08 del 2018). *¿Cómo capacitarse en Machine Learning?* Obtenido de Sitio Web Reclu IT: <https://www.recluit.com/como-capacitarse-en-machine-learning/>
- Janssen, C. (02 de 10 del 2018). *Software Framework*. Obtenido de <http://www.techopedia.com>: <http://www.techopedia.com/definition/14384/software-framework>
- Jorge, F. D. (2011). *Mercadotecnia* . España: McGraw-Hill Interamericana de España S.L.; Edición: 4.
- Kevin Klues, S. P. (29 de 10 del 2017). *Mesosphere*. Obtenido de TensorFlow distribuido en cualquier infraestructura con Mesosphere DC / OS: <https://mesosphere.com/blog/tensorflow-gpu-support-deep-learning/>

Knight, W. (10 de abril del 2017). *Reinforcement Learning, By experimenting, computers are figuring out how to do things that no programmer could teach them*. Obtenido de MIT Technology Review.

Kotler Philip, A. G. (2012). *Marketing, Decimocuarta edición*. México, 2012: PEARSON EDUCACIÓN, S.A.

L., M. V. (05 de 02 del 2016). *La Nación Costa Rica*. Obtenido de Microsoft impulsará compañía de Costa Rica simMachines:
<https://www.nacion.com/tecnologia/informatica/microsoft-impulsara-compania-de-costa-rica-simmachines/UQHI3OCPJJEU3GN3FLQCA7J6AU/story/>

Li, M. (12 de 10 del 2017). *Clasificación de bibliotecas de Deep Learning populares para la ciencia de datos*. Obtenido de The Data Incubator:
<https://blog.thedataincubator.com/2017/10/ranking-popular-deep-learning-libraries-for-data-science/>

Lichtenstein, S. &. (2006). *The Construction of Preference*. New York: Cambridge University Press. Recuperado el 08 de 06 del 2017

Lorica, B. (13 de 10 de 2015). *oreilly.com*. Obtenido de Six reasons why I recommend scikit-learn: <https://www.oreilly.com/ideas/six-reasons-why-i-recommend-scikit-learn>

MacQueen, J. (1967). *Some methods for classification and analysis of multivariate observations*. Los Angeles: University of California.

- Madrigal, A. (12 de diciembre del 2014). *La Inteligencia artificial y su grado de avance en Costa Rica*. Obtenido de Costa Rica 2050:
<http://costarica2050.cr/la-inteligencia-artificial-y-su-grado-de-avance-en-costa-rica.html>
- Management, E. E. (07 de 03 del 2017). *Capacitación y desarrollo personal en una empresa*. Obtenido de Sitio Web Escuela Europea de Management:
<http://www.escuelamanagement.eu/desarrollo-personal/capacitacion-desarrollo-personal-una-empresa>
- Marsland, S. (2011). *Machine Learning: An Algorithmic Perspective*. Boca Raton London New York: Chapman and Hall/CRC; Edición: 1. Recuperado el 11 de 06 del 2017
- Martín, E. (28 de Noviembre de 2017). *Por qué 'machine learning' sera la tecnología más importante en 2018*. Obtenido de Periódico El País:
https://elpais.com/tecnologia/2017/11/28/actualidad/1511866764_933798.html
- MathWorks. (20 de junio del 2018). *¿Qué es el aprendizaje automático?* Obtenido de <https://www.mathworks.com>:
<https://www.mathworks.com/discovery/machine-learning.html>
- MathWorks. (10 de 09 del 2018). *Deep Learning, Tres cosas que es necesario saber*. Obtenido de <https://la.mathworks.com/discovery/deep->

learning.html#howitworks: <https://la.mathworks.com/discovery/deep-learning.html#howitworks>

McKinsey&Company. (2016). *The Age of analytics: Competing in a Data-Driven World*. North America: McKinsey Global Institute.

Mejía, F. C. (2006). *Diccionario Económico Financiero*. Medellín: Universidad de Medellín.

Mesa, A. R. (02 de julio del 2018). *Qué es Apache Spark*. Obtenido de openwebinars.net: <https://openwebinars.net/blog/que-es-apache-spark/>

Microsoft. (23 de 09 de 2018). *What is Azure Machine Learning service (preview)?* Obtenido de Microsoft: <https://docs.microsoft.com/en-us/azure/machine-learning/service/overview-what-is-azure-ml>

Microsoft. (27 de 03 del 2018). *What is Azure Machine Learning Studio?* Obtenido de Microsoft.com: <https://docs.microsoft.com/en-us/azure/machine-learning/studio/what-is-ml-studio>

Mitchell, T. M. (1997). *Machine learning (Mcgraw-Hill International Edit)*. New York: McGraw-Hill Education. Recuperado el 12 de 06 del 2017

Mizrahi, G. (11 de junio del 2018). *¿Los chatbots de FB Messenger son un buen canal para la empresa o NO?— ¡Esa es la pregunta!* Obtenido de planetachatbot.com: <https://planetachatbot.com/los-chatbots-de-fb->

messenger-son-un-buen-canal-para-la-empresa-o-no-esa-es-la-pregunta-4b6f49002b4e

MOOC. (03 de 02 del 2019). *¿Qué es un mooc?* Obtenido de Sitio Web MOOC.es:
<http://mooc.es/que-es-un-mooc/>

Nacional, U. (21 de 08 del 2018). *Plan de Estudios Ingeniería en Sistemas con Grado Bachillerato y Salida Lateral de Diplomado*. Obtenido de Escuela de Informática: <http://www.escinf.una.ac.cr/index.php/oferta-academica/bachillerato/plan-de-estudios-ing-sistemas>

Noguez, O. (23 de 04 del 2017). *Merca 2.0*. Obtenido de INTELIGENCIA ARTIFICIAL EN AMÉRICA LATINA, QUÉ NÚMEROS HAY HASTA AHORA: <https://www.merca20.com/inteligencia-artificial-en-america-latina-que-numeros-hay-hasta-ahora/>

Nolla, M. (31 de agosto del 2016). *Machine Learning, la era de la empresa inteligente*. Recuperado el 20 de noviembre del 2017, de TicBeat: <http://www.ticbeat.com/educacion/machine-learning-la-era-de-la-empresa-inteligente/>

Norvig, S. J. (2010). *Artificial Intelligence A Modern Approach*. Upper Saddle River, New Jersey 07458: Pearson Education, Inc.

OperSource. (08 de 11 del 2018). *What is open source?* Obtenido de Opensource.com: <https://opensource.com/resources/what-open-source>

Oscar, L. R., Acosta, J. C., Mata, L. E., Bachmann, N. G., & Vallejos. (01 de 06 de 2012). *Aprendizaje combinado, aprendizaje electrónico centrado en el alumno y nuevas tecnologías*. Recuperado el 26 de 06 del 2017, de SEDICI Argentina: <http://sedici.unlp.edu.ar/handle/10915/19306>

P., J. D. (1 de junio del 2012). *¿Qué son las actividades económicas?* Obtenido de www.actividadeseconomicas.org:
<https://www.actividadeseconomicas.org/2012/05/que-son-las-actividades-economicas.html>

Pacheco, S. S. (16 de 08 del 2018). *¿Cómo ha evolucionando la inteligencia artificial en Costa Rica?* Obtenido de Sitio Web Revista IT Now:
<https://revistaitnow.com/inteligencia-artificial-futuro-hoy/>

País, P. E. (s.f.). Obtenido de https://elpais.com/tecnologia/2017/11/28/actualidad/1511866764_933798.html.

Patrick Hall, J. D. (2014). *An Overview of Machine Learning with SAS® Enterprise Miner™, Paper SAS313-2014*. Cary, NC 27513: SAS Institute Inc.

Pérez, C. C. (14 de 04 del 2017). *El Financiero Costa Rica*. Obtenido de Firmas locales brindan servicios de inteligencia artificial:
<https://www.elfinancierocr.com/tecnologia/firmas-locales-brindan-servicios-de-inteligencia-artificial/ROZBI3SP3ZF6HLYTLFXTGWH4ZU/story/>

Pérez, C. C. (22 de 01 del 2017). *Nuevas tecnologías impulsarán empleo de*

informáticas ticas en 2017. Obtenido de Periódico El Financiero:

<https://www.elfinancierocr.com/tecnologia/nuevas-tecnologias-impulsaran-empleo-de-informaticas-ticas-en-2017/JHOMIA2EYZBYTIJPECLQJCL7I4/story/>

Piatetsky, G. (17 de diciembre de 2017). *KDNuggets*. Obtenido de Top Data

Science and Machine Learning Methods Used in 2017:

<https://www.kdnuggets.com/2017/12/top-data-science-machine-learning-methods.html>

Portolés, E. (2012). *Costa Rica: Estructura Económica*. Panamá: Instituto Español

de comercio exterior.

Recognition, h. B. (21 de 08 del 2018). *What is computer vision?* Obtenido de

BMVA: <http://www.bmva.org/visionoverview>

Reno, T. (19 de diciembre del 2017). *¿Qué significa "biblioteca" en el caso de los*

lenguajes de programación? Obtenido de SitioWeb Quora:

<https://www.quora.com/What-does-library-mean-in-the-case-of-programming-languages>

Rica, M. L. (29 de 10 del 2018). *Machine Learning Costa Rica, Facebook*.

Obtenido de Facebook:

<https://www.facebook.com/search/top/?q=machine%20learning%20costa%20Orica%20-%20mlcr>

- Rica, T. d. (21 de 08 del 2018). *El plan de Bachillerato en ingeniería en computación* . Obtenido de Sede Cartago: <https://www.tec.ac.cr/planes-estudio/bachillerato-ingenieria-computacion>
- Rica, U. C. (11 de julio del 2018). *XTOL: DATA ANALYTICS - BIG DATA (Certificado)*. Obtenido de <https://www.ucenfotec.ac.cr/data-analytics-big-data>
- Rica, U. d. (21 de 08 del 2018). *Carreras por facultades, Universidad de Costa Rica*. Obtenido de <https://www.ucr.ac.cr/estudiantes/carreras/>
- Rica, U. d. (21 de 08 del 2018). *Escuela de Ingeniería de Ciencias de la Computación e Informática*. Obtenido de Plan de Bachillerato y Licenciatura: <https://www.ucr.ac.cr/estudiantes/carreras/#ingenierias>
- Rica, U. d. (21 de 08 del 2018). *Escuela: CIENCIAS DE LA COMPUTACION E INFORMATICA*. Obtenido de ÉNFASIS EN CIENCIAS DE LA COMPUTACIÓN:
<http://www.cea.ucr.ac.cr/index.php/component/jdownloads/send/32-informatica/372-bachillerato-en-computacion-varios-enfasis-plan-1>
- Richard, S. L. (2002). *Mercadotecnia*.. México D.F: Compañía Editorial Continental.
- Roberto Hernández Sampieri, C. F. (2010). *Metodología de la investigación, Quinta Edición*. México D.F.: MCGRAW-HILL / INTERAMERICANA EDITORES, S.A. DE C.V. Recuperado el 14 de 06 del 2017

Rojas, J. C. (30 de marzo de 2017). *Parma Group TEC Costa Rica*. Obtenido de SIMPOSIO INTERNACIONAL SOBRE APLICACIONES DEL APRENDIZAJE AUTOMÁTICO: http://parma.ic-itcr.ac.cr/public/noticias_vermas.html

Rouse, M. (01 de 09 del 2018). *Infrastructure as a Service (IaaS)*. Obtenido de searchcloudcomputing.techtarget.com:
<https://searchcloudcomputing.techtarget.com/definition/Infrastructure-as-a-Service-iaaS>

Sampieri, R. H. (2014). *Metodología de la investigación 6a Edición*. México: MCGRAW-HILL / INTERAMERICANA EDITORES, S.A. DE C.V.

Savaram, R. (21 de junio del 2018). *Top Python Libraries for Machine Learning*. Obtenido de [datascienceplus](http://datascienceplus.com): <https://datascienceplus.com/top-python-libraries-for-machine-learning/>

Saxena, R. (30 de enero del 2017). *HOW DECISION TREE ALGORITHM WORKS*. Obtenido de [Dataaspirant](http://dataaspirant.com):
<http://dataaspirant.com/2017/01/30/how-decision-tree-algorithm-works/>

Schieltz, M. (05 de 03 del 2013). *Formas y características de una estructura organizativa*. Obtenido de pyme.lavoztx.com:
<https://pyme.lavoztx.com/formas-y-caractersticas-de-una-estructura-organizativa-5855.html>

SciKit-Learn. (25 de 01 del 2018). *Scikit-Learn*. Obtenido de ¿Quién está usando scikit-learn? : <https://scikit-learn.org/stable/testimonials/testimonials.html>

scikit-learn. (08 de 11 del 2018). *scikit-learn machine learning in Python*. Obtenido de scikit-learn.org: <https://scikit-learn.org/stable/index.html>

Seif, G. (26 de 12 del 2018). *towardsdatascience.com*. Obtenido de An Introduction to Scikit Learn: The Gold Standard of Python Machine Learning: <https://towardsdatascience.com/an-introduction-to-scikit-learn-the-gold-standard-of-python-machine-learning-e2b9238a98ab>

ServiceNow. (2017). *The Global CIO*. Santa Clara: ServiceNow.

Services, A. W. (08 de 11 del 2018). *¿Qué es Amazon Machine Learning?*
Obtenido de aws.amazon.com:
https://docs.aws.amazon.com/es_es/machine-learning/latest/dg/what-is-amazon-machine-learning.html

Services, A. W. (25 de 01 del 2018). *Amazon Web Services*. Obtenido de Crear un modelo de aprendizaje automático: <https://aws.amazon.com/es/getting-started/projects/build-machine-learning-model/>

Services, A. W. (13 de 11 del 2018). *Amazon Web Services*. Obtenido de Amazon SageMaker, Elija y optimice su algoritmo de aprendizaje automático: <https://aws.amazon.com/es/sagemaker/>

Services, A. W. (25 de 01 del 2019). *Amazon Web Services*. Obtenido de Características de Amazon SageMaker:
<https://aws.amazon.com/es/sagemaker/features/>

Services, A. W. (25 de 01 del 2019). *Amazon Web Services*. Obtenido de Amazon Machine Learning, Introducción: <https://aws.amazon.com/es/aml/>

Simple, E. (16 de 10 del 2018). *Definición de proveedor*. Obtenido de EconomíaSimple.net: <https://www.economiasimple.net/glosario/proveedores>

Simsek., B. (18 de 08 del 2004). *Libraries*. Obtenido de enderunix.org:
<http://www.enderunix.org/simsek/articles/libraries.pdf>

Skyose. (09 de 08 del 2018). *Skyose.com*. Obtenido de Full Microsoft Azure Machine Learning Studio Predictive Analysis Software Review:
<https://skyose.com/microsoft-azure-machine-learning-studio/>

Spark, A. (21 de 08 del 2018). *Apache Spark™ is a unified analytics engine for large-scale data processing*. Obtenido de Generalidades:
<https://spark.apache.org/>

Std, I. (1993). *IEEE Software Engineering Standard: Glossary of Software Engineering Terminology*. New York: IEEE Computer Society Press.

System, S. S. (s.f.). *Importancia del Machine Learning*. Recuperado el 07 de 10 del 2016, de SAS Statistical Analysis System:

http://www.sas.com/en_us/insights/analytics/machine-learning.html#machine-learning-importance

Tenerife, C. d. (13 de enero del 2019). *Producción y Operaciones*. Obtenido de creacionempresas.com: <http://www.creacionempresas.com/plan-de-viabilidad/que-es-un-plan-de-empresa-viabilidad/produccion-y-operaciones>

TensorFlow. (21 de 08 de 2018). *Información sobre TensorFlow*. Obtenido de TensorFlow.org: <https://www.tensorflow.org/>

Tocci, R. (2006). *Digital Systems: Principles and Applications (10th Edition)*. New Jersey: Prentice Hall.

UCR, E. d. (23 de 07 del 2018). *El Laboratorio de Reconocimiento de Patrones y Sistemas Inteligentes*. Obtenido de PRIS-LAB: <https://pris.eie.ucr.ac.cr/>

UCR, P.-L. E. (23 de 07 del 2018). *RISE, Pris-Lab EIE UCR*. Obtenido de <https://pris.eie.ucr.ac.cr/rise/>

University, C. (2016). *Meaning of “algorithm” in the English Dictionary*. Obtenido de Cambridge University.

University, C. M. (11 de 02 del 2019). *Master of Science in Machine Learning Financial Information*. Obtenido de Sitio Web Carnegie Mellon University: <https://www.ml.cmu.edu/academics/ms-finance.html>

University, O. (14 de 01 del 2019). *Definición de local en español*. Obtenido de Oxford Dictionary: <https://es.oxforddictionaries.com/definicion/local>

University, O. (15 de 01 del 2019). *Definición de proyectar*. Obtenido de Oxford

University: <https://es.oxforddictionaries.com/definicion/proyectar>

Vargas, D. (31 de 05 del 2018). *¿Qué hará la inteligencia artificial por nosotros?*

Obtenido de Sitio Web La Nación:

<https://www.nacion.com/brandvoice/proyectos/innova/que-hara-la-inteligencia-artificial-por-nosotros/HUJARKXOGVDTLGL3GBQLY3PIG4/story/>

Vargas, D. (07 de junio de 2018). *La Nación*. Obtenido de *¿Cómo puede ayudar la inteligencia artificial a mejorar los servicios de salud?:*

<https://www.nacion.com/brandvoice/proyectos/innova/como-puede-ayudar-la-inteligencia-artificial-a-mejorar-los-servicios-de-salud/GWMHIW7AABFLHI646B4USYLQUM/story/>

Venegas, J. U. (17 de 04 del 2017). *Simposio internacional mostró el presente y futuro en el innovador campo del aprendizaje automático*. Obtenido de Sitio Web TEC Costa Rica:

<https://www.tec.ac.cr/hoyeneltec/2017/04/17/simposio-internacional-mostro-presente-futuro-innovador-campo-aprendizaje-automatico>

Zaforas, M. (18 de 04 del 2016). *Machine Learning, la forma de hacer Big Data inteligente*. Recuperado el 03 de 11 del 2016, de ParadigmaDigital.com:

<https://www.paradigmadigital.com/dev/machine-learning-la-forma-big-data-inteligente/>

Apéndice

Apéndice A Respuesta de CAMTIC sobre la investigación.

RV: Tesis: Características de la oferta de procesamiento de datos utilizando Machine Lear...

🕒 Mensaje enviado con importancia Alta.

 Marco Zúñiga Ugalde
 Jue 22/6/2017, 15:27
 info@camtic.org

↩ ↶ → ✓

Buenas tardes,

El jueves 15 de junio asistí a las oficinas de CAMTIC para que me dieran respuesta de mi petición para mi tesis. Me mencionaron que **no** era posible brindarme el nombre de las empresas que participaron en el mapeo sectorial por cuestiones de privacidad.

Dado esto les agradecería que me respondieran por correo confirmando lo dicho para respaldar mi tesis y buscar otra fuente de información.

Gracias.
 Marco Zúñiga Ugalde.

De: Marco Zúñiga Ugalde
Enviado: viernes 2 de junio de 2017 05:04 p.m.
Para: info@camtic.org
Asunto: RV: Tesis: Características de la oferta de procesamiento de datos utilizando Machine Learning.

Buenas tardes,

Dirigido a Otto Rivera o encargado que pueda facilitar dicha información.

Soy Marco Zúñiga Ugalde, estudiante de la Universidad Técnica Nacional, estoy cursando licenciatura y estoy realizando un anteproyecto con el tema: ***Características de la oferta de procesamiento de datos utilizando Machine Learning.***

Para el anteproyecto necesito definir una población de estudio. Mi objetivo es tener una fuente fiable de información de las **empresas proveedoras de servicios de análisis y procesamiento de datos (Minería de datos, Big Data o Machine Learning) en Costa Rica**. Estas empresas supondrían estar en el Mapeo de sectorial en el subarea de **Tecnologías de la información.**

Si me pudieran facilitar la base de datos de las **empresas** (nombre, no requiero información privada por cuestiones de privacidad) de este tipo se los agradecería enormemente.

Gracias.
 Marco Zúñiga Ugalde.

Apéndice B Respuesta CINDE sobre la investigación.

Información solicitada

CM Carlos Morales <cmorales@cinde.org> ↶ ↷ → ...
Mié 26/4/2017, 13:44
Usted ☺

Buenas tardes, sobre la información que usted nos solicita, lamento informarle que al menos en nuestro caso no contamos con estadísticas o un estudio al respecto. Conocemos del caso de empresas multinacionales que realizan esa tarea; sin embargo, son ejemplos casuísticos y sería tratar directamente con ellas la información.

Saludos.

Carlos
CINDE

Apéndice C Respuesta del CONICIT sobre la investigación.

Re: Tesis: Características de la oferta de procesamiento de datos utilizando Machine Learn...

VJ Vinicio Porras Jiménez <vporras@conicit.go.cr>
 Jue 27/4/2017, 09:50
 Usted ☾

↩ ☹ → ∨

Estimado Marco.

Nosotros no dedicamos a la Minería de datos saludos

De: "Marco Zúñiga Ugalde" <MZU_92@hotmail.com>
Para: "conicit" <conicit@conicit.go.cr>
Enviados: Miércoles, 26 de Abril 2017 11:26:59
Asunto: Tesis: Características de la oferta de procesamiento de datos utilizando Machine Learning.

Buenos días,

Soy Marco Zúñiga Ugalde, estudiante de la Universidad Técnica Nacional, estoy cursando licenciatura y estoy realizando un anteproyecto con el tema: **Características de la oferta de procesamiento de datos utilizando Machine Learning**.

Para el anteproyecto necesito definir una población de estudio. Mi objetivo es tener una fuente fiable de información de las **empresas proveedoras de servicios de análisis y procesamiento de datos (Minería de datos, Big Data o Machine Learning) en Costa Rica**.

Si fuera el caso de no poseer una base de datos o información estadística de algún informe emitido favor responder el correo para justificar que no pude encontrar dicha base de información.

Gracias.
 Marco Zúñiga Ugalde.

--

 **Lic. Vinicio Porras Jiménez** | Tel.: (506) 2216-1500 | email: vporras@conicit.go.cr
 Coordinador Unidad de TICs | Fax: (506) 2216-1565 | www.conicit.go.cr

Apéndice D Respuesta MICITT sobre la investigación.

Re: Tesis: Características de la oferta de procesamiento de datos utilizando Machine Learn...

M Ministerio de Ciencia, Tecnología y Telecomunicaciones <micitt@micit.go.cr> ↶ ↷ → ✓
 Mié 26/4/2017, 12:59
 Usted; Johnny Pan ✓

Buenas tardes Marco.

En atención a su correo, copiamos el mismo al señor Pan, Jefe de la Unidad de Servicios Tecnológicos del MICITT para que le pueda orientar.

Atentamente,

El 26 de abril de 2017, 10:47, Marco Zúñiga Ugalde <MZU_92@hotmail.com> escribió:

Buenos días,

Soy Marco Zúñiga Ugalde, estudiante de la Universidad Técnica Nacional, estoy cursando licenciatura y estoy realizando un anteproyecto con el tema: **Características de la oferta de procesamiento de datos utilizando Machine Learning**.

Para el anteproyecto necesito definir una población de estudio. Mi objetivo es tener una fuente fiable de información de las **empresas proveedoras de servicios de análisis y procesamiento de datos (Minería de datos, Big Data o Machine Learning)**.

Si fuera el caso de no poseer una base de datos o información estadística de algún informe emitido favor responder el correo para justificar que no pude encontrar dicha base de información.

Gracias.

Ministerio de Ciencia, Tecnología y Telecomunicaciones de Costa Rica

Ministerio de Ciencia, Tecnología y Telecomunicaciones de Costa Rica
 A 29 mil personas les gusta esta página, incluidos Sofia Zn y 2 amigos
 Organización gubernamental

26/04/2017 11:08

Buenos días,
 Soy Marco Zúñiga Ugalde, estudiante de la Universidad Técnica Nacional, estoy cursando licenciatura y estoy realizando un anteproyecto con el tema: Características de la oferta de procesamiento de datos utilizando Machine Learning.
 Para el anteproyecto necesito definir una población de estudio. Mi objetivo es tener una fuente fiable de información de las empresas proveedoras de servicios de análisis y procesamiento de datos (Minería de datos, Big Data o Machine Learning).
 Si fuera el caso de no poseer una base de datos o información estadística de algún informe emitido favor responder el correo para justificar que no pude encontrar dicha base de información.
 Gracias.
 Marco Zúñiga Ugalde.

Si pueden responder a mzu_92@hotmail.com

Opciones

- Buscar en la conversación
- Administrar mensajes
- Calificar experiencia
- Notificaciones

Enlace de Messenger

m.me/micitr

Escribe un mensaje...

📎 🗨️ GIF 😊 🗣️ 📷 👍

Re: Tesis - Machine Learning

DK Diana Montero Katchan <diana.montero@micitt.go.cr> ↶ ↷ → ✓
 Jue 16/2/2017, 10:03
 Usted; Federico Torres Carballo; David Bullon Patton; Jose Luis Orozco Roses ☺

Estimado Marco.

Reciba un cordial saludo de mi parte.

Le comento que el Ministerio de Ciencia, Tecnología y Telecomunicaciones, actualmente no cuenta con proyectos en el tema de inteligencia artificial.

Le brindo a continuación los contactos de los profesionales que le podrán facilitar información en dicho tema.

- José Castro, profesor de computación del Instituto Tecnológico de Costa Rica (TEC), jose.r.castro@gmail.com
- Tomás de Camino, profesor de computación en maestría del TEC, tomas.decamino@gmail.com
- Alex Murillo del Centro de Investigaciones en Matemáticas Puras y Aplicadas (CIMPA) de la Universidad de Costa Rica (UCR), alex.murillo@ucr.ac.cr, tel: 2511-5889

Cualquier otra consulta, estoy a su disposición.

Cordialmente,

--



Ph.D. Diana Montero Katchan
 Asesora, Dirección de Investigación y Desarrollo Tecnológico
 Ministerio de Ciencia, Tecnología y Telecomunicaciones
 E-mail: diana.montero@micitt.go.cr
 Teléfono: (506) 2539-2311
www.micitt.go.cr

----- Mensaje reenviado -----

De: **Ministerio de Ciencia, Tecnología y Telecomunicaciones** <micitt@micitt.go.cr>

Fecha: 8 de febrero de 2017, 23:20

Asunto: Re: Tesis - Machine Learning

Para: Marco Zúñiga Ugalde <MZU_92@hotmail.com>

Cc: David Bullon Patton <david.bullon@micitt.go.cr>, Federico Torres Carballo <federico.torres@micitt.go.cr>, Jose Luis Orozco Roses <contraloria@micitt.go.cr>

Buenas noches.

En seguimiento a su consulta copiamos a los señores Bullón y Torres, Directores de Innovación y de Investigación y Desarrollo del MICITT, para que le puedan orientar.

Apéndice E Respuesta de la empresa CRUX de la encuesta.

Tesis: Estudio de características tecnológicas Machine Learning a nivel nacional - UTN

Randald Alfonso Rodriguez Rodriguez <rrodriguez@cruxconsultores.com>
 Para: "Marco Zúñiga U." <mazuga92@gmail.com>

27 de marzo de 2018, 17:27

Buenas tardes,

Disculpe la tardanza para contestar.

Solo para aclarar, esta empresa trabaja con servicios de inteligencia artificial que no corresponden a Machine Learning, por lo que no podríamos responder a las preguntas del cuestionario, entonces no le sería de mucha ayuda.

Disculpas por la demora y el malentendido.

Saludos

De: Marco Zúñiga U. [mailto:mazuga92@gmail.com]
Enviado el: miércoles, 21 de marzo de 2018 10:24 a. m.
Para: Randald Alfonso Rodriguez Rodriguez <rrodriguez@cruxconsultores.com>
Asunto: Tesis: Estudio de características tecnológicas Machine Learning a nivel nacional - UTN

Apéndice F Respuesta empresa Grupo Prides sobre la encuesta.

Byron Roman Villalobos <broman@prides.net>
 Para: "Marco Zúñiga U." <mazuga92@gmail.com>

25 de abril de 2018, 17:01

Buenas tardes Marco,

Primero las disculpas por no contestar su primer correo, con respecto ala participación en la encuesta con gusto, por temas de tiempo prefiero el cuestionario.

Muchas gracias
 Lic. Byron Román Villalobos
 CSM, PSM I, SMAC ,MCP,MS, MCTS, MCPD, MCSA, MCSA
 División de Investigación y Desarrollo
 Grupo Prides
broman@prides.net

Apéndice G Respuesta empresa Logosoft sobre la encuesta.

Tesis: Estudio de características tecnológicas Machine Learning a nivel nacional - UTN

6 mensajes

Marco z <mazuga92@gmail.com>
Para: gabriela@logosoftcr.com

12 de marzo de 2018, 16:32

Buenas tardes,

Gabriela,

Soy Marco Antonio Zúñiga Ugalde estudiante de la Universidad Técnica Nacional, estoy realizando un trabajo de tesis para el grado de licenciatura. Quiero solicitar su participación en mi tesis.

Mi método de recolección de información son **entrevistas personales**, pero en dado caso no se pueda concretar la entrevista personal estaré enviando un **cuestionario de Google Forms** para que pueda ser participe de este estudio.

¿Tiene disponibilidad para una entrevista personal para agendar?

Adjunto un volante informativo para poder conocer de que trata mi tesis, además de las preguntas de cuestionario que serán utilizadas en la entrevista personal.

Un cordial saludo.

--

Marco Antonio Zúñiga Ugalde

2 archivos adjuntos

 **Volante Informativov2.pdf**
1398K

 **Cuestionariov2 - Machine Learning.pdf**
25K

Gabriela Sánchez <gabriela@logosoftcr.com>
Para: "Marco Zúñiga U." <mazuga92@gmail.com>

25 de abril de 2018, 17:04

Gracias por ponerse en contacto, le informamos que esta cuenta de correo está pronta a cerrarse por cambio de dominio,

Sírvase escribir a gabriela@logosoftla.com.

gabriela@logosoftla.com <gabriela@logosoftla.com>
Para: "Marco Zúñiga U." <mazuga92@gmail.com>

30 de abril de 2018, 10:22

Hola Marco,

Gracias por tomarnos en cuenta. Me gustaría me cuentes un poco más sobre la tesis, que veo que el tema es sobre "Machine Learning", para así saber quién sería la persona idónea, ya sea para una entrevista o para contestar el formulario.

Saludos,

Gabriela Sánchez Cortés
Especialista en Capital Humano



Tel. (506) 2253 0800
Cel. (506) 7013 0805
gabriela@logosoftla.com
www.logosoftla.com

Logosoft



Antes de imprimir este mensaje, asegúrese de que sea necesario.
¡Proteger el medio ambiente está también en sus manos!

Marco Zúñiga U. <mazuga92@gmail.com>
Para: gabriela@logosoftla.com

30 de abril de 2018, 11:50

Buenos días,

Gabriela,

Soy Marco Antonio Zúñiga Ugalde estudiante de la Universidad Técnica Nacional, estoy realizando un trabajo de tesis para el grado de licenciatura. La tesis trata sobre tener un mejor panorama de la tecnológica de machine learning en Costa Rica, analizando las características de dicha tecnología, retos a futuro. No hay preguntas sobre clientes, costos, todo es sobre la tecnología, igualmente dentro de mi estudio solo se da un resultado global y el nombre de la empresa no figura dentro de las conclusiones.

El cuestionario esta dirigido a alguien con conocimientos en machine learning, minería de datos, inteligencia de negocio, desarrollados de software y científicos de datos.

Adjunto un volante informativo para poder conocer de que trata mi tesis, ademas de las preguntas de cuestionario que serán utilizadas en la entrevista personal.

Un cordial saludo.

[El texto citado está oculto]

--

Marco Zúñiga Ugalde

2 archivos adjuntos

 **Cuestionariov2 - Machine Learning.pdf**
25K

 **Volante Informativov2.pdf**
1398K

Apéndice H Respuesta de la empresa MediaLAB sobre la encuesta.

Tesis: Estudio de características tecnológicas Machine Learning a nivel nacional - UTN

Diana Flaque <diana@medialabla.com>
 Para: "Marco Zúñiga U," <mazuga92@gmail.com>
 Cc: Guillermo Montealegre <guillermo@medialabla.com>

22 de mayo de 2018, 8:39

Buenos días Marco,

Estuve revisando las preguntas y lamentablemente nosotros no le podemos ayudar porque no hacemos Machine Learning.

Gracias,



Apéndice I Respuesta de la empresa SimMachines sobre la encuesta.

Tesis: Estudio de características tecnológicas Machine Learning a nivel nacional - UTN

Julián Rodríguez <julianrodriguez@simmachines.com>
 Para: "Marco Zúñiga U," <mazuga92@gmail.com>

4 de abril de 2018, 10:45

Buenas tardes Marco,

Me gustaria poder ayudarle con la encuesta solicitada, Julio sibaja me remitio el google form que ud elaboro y lo examine a profundidad.

Lamentablemente, por motivos de privacidad y confidencialidad empresarial, las respuestas a algunas de las preguntas que ud plantea son de caracter confidencial y por ende no me es posible compartirle dicha informacion.

Saludos,

Julian Rodriguez
 Site Manager simMachines Costa Rica
 +506-8898-0291
www.simMachines.com



[El texto citado está oculto]

Apéndice J Respuesta de la empresa Xchematic sobre la encuesta.

Tesis: Estudio de características tecnológicas Machine Learning a nivel nacional - UTN

3 mensajes

Marco Zúñiga U. <mazuga92@gmail.com>
Para: ccerrato@xchematic.com

4 de abril de 2018, 18:27

Buenas tardes,

Carlos,

Me hizo referencia el señor Alejandro Sanchez el laboró en Cognitiva.

Soy Marco Antonio Zúñiga Ugalde estudiante de la Universidad Técnica Nacional, estoy realizando un trabajo de tesis para el grado de Licenciatura. Quiero solicitar su participación en mi tesis.

Mi método de recolección de información son **entrevistas personales**, pero en dado caso no se pueda concretar la entrevista personal estaré enviando un **cuestionario de Enlace Google Forms** para que pueda ser partícipe de este estudio.

¿Tiene disponibilidad para una entrevista personal para agendar?

Adjunto un volante informativo para poder conocer de que trata mi tesis, ademas de las preguntas de cuestionario que serán utilizadas en la entrevista personal.

Un cordial saludo,

--

Marco Zúñiga Ugalde

2 archivos adjuntos

 **Cuestionariov2 - Machine Learning.pdf**
25K

 **Volante Informativov2.pdf**
1398K

Carlos Cerrato <ccerrato@xchematic.com>
Para: "Marco Zúñiga U." <mazuga92@gmail.com>

20 de abril de 2018, 12:40

Hola Marco

Ya te envíe el formulario con los datos de Xchematic y lo que estamos haciendo con Machine Learning

Saludos

[El texto citado está oculto]

Apéndice K Volante informativo de tesis.



Ingeniería del Software
Tesis de Licenciatura

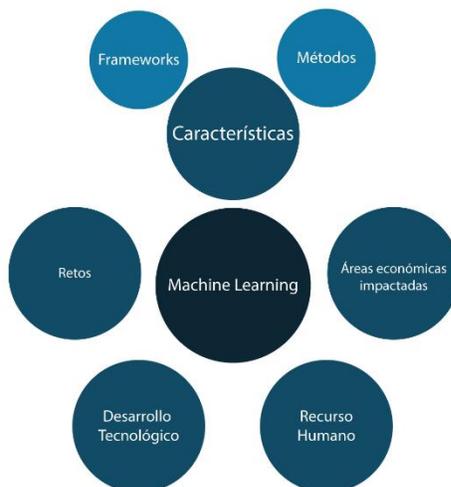
Estudio técnico de la oferta de servicios tecnológicos digitales con machine learning por medio de las empresas proveedoras que permita análisis las características principales del machine learning en la gran área metropolitana de Costa Rica en el periodo 2017-2018.

Estudiante: Marco Antonio Zúfiga Ugalde

La investigación hace referencia acerca del análisis de la oferta de los servicios de machine learning de las empresas proveedoras en Costa Rica en el periodo 2017 al 2018. Dentro de la investigación se tiene como objetivo identificar y analizar las principales características tecnológicas de machine learning. Permite compartir información oportuna y relevante sobre empresas de tecnologías digitales localizadas en el territorio costarricense.

La investigación, también tiene su importancia en que puede ser considerada como un panorama de largo plazo en la toma de decisiones, para empresas que piensen utilizar esta tecnología en Costa Rica.

Este es un gran esfuerzo que se hace primeramente para identificar a las empresas proveedoras de esta tecnología, segundo para analizar las características que componen esta tecnología y sus casos de uso.



Reglas de la visita

Se debe explicar e informar sobre el proceso de la investigación a cada empresa y cómo será el protocolo de recolección de información y aclarar dudas. A continuación, una lista del protocolo:

- Propósito general del estudio.
- Motivaciones para el sujeto entrevistado (importancia de su participación).
- Agradecimiento.
- Tiempo aproximado de respuesta (un promedio o rango). Lo suficientemente abierto para no presionar al participante, pero tranquilizarlo.
- Espacio para que firme o indique su consentimiento.
- Identificación de quién o quiénes lo aplican.
- Explicar brevemente cómo se procesarán los cuestionarios y una cláusula de confidencialidad del manejo de la información individual.
- Instrucciones iniciales claras y sencillas (cómo responder en general, con ejemplos si se requiere).

Agradecimiento

Este estudio ayuda a comprender el nivel de desarrollo de la tecnología machine learning, esto es gracias al conocimiento y experiencia que las empresas generan y por ello se solicita de manera mas formal la participación.

Los resultados serán útiles para conocer temas importantes en cuanto a desarrollo, fuerza laboral y retos. Todas las estadísticas serán compartidas con la comunidad interesada así con las universidades que son un factor determinante en la generación de profesionales.

La comunidad estará muy agradecida por alcanzar el reto de conocer en primera instancia lo que acontece a nivel tecnológico en el país.

Apéndice L Cuestionario para la recolección de información.



Tesis de grado Licenciatura, Universidad Técnica Nacional

Elaborar un estudio técnico de la oferta de servicios tecnológicos digitales con *machine learning* por medio de las empresas proveedoras que permita el análisis de sus características principales en la gran área metropolitana de Costa Rica durante el periodo 2017-2018.

Estudiante Marco Antonio Zúñiga Ugalde

Fecha:

Nombre de la empresa:

Nombre del entrevistados:

Puesto del entrevistado:

PREGUNTAS

FRAMEWORKS

1. ¿Cuáles son los framework más utilizados en sus productos? (Azure ML, Amazon ML, Tensorflow, Torch)

MÉTODOS

2. ¿Qué tipos de métodos son lo más utilizados en sus productos?
3. ¿Qué estilos de aprendizaje utilizan sus productos? Supervised, Unsupervised, Semi supervised

4. ¿Qué tipo de problema está enfocado el machine learning? (predicción, clasificación, generación)

AREAS ECONÓMICAS / INDUSTRIA IMPACTADA / RECURSO HUMANO

5. ¿Cuáles áreas económicas están destinados los productos que utilizan machine learning en su compañía?
6. ¿Cuáles son los casos de uso donde se implementa machine learning? (poner los casos de uso)
7. ¿Qué ventajas son las que puede contar un cliente con sus productos que utilizan machine learning?
8. ¿Cuáles han sido los retos de implementar una tecnología como machine learning en sus procesos o productos?
9. ¿Se consume el producto de manera local o en el extranjero?
10. ¿Cuáles son las necesidades o retos para desarrollar más productos con Machine Learning?
11. ¿Que se espera para los siguientes años en materia de desarrollo de machine learning en Costa Rica?

12. ¿Cuáles son los puestos de trabajo que desarrollan tecnologías de machine learning en su empresa? (Programador, científico de datos, especialista en machine learning)

13. ¿Cuál es el método de capacitación de su personal, ya sea en la misma empresa, desde la academia o que otra fuente de capacitación manejan?

Marco Zúñiga Ugalde

Firma del entrevistado

Apéndice M Formulario de Google con las preguntas del cuestionario.



The image shows a Google Form interface. At the top left is the logo for Universidad Técnica Nacional (UTN). The main title of the form is 'MACHINE LEARNING' in large, bold, white letters against a dark background with a futuristic, industrial aesthetic. Below the title, the form content is displayed in a white box on a light blue background. The text in the form includes the title 'Machine Learning una tecnología de impacto', the author 'Tesis de grado Licenciatura, Universidad Técnica Nacional Ingeniería del Software', the objective 'Objetivo: Elaborar un estudio técnico de la oferta de servicios tecnológicos digitales con machine learning por medio de las empresas proveedoras que permita el análisis de sus características principales en la gran área metropolitana de Costa Rica durante el periodo 2017-2018.', and the author's name 'Estudiante Marco Antonio Zuñiga Ugalde'. A 'SIGUIENTE' button is visible below the text. At the bottom of the form, there is a small disclaimer: 'Este contenido no ha sido creado ni aprobado por Google. Notificar uso inadecuado - Condiciones del servicio' and the 'Google Formularios' logo.

UTN
Universidad
Técnica Nacional

MACHINE LEARNING

Machine Learning una tecnología de impacto

Tesis de grado Licenciatura, Universidad Técnica Nacional
Ingeniería del Software

Objetivo:

Elaborar un estudio técnico de la oferta de servicios tecnológicos digitales con machine learning por medio de las empresas proveedoras que permita el análisis de sus características principales en la gran área metropolitana de Costa Rica durante el periodo 2017-2018.

Estudiante Marco Antonio Zuñiga Ugalde

SIGUIENTE

Nunca envíes contraseñas a través de Formularios de Google.

Este contenido no ha sido creado ni aprobado por Google. Notificar uso inadecuado - Condiciones del servicio

Google Formularios

Apéndice N Machote de Carta de confidencialidad.

Alajuela, Fecha.

Marco Antonio Zúñiga Ugalde

Estudiante de Licenciatura Universidad Técnica Nacional

Yo, **Marco Antonio Zúñiga Ugalde**, mayor de edad, cédula 206970733, me comprometo a no divulgar ni utilizar en mi conveniencia personal la distinta información empresarial obtenida mediante mi trabajo de tesis de grado **de Licenciatura en la Universidad Técnica Nacional**, ni proporcionaré a otras personas o empresas, ni de manera verbal, ni de manera escrita, ya sea directa o indirectamente, información de las empresas participantes o alguna que pueda perjudicar los intereses de la **Empresa**.

No revelaré ninguna información técnica sobre los productos ni quién es su desarrollador.

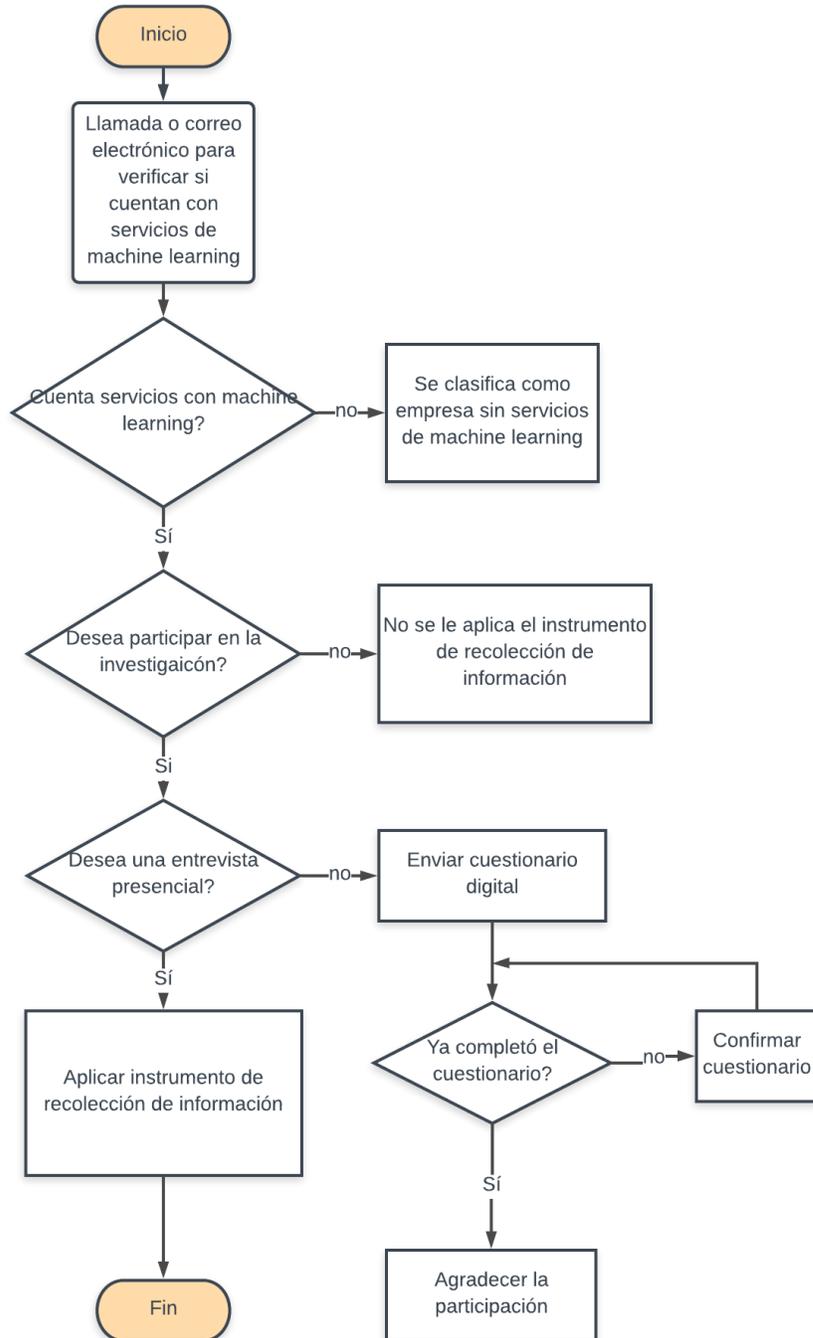
Mi tesis únicamente se basará en determinar cuáles son las características más importantes del *machine learning* y sus áreas económicas en Costa Rica.

Se recopilará información de manera anónima por lo que en los resultados no se darán a conocer las empresas incluidas en la investigación.

Las respuestas se filtrarán y se analizarán para brindar resultados a estudiantes, empresas, universidades y público en general que quieran obtener conocimiento de las características del *machine learning* en Costa Rica.

Las empresas investigadas podrán tener comunicación conmigo por medio del correo mazuga92@gmail.com

Apéndice O Diagrama de flujo del método de medición.



Anexo

Anexo 1 Visión por computadora. Plan de estudio Ingeniería Eléctrica, UCR.

Universidad de Costa Rica
Sistema de Aplicaciones Estudiantiles SAE
Módulo Planes de Estudio

FASCÍCULO UNIDAD ACADÉMICA

Escuela: INGENIERIA ELECTRICA

Carrera 420201 BACH.Y LIC.EN INGENIERIA ELECTRICA

Plan de Estudio 1 PLAN DE 1985

Bloques Optativos

Bloque Optativo OPT399 BLOQUE OPTATIVO

Observaciones:

Curso	Nombre del curso	*Horas*				Cred.	Requisitos y Req. Equivalentes	Correquisit
		T	P	L	TP			
IE0407	ELECTROMAGNETISMO II	3	2	0	3	IE0307		
IE0523	CIRCUITOS DIGITALES II	3	1	0	3	IE0323 Equiv.: IE0423		
IE1003	TEMAS ESPECIALES EN INGENIERÍA	3	0	0	3	IE0409		
IE0117	PROGRAMACIÓN BAJO PLATAFORMAS ABIERTAS	3	0	3	0	CI0202		
IE0217	ESTRUCTURAS ABSTRACTAS DE DATOS Y ALGORITMOS PARA INGENIERÍA	3	0	3	0	IE0117		
IE0445	METROLOGÍA EN INGENIERÍA ELÉCTRICA	3	2	0	0	3	IE0308; IE0315; IE0405	
IE0559	ENERGÍA SOLAR FOTOVOLTAICA	3	0	0	0	3	IE0471 Equiv.: IE1071	
IE0571	DISEÑO ELÉCTRICO INDUSTRIAL II	3	0	0	0	3	IE0471 Equiv.: IE1071	
IE0633	AUTOMATIZACIÓN INDUSTRIAL	3	0	3	0	3	IE0323; IE0431	
IE0311	DISPOSITIVOS SEMICONDUCTORES	4	0	0	0	3	FS0410; IE0405 Equiv.: I0305	
IE0321	ESTRUCTURAS DE COMPUTADORES DIGITALES I	3	1	0	0	3	IE0323 Equiv.: IE0423	
IE0411	MICROELECTRÓNICA: SISTEMAS EN SILICIO	4	0	0	0	3	IE0311 Equiv.: IE0109; IE0323 Equiv.: IE0423	
IE0424	LABORATORIO DE CIRCUITOS DIGITALES	1	0	5	0	3	IE0308; IE0321 Equiv.: IE0503; IE0323 Equiv.: IE0423	
IE0425	REDES DE COMPUTADORES	4	0	0	0	3	IE0321 Equiv.: IE0503; IE0527	
IE0433	CONTROL E INSTRUMENTACIÓN DE PROCESOS INDUSTRIALES	3	0	3	0	3	IE0431	
IE0437	FOTÓNICA	4	1	0	0	3	IE0305; IE0307	
IE0439	OPTOELECTRÓNICA	4	1	0	0	3	IE0307; IE0513	
IE0447	PROCESAMIENTO DIGITAL DE SEÑALES I	3	1	0	0	3	IE0409; IE0413	
IE0449	VISIÓN POR COMPUTADOR	3	0	0	3	3	IE0305	
IE0457	CONSERVACIÓN DE ENERGÍA ELÉCTRICA	3	1	0	0	3	IE0479 Equiv.: IE0470; IE0615 Equiv.: IE0516	
IE0461	CENTRALES Y SUBESTACIONES	3	1	1	0	3	IE0615 Equiv.: IE0516	
IE0465	REDES DE DISTRIBUCIÓN Y TRANSMISIÓN	3	1	0	0	3	IE0365 Equiv.: IE0665	
IE0467	DISTRIBUCIÓN ELÉCTRICA SUBTERRÁNEA	4	0	0	0	3	IE0307; IE0416 Equiv.: IE0315	
IE0469	SISTEMAS DE POTENCIA I	3	2	0	0	3	IE0365 Equiv.: IE0665; IE0616 Equiv.: IE0516	
IE0505	TÉCNICAS DE OPTIMIZACIÓN	3	0	0	0	3	IE0409	

Anexo 2 Inteligencia Artificial, Ciencias de la computación e informática, UCR.

Universidad de Costa Rica
Sistema de Aplicaciones Estudiantiles SAE
Módulo Planes de Estudio

FASCÍCULO UNIDAD ACADÉMICA

Escuela: CIENCIAS DE LA COMPUTACION E INFORMATICA

Carrera	420705	BACH. EN COMPUTACION CON VARIOS ENFASIS						
Plan de Estudio	1	BACHILLERATO EN COMPUTACIÓN CON VARIOS ÉNFASIS						
Enfasis	1	ÉNFASIS EN CIENCIAS DE LA COMPUTACIÓN						
			Horas					
Ciclo	Curso	Nombre del curso	T	P	L	TP	Cred. Requisitos y Req. Equivalentes	
5	CI0120	ARQUITECTURA DE COMPUTADORAS	5	0	0	0	4 CI0114; CI0118	
5	MA1005	ECUACIONES DIFERENCIALES	5	0	0		4 MA1002 Equiv.: MA1023 o MA0292 o MA0232; MA1004 Equiv.: MA0294 o MA0232	
			Grado: BACHILLERATO				Créditos ciclo 5 : 8	
6	CI0129	INTELIGENCIA ARTIFICIAL	5	0	0	0	4 CI0115	
			Grado: BACHILLERATO				Créditos ciclo 6 : 4	
7	CI0130	METODOS DE MODELADO Y OPTIMIZACIÓN	5	0	0	0	4 CI0115	
7	CI0131	DISEÑO DE EXPERIMENTOS	5	0	0	0	4 CI0115	
7	CI0132	TEORÍA DE LA COMPUTACIÓN	5	0	0	0	4 CI0115; CI0129	
7	CI0146	COMPILADORES	5	0	0	0	4 CI0120; CI0129	
			Grado: BACHILLERATO				Créditos ciclo 7 : 16	
8	CI0134	INVESTIGACIÓN EN CIENCIAS DE LA COMPUTACIÓN	5	0	0	0	4 CI0130; CI0131; CI0132	
8	OPT1061	OPTATIVOS DEL ÉNFASIS EN CIENCIAS DE LA COMPUTACIÓN					8	
			Grado: BACHILLERATO				Créditos ciclo 8 : 12	
			Total créditos enfasis 1 : 40					
			Total cursos enfasis 1 : 9					

Anexo 3 Bloque optativo en los planes de Bachillerato y Licenciatura, Computación e informática, UCR

Universidad de Costa Rica
Sistema de Aplicaciones Estudiantiles SAE
Módulo Planes de Estudio

FASCÍCULO UNIDAD ACADÉMICA

Escuela: CIENCIAS DE LA COMPUTACION E INFORMATICA

Carrera 420703 BACH.Y LIC.EN COMPUTACION E INFORMATICA
 Plan de Estudio 4 PLAN NUEVO A REEMPLAZAR LOS ÉNFASIS VIGENTES

Bloques Optativos

Bloque Optativo OPT781 BLOQUE OPTATIVO

Observaciones:

Curso	Nombre del curso	*Horas*				Cred.	Requisitos y Req. Equivalentes
		T	P	L	TP		
CI2400	TEMAS ESPECIALES DE INGENIERÍA DE SISTEMAS	4	0	0		4	CI1430; CI1431
CI2402	GRAFICACION INTERACTIVA I	4	0	0		4	CI1430; CI1431
CI2403	ORGANIZACION Y LENGUAJES DE PROGRAMACION	4	0	0		4	CI1322
CI2405	PROGRAMACION LINEAL	4	0	0		4	CI1453
CI2406	GRAFICACION INTERACTIVA II	4	0	0	0	4	CI2402
CI2408	PROGRAMACION CONCURRENTE	4	0	0		4	CI1430; CI1431
CI2411	INTRODUCCIÓN A LA MULTIMEDIA	4	0	0		4	CI1330; CI1331
CI2412	SISTEMAS DE INFORMACIÓN GEOGRÁFICOS	4	0	0		4	CI1312
CI2413	DESARROLLO DE APLICACIONES PARA INTERNET	4	0	0		4	CI1320; CI1321; CI1330; CI1331
CI2414	RECUPERACIÓN DE INFORMACIÓN	4	0	0		4	CI1330; CI1331
CI2451	DISEÑO DE INTERACCIÓN PERSONA - COMPUTADOR	4	0	0	0	4	CI1430; CI1431
CI2452	VALIDACIÓN Y VERIFICACIÓN DE SOFTWARE	4	0	0	0	4	CI1430; CI1431
CI2453	INGENIERIA DE SOFTWARE ASISTIDA POR COMPUTADORA	4	0	0	0	4	CI1330; CI1331
CI2454	SERVICIOS WEB	4	0	0	0	4	CI1320; CI1321; CI1430; CI1431
CI2455	ARQUITECTURA DE SOFTWARE	4	0	0	0	4	CI1330; CI1331
CI2600	TEMAS ESPECIALES PARA INTELIGENCIA ARTIFICIAL	4	0	0		4	CI1441
CI2602	LÓGICA FORMAL PARA COMPUTACIÓN	4	0	0		4	CI1441
CI2603	APRENDIZAJE AUTOMATICO	4	0	0	0	4	CI2607
CI2604	LA MENTE Y EL COMPUTADOR	4	0	0		4	CI1441
CI2605	REPRESENTACION DEL CONOCIMIENTO	4	0	0	0	4	CI2607
CI2606	SISTEMAS EXPERTOS	4	0	0		4	CI1441
CI2607	PROGRAMACION LOGICA	4	0	0		4	CI1441
CI2608	INTRODUCCION A LA NEUROCOMPUTACION	4	0	0	0	4	CI2604

Anexo 4 Bloque optativo en los planes de Bachillerato y Licenciatura, Computación e informática, UCR

Universidad de Costa Rica
Sistema de Aplicaciones Estudiantiles SAE
Módulo Planes de Estudio

FASCÍCULO UNIDAD ACADÉMICA

Escuela: CIENCIAS DE LA COMPUTACION E INFORMATICA

Carrera 420703 BACH.Y LIC.EN COMPUTACION E INFORMATICA
 Plan de Estudio 4 PLAN NUEVO A REEMPLAZAR LOS ÉNFASIS VIGENTES

Bloques Optativos

Bloque Optativo OPT781 BLOQUE OPTATIVO

Observaciones:

Curso	Nombre del curso	*Horas*				Cred.	Requisitos y Req. Equivalentes
		T	P	L	TP		
CI2651	INTELIGENCIA ARTIFICIAL	4	0	0	0	4	CI1441
CI2652	PROCESAMIENTO DIGITAL DE SEÑALES BIOLÓGICAS	4	0	0	0	4	CI1310; CI1311; CI1352; MA0292 Equiv.: MA0429
CI2653	PROCESAMIENTO DIGITAL DE IMÁGENES	4	0	0		4	CI1352; MA0294
CI2654	RECONOCIMIENTO, ANÁLISIS Y CLASIFICACIÓN DE FORMAS	4	0	0	0	4	CI2652
CI2655	REDES NEURONALES	4	0	0	0	4	CI1441; CI1453
CI2656	AGENTES Y SISTEMAS MULTIAGENTES	4	0	0	0	4	CI1441
CI2657	ROBÓTICA	4	0	0	0	4	CI1322; CI1323
CI2700	TEMAS ESPECIALES DE TEORÍA DE LA COMPUTACIÓN	4	0	0		4	CI1322; MA0292 Equiv.: MA0429
CI2701	EVALUACIÓN DE LENGUAJES DE PROGRAMACIÓN	4	0	0		4	CI1322; CI1330; CI1331
CI2702	TEORÍA DE LA INFORMACIÓN	4	0	0		4	CI1320; CI1321; CI1323
CI2703	COMPUTABILIDAD	4	0	0	0	4	CI2704
CI2704	TEORÍA DE AUTÓMATAS Y LENGUAJES FORMALES	4	0	0		4	CI1322; CI1330; CI1331
CI2706	COMPLEJIDAD COMPUTACIONAL	4	0	0		4	CI1221; MA0292 Equiv.: MA0429
CI2707	ANÁLISIS DE ALGORITMOS	4	0	0		4	CI1221; MA0292 Equiv.: MA0429
CI2800	TEMAS ESPECIALES DE INVESTIGACION DE OPERACIONES	4	0	0		4	CI1453
CI2801	TECNICAS ESTADISTICAS PARA COMPUTACION	4	0	0	0	4	CI1453
CI2802	MODELOS MATEMATICOS	4	0	0	0	4	CI1453
CI2803	SIMULACION	4	0	0		4	CI1320; CI1321
CI2804	MODELOS LINEALES Y NO LINEALES	4	0	0	0	4	CI1453
CI2805	PROGRAMACION DINAMICA	4	0	0	0	4	CI1453
CI2806	METODOS Y ANALISIS NUMERICOS I	4	0	0	0	4	CI1453
CI2807	TEORIA DE JUEGOS	4	0	0	0	4	CI1453

Anexo 5 Inteligencia artificial, Ingeniería en sistemas de información, UNA.

INGENIERIA EN SISTEMAS DE INFORMACIÓN CON GRADO DE BACHILLERATO
Y SALIDA LATERAL DE DIPLOMADO EN PROGRAMACIÓN DE APLICACIONES INFORMÁTICAS
Estudiantes que ingresan en el 2012



II Ciclo

CÓDIGO	NOMBRE	CRÉDITOS	REQUISITOS
EIF400	Paradigmas de programación	4	EIF206
EIF401	Ingeniería de sistemas II	4	EIF210
EIF402	Administración de bases de datos	4	EIF212 EIF211
EIF412	Investigación de operaciones y sus aplicaciones	3	MAT005 MAT006
EIF407	Liderazgo y organización	3	EIF404

18 créditos

IV Nivel

I Ciclo

CÓDIGO	NOMBRE	CRÉDITOS	REQUISITOS
EIF411	Diseño y programación de plataformas móviles	4	EIF209
EIF406	Ingeniería de sistemas III	4	EIF401
	Optativa	3	Ver cursos optativos
	Optativa	3	Ver cursos optativos
EIF413	Métodos de investigación científica en informática	3	MAT006

17 créditos

II Ciclo

CÓDIGO	NOMBRE	CRÉDITOS	REQUISITOS
EIF408	Proyectos y su aplicación en la organización (PPS)	5	EIF209 EIF401 EIF402
EIF409	Aplicaciones informáticas globales	4	EIF209 EIF401 EIF402
	Optativa	3	Ver cursos optativos
	Optativa	3	Ver cursos optativos
EIF410	Informática y Sociedad	2	EIF401

17 créditos

BACHILLERATO 140 CRÉDITOS.

OPTATIVOS DISCIPLINARIOS

CODIGO	NOMBRE	CRÉDITOS	REQUISITOS
EIF1000	Dispositivos para comunicaciones de datos	3	EIF208
EIF4200	Inteligencia artificial	3	EIF207
EIF4210	Análisis de algoritmos	3	EIF207 MAT002
EIF4220	Diseño de interfaces de usuario	3	EIF206
EIF4230	Introducción a la programación de dispositivos móviles		EIF209
EIF4240	Desarrollo de aplicaciones educativas	3	EIF206
EIF4250	Diseño de ambientes multimediales	3	EIF206
EIF4260	Diseño de ambientes de aprendizaje	3	EIF206
EIF4270	Robótica	3	EIF204
EIF4280	Fundamentos de programación web	3	EIF206

Anexo 6 Minería de datos 1 y 2, Ingeniería en sistemas de información, UNA

INGENIERIA EN SISTEMAS DE INFORMACIÓN CON GRADO DE BACHILLERATO
Y SALIDA LATERAL DE DIPLOMADO EN PROGRAMACIÓN DE APLICACIONES INFORMÁTICAS

Estudiantes que ingresan en el 2012



			EIF211
EIF431O	Administración de servidores basados en software libre	3	EIF212 EIF402
EIF432O	Introducción a la programación paralela en arquitecturas multinúcleo	3	EIF207 EIF212
EIF433O	Contexto, desarrollo y aplicación de software libre y de código abierto	3	EIF212 EIF211
EIF434O	Minería de datos I	3	EIF402 MAT006
EIF435O	Minería de datos II	3	EIF434O
EIF436O	Gráficos por computadora	3	EIF207 MAT005
EIG416O	Gestión de tecnología educativa	3	EIF425O
EIF439O	Uso estratégico del comercio electrónico		EIF209 EIF211
EIF440O	Seguridad informática	3	EIF208 EIF212
EIF442O	Entornos de programación para videojuegos	3	EIF204 MAT002
EIF443O	Diseño y programación de videojuegos en 3D	3	EIF201 MAT030
EIF441O	Soporte técnico avanzado	3	EIF202

Anexo 7 Inteligencia artificial, Bachillerato Ingeniería en Computación, TEC.

BLOQUE 7			
⊕ AE4208 - DESARROLLO DE EMPRENDEDORES			
Créditos	Horas	Requisitos	Correquisitos
4	4	NO HAY REQUISITOS EN EL PLAN	NO HAY CORREQUISITOS EN EL PLAN
⊕ IC5001 - ELECTIVA II			
Créditos	Horas	Requisitos	Correquisitos
3	4	NO HAY REQUISITOS EN EL PLAN	NO HAY CORREQUISITOS EN EL PLAN
⊕ IC6200 - INTELIGENCIA ARTIFICIAL			
Créditos	Horas	Requisitos	Correquisitos
4	4	IC5701, IC6400	NO HAY CORREQUISITOS EN EL PLAN
⊕ IC7602 - REDES			
Créditos	Horas	Requisitos	Correquisitos
4	12	IC6600	NO HAY CORREQUISITOS EN EL PLAN

Anexo 8 Mapa de algoritmos de machine learning



CRONOGRAMA

Año	2017			2018						
Actividad	Mayo	Junio	Julio a Mayo, 2018	Junio	Julio	Agosto	Setiembre	Octubre	Noviembre	Diciembre
Matrícula Tesis (TFG)										
Capítulo 1,2,3										
Asignación de tutores										
Revisión tutor										
Capítulo 4										
Desarrollo de la investigación.										
Revisión tutor										
Entrega y correcciones										
Capítulo 5										
Conclusiones y Recomendaciones										
Revisión, Tutor, Conclusiones y recomendaciones.										
Revisión Lectores										
Correcciones solicitadas										
Revisión final en conjunto tutor y lectores.										
Entrega de tesis final al comité evaluador.										
Defensa de tesis										