

**FACULTAD DE INGENIERÍA**

Escuela Académico Profesional de Ingeniería de Sistemas e Informática

Tesis

**Aplicación web para la clasificación de clientes de  
la Empresa CFCGROUP - Arequipa 2021**

Diego Esteban Santa Cruz Álvarez  
Renato Boris Zapana Llanquiche

Para optar el Título Profesional de  
Ingeniero de Sistemas e Informática

Arequipa, 2021

Repositorio Institucional Continental  
Tesis digital



Esta obra está bajo una Licencia "Creative Commons Atribución 4.0 Internacional" .

## **AGRADECIMIENTOS**

Agradecemos a Dios por bendecirnos, guiarnos a lo largo de toda nuestra vida y ser quien nos dio la fortaleza en los momentos difíciles, a nuestros padres, por ser las principales personas en confiar que lograríamos culminar el presente proyecto.

## **DEDICATORIA**

El Presente trabajo investigativo lo dedicamos principalmente a Dios, por ser quien nos dio sabiduría para poder realizarlo, a nuestros padres, por su amor, trabajo y el sacrificio que realizaron todos estos años para poder cumplir nuestros sueños de ser cada vez mejores.

## ÍNDICE DE CONTENIDOS

AGRADECIMIENTOS.....	ii
DEDICATORIA .....	iii
ÍNDICE DE CONTENIDOS.....	iv
ÍNDICE DE FIGURAS.....	vii
ÍNDICE DE TABLAS.....	x
RESUMEN .....	xii
ABSTRACT .....	xiii
INTRODUCCIÓN.....	xiv
Capítulo I. PLANTEAMIENTO DEL ESTUDIO .....	1
1.1 Planteamiento y formulación del problema .....	1
1.1.1 Formulación del problema .....	3
1.2 Objetivos.....	4
1.2.1 Objetivo General.....	4
1.2.2 Objetivos específicos.....	4
1.3 Justificación e importancia .....	4
1.3.1 Justificación .....	4
1.3.2 Importancia:.....	5
Capítulo II. MARCO TEÓRICO .....	6
2.1 Antecedentes del Problema .....	6
2.1.1 Tesis Internacionales.....	6
2.1.2 Tesis Nacionales .....	9
2.2 Bases Teóricas.....	13
2.2.1 Sistemas de Información: .....	13
2.2.2 Aplicación Web:.....	13
2.2.3 Inteligencia artificial: .....	16
2.3 Empresa .....	24
2.3.1 Datos generales de la institución .....	24
2.4 SCRUM: .....	26
2.4.1 Metodología aplicada para el desarrollo de la solución:.....	26
2.4.2 Roles de Scrum .....	26
2.4.3 Eventos de Scrum .....	28
2.4.4 Artefactos de Scrum .....	31
2.5 Definición de Términos Básicos.....	32

Capítulo III. METODOLOGÍA .....	34
3.1 Metodología aplicada para el desarrollo de la solución .....	34
Capítulo IV. ANÁLISIS Y DISEÑO DE LA SOLUCIÓN .....	37
4.1 Designación de Roles .....	37
4.2 Alcance General .....	37
4.3 Identificación de Historias de usuario.....	38
4.4 Validación de Pila de usuarios .....	40
4.5 Historias de Usuario .....	42
4.6 Planificación de Sprints.....	51
4.6.1 Historias de usuarios de los usuarios registrados .....	51
4.6.2 Historias de usuarios de usuario Administrador .....	52
4.7 Arquitectura de solución .....	54
4.8 Diseño de interfaces .....	55
4.9 Validación de diseño de interfaces.....	67
4.10 Diagrama de navegación .....	69
4.11 Diseño de base de datos .....	70
4.11.1 Diccionario de datos .....	71
4.11.2 Validación de base de datos básica.....	74
Capítulo V. CONSTRUCCIÓN.....	77
5.1 Construcción de la aplicación web .....	77
5.1.1 Pantallas de la Aplicación Web.....	77
5.1.2 Proceso de clasificación en la aplicación web.....	89
5.2 Construcción del modelo de clasificación.....	91
5.2.1 Procesamiento de datos .....	91
5.2.2 Selección del modelo de clasificación.....	91
5.2.3 Diseño del Modelo de la red neuronal.....	92
5.2.4 Construcción del modelo de la red neuronal .....	93
5.2.5 Resultados y Validación del Modelo .....	99
5.3 Pruebas y resultados .....	101
5.3.1 Pruebas de aceptación .....	102
5.3.2 Pruebas de exactitud .....	104
5.3.3 Resultados de exactitud.....	104
5.3.4 Pruebas de eficiencia .....	105
5.3.5 Resultados de eficiencia .....	105
5.3.6 Pruebas de fiabilidad .....	106

5.3.7 Resultados de fiabilidad.....	106
CONCLUSIONES.....	108
TRABAJOS FUTUROS .....	109
REFERENCIAS BIBLIOGRÁFICAS.....	110
ANEXOS .....	113

## ÍNDICE DE FIGURAS

Figura 1.	Incremento anual de los clientes sin clasificar .....	2
Figura 2.	Sub campos de la inteligencia artificial .....	16
Figura 3.	Representación gráfica del modelo de redes bayesianas. ....	19
Figura 4.	Representación gráfica del modelo KNN. ....	19
Figura 5.	Representación gráfica del modelo SVM. ....	20
Figura 6.	Representación gráfica del modelo de redes neuronales.....	20
Figura 7.	Representación gráfica del modelo de árboles de decisión. ....	22
Figura 8.	Fórmula para hallar la exactitud.....	22
Figura 9.	Fórmula para hallar la precisión.....	22
Figura 10.	Fórmula para hallar la sensibilidad. ....	23
Figura 11.	Matriz de confusión.....	24
Figura 12.	Organigrama empresa CFCGROUP.....	25
Figura 13.	Proceso de clasificación manual de clientes .....	26
Figura 14.	Arquitectura de la aplicación web .....	54
Figura 15.	Interfaz Historia de usuario 1 .....	55
Figura 16.	Interfaz Historia de usuario 2 .....	56
Figura 17.	Interfaz Historia de usuario 3 .....	57
Figura 18.	Interfaz Historia de usuario 4 .....	58
Figura 19.	Interfaz Historia de usuario 5 .....	58
Figura 20.	Interfaz Historia de usuario 6 .....	59
Figura 21.	Interfaz Historia de usuario 7 .....	59
Figura 22.	Interfaz Historia de usuario 8 .....	60
Figura 23.	Interfaz Historia de usuario 9 .....	60
Figura 24.	Interfaz Historia de usuario 10 .....	61
Figura 25.	Interfaz Historia de usuario 11 .....	62
Figura 26.	Interfaz Historia de usuario 12 .....	63
Figura 27.	Interfaz Historia de usuario 13 .....	64
Figura 28.	Interfaz Historia de usuario 14 .....	64
Figura 29.	Interfaz Historia de usuario 15 .....	65
Figura 30.	Interfaz Historia de usuario 16 .....	65

Figura 31.	Interfaz Historia de usuario 17 .....	66
Figura 32.	Interfaz Historia de usuario 18 .....	67
Figura 33.	Diagrama de navegación.....	70
Figura 34.	Diagrama de base de datos.....	71
Figura 35.	Interfaz de aplicación web Historia de usuario 1 .....	77
Figura 36.	Interfaz de aplicación web Historia de usuario 2 .....	78
Figura 37.	Interfaz de aplicación web Historia de usuario 3 .....	79
Figura 38.	Interfaz de aplicación web Historia de usuario 4 .....	80
Figura 39.	Interfaz de aplicación web Historia de usuario 5 .....	80
Figura 40.	Interfaz de aplicación web Historia de usuario 6 .....	81
Figura 41.	Interfaz de aplicación web Historia de usuario 7 .....	81
Figura 42.	Interfaz de aplicación web Historia de usuario 8 .....	82
Figura 43.	Interfaz de aplicación web Historia de usuario 9 .....	82
Figura 44.	Interfaz de aplicación web Historia de usuario 10 .....	83
Figura 45.	Interfaz de aplicación web Historia de usuario 11 .....	84
Figura 46.	Interfaz de aplicación web Historia de usuario 12 .....	85
Figura 47.	Interfaz de aplicación web Historia de usuario 13 .....	86
Figura 48.	Interfaz de aplicación web Historia de usuario 14 .....	86
Figura 49.	Interfaz de aplicación web Historia de usuario 15 .....	87
Figura 50.	Interfaz de aplicación web Historia de usuario 16 .....	87
Figura 51.	Interfaz de aplicación web Historia de usuario 17 .....	88
Figura 52.	Interfaz de aplicación web Historia de usuario 18 .....	88
Figura 53.	Registro de clientes a la base de datos .....	89
Figura 54.	Obtención de los registros de la base de datos .....	89
Figura 55.	Rutas web para la obtención de datos .....	90
Figura 56.	Obtención de datos de clientes.....	90
Figura 57.	Selección de variables para el modelo .....	90
Figura 58.	Evaluación de los algoritmos de clasificación .....	91
Figura 59.	Diseño del modelo de clasificación .....	93
Figura 60.	Parámetros principales .....	93
Figura 61.	Creación del modelo.....	94
Figura 62.	Definición de capa oculta y función de activación .....	94

Figura 63.	Definición de la capa de salida .....	94
Figura 64.	Definición de la función de pérdida .....	95
Figura 65.	Parámetros de entrenamiento.....	95
Figura 66.	Época 1 al 7 de modelo de clasificación .....	96
Figura 67.	Época 40 al 50 de modelo de clasificación .....	96
Figura 68.	Época 100 al 110 de modelo de clasificación .....	97
Figura 69.	Época 140 al 150 de modelo de clasificación .....	97
Figura 70.	Época 190 al 200 de modelo de clasificación .....	98
Figura 71.	Época 240 al 250 de modelo de clasificación .....	98
Figura 72.	Accuracy y Validation Accuracy .....	99
Figura 73.	Lost y Validation Loss.....	99
Figura 74.	Matriz de confusión.....	100
Figura 75.	Clasificación obtenida en la aplicación web .....	101
Figura 76.	Resultados de prueba de aceptación.....	103
Figura 77.	Prueba de Kolmogorov-Smirnov para exactitud.....	104
Figura 78.	Prueba U de Mann Whitney para exactitud.....	105
Figura 79.	Prueba de Kolmogorov-Smirnov para eficiencia .....	105
Figura 80.	Prueba U de Mann Whitney para eficiencia .....	106
Figura 81.	Prueba de Kolmogorov-Smirnov para fiabilidad .....	107
Figura 82.	Prueba U de Mann Whitney para fiabilidad .....	107

## ÍNDICE DE TABLAS

Tabla 1.	Incremento anual de clientes sin clasificar 2015-2020.....	2
Tabla 2.	Tiempo empleado con el proceso de clasificación actual .....	3
Tabla 3.	Tipos de sistemas de información .....	13
Tabla 4.	Cuadro Comparativo Lenguaje PHP .....	15
Tabla 5.	Cuadro comparativo Aprendizaje supervisado y no supervisado.....	17
Tabla 6.	Cuadro comparativo Pytorch y Tensorflow .....	18
Tabla 7.	Fases de desarrollo.....	36
Tabla 8.	Asignación de roles .....	37
Tabla 9.	Product backlog .....	39
Tabla 10.	Validación de Historias de usuario.....	41
Tabla 11.	Historia de usuario 1 .....	42
Tabla 12.	Historia de usuario 2.....	43
Tabla 13.	Historia de usuario 3.....	43
Tabla 14.	Historia de usuario 4.....	44
Tabla 15.	Historia de usuario 5.....	44
Tabla 16.	Historia de usuario 6.....	45
Tabla 17.	Historia de usuario 7.....	45
Tabla 18.	Historia de usuario 8.....	46
Tabla 19.	Historia de usuario 9.....	46
Tabla 20.	Historia de usuario 10.....	47
Tabla 21.	Historia de usuario 11 .....	47
Tabla 22.	Historia de usuario 12.....	48
Tabla 23.	Historia de usuario 13.....	49
Tabla 24.	Historia de usuario 14.....	49
Tabla 25.	Historia de usuario 15.....	50
Tabla 26.	Historia de usuario 16.....	50
Tabla 27.	Historia de usuario 17 .....	51
Tabla 28.	Historia de usuario 18.....	51
Tabla 29.	Planificación de Historias de Usuario de los usuarios registrados.....	52
Tabla 30.	Planificación de Historias de Usuario del usuario administrador .....	53
Tabla 31.	Leyenda de prioridad.....	53

Tabla 32.	Diccionario de datos .....	69
Tabla 33.	Diccionario de datos .....	74
Tabla 34.	Validación de la base de datos .....	76
Tabla 35.	Selección del número de neuronas en la capa oculta .....	92
Tabla 36.	Métricas Precision, Recall , F1 Score y Exactitud .....	101
Tabla 37.	Pruebas de usabilidad .....	102
Tabla 38.	Pruebas de funcionalidad .....	103
Tabla 39.	Resultados de encuestas.....	103

## RESUMEN

La presente tesis plantea la creación de una aplicación web para la clasificación de los clientes de la empresa CFCGROUP. En promedio se tiene 8000 clientes con baja exactitud en la clasificación, inconsistencias en los datos y, debido a la forma manual, uso de tiempo excesivo para clasificarlos; motivos que generan que los clientes se desvinculen de la empresa. Por esta razón, se plantearon los objetivos: aumentar la exactitud de la clasificación, reducir los tiempos e inconsistencias en los datos con respecto al proceso manual con que se desarrollaba. Ante esto, se aplicó una metodología de desarrollo SCRUM para el proceso de construcción y desarrollo de la solución. Se desarrolló un Sistema web utilizando el framework Laravel y base de datos Mysql, se usó la biblioteca Tensorflow para el modelo de clasificación, mediante el cual se implementó una red neuronal artificial de tipo secuencial (Perceptrón Multicapa).

Se logró el objetivo de mejorar la exactitud de la clasificación de clientes, que pasó de 60.75% a 90.50%, además se obtuvo una reducción del tiempo de clasificación, que en promedio pasó de 7.5052 a 0.0040 segundos; también se redujeron las inconsistencias pasando de 5.063% a 0%. Para la validación de los resultados se aplicó la prueba de U de Mann Whitney con significancia asintótica (p valor) menor a 0.05, por lo que al haberse demostrado el logro de los objetivos específicos se asume el logro del objetivo general referente a la mejora de la clasificación de clientes.

**Palabras clave:** Aplicación Web, Modelo de Clasificación, Aprendizaje Automático, Inteligencia Artificial, Red Neuronal, Tensorflow, Framework, Scrum.

## ABSTRACT

This thesis proposes the creation of a web application for the classification of clients of the company CFCGROUP. On average, there are 8,000 clients with low accuracy in the classification, inconsistencies in the data and, due to the manual form, excessive time used to classify them; Reasons that cause customers to leave the company. For this reason, the objectives were set: increase the accuracy of the classification, reduce time and inconsistencies in the data with respect to the manual process with which it was developed. Given this, a SCRUM development methodology was applied for the process of construction and development of the solution. A web system was developed using the Laravel framework and Mysql database, the Tensorflow library was used for the classification model, through which a sequential type artificial neural network (Multilayer Perceptron) was implemented.

The objective of improving the accuracy of customer classification was achieved, which went from 60.75% to 90.50%, in addition, a reduction in classification time was obtained, which on average went from 7.5052 to 0.0040 seconds; inconsistencies were also reduced from 5.063% to 0%. For the validation of the results, the Mann Whitney U test was applied with asymptotic significance (p value) less than 0.05, so that, having demonstrated the achievement of the specific objectives, the achievement of the general objective referring to the improvement of customer classification.

**KeyBoards:** Web Application, Classification model, Machine Learning, Artificial intelligence, Neural Network, Tensorflow, Framework, Scrum.

## INTRODUCCIÓN

La razón que motivó la realización de esta tesis fue que se evidenciaron los resultados poco eficientes del proceso manual de clasificación, y por consecuencia, el hecho de que podían ser mejorados mediante el uso de un sistema web y la creación de un modelo de Aprendizaje Automático que ayudaría a mejorar la eficiencia en la clasificación.

Un problema que se viene presentando en la empresa CFCGROUP en los últimos años es el incremento de datos de clientes sin clasificar, para ello la empresa emplea un proceso de clasificación manual poco eficiente, dificultando el contacto y envío de información acorde con los intereses de dichos clientes.

Para la elaboración de esta tesis se utilizó la metodología ágil Scrum, que es un marco de trabajo para el desarrollo de productos de software. Se basa en prácticas iterativas e incrementales que incrementan significativamente la productividad y reducen el tiempo para la obtención de resultados. Para la construcción del modelo de clasificación se usó la biblioteca Tensorflow, mediante la cual se implementó una red neuronal artificial de tipo secuencial.

La solución de este problema es importante, ya que de no implementarse, se incrementaría la desvinculación del contacto con la empresa por parte de los clientes, debido a que gran parte de ellos no son clasificados correctamente.

El problema anteriormente descrito llevó a formular la siguiente interrogante:

¿Cómo mejorar la clasificación de clientes para la empresa CFCGROUP Arequipa - 2021?

Por lo cual este estudio se estructura en cinco capítulos:

Capítulo I: Planteamiento del Estudio, donde se explica la situación problemática del proceso de clasificación de clientes y la descripción del incremento de clientes sin clasificación anual.

Capítulo II: Marco teórico, que analiza el estado del arte relacionado a sistemas web, Inteligencia Artificial, Aprendizaje Automático; y modelos de clasificación que detallan puntos relevantes para mejorar el entendimiento de sistemas web y clasificación mediante el uso del Aprendizaje Automático.

Capítulo III: Metodología, donde se detalla las características de las fases de la metodología Scrum aplicada al desarrollo del proyecto.

Capítulo IV: Análisis y diseño de la solución, donde se identifican los requerimientos y el diseño de interfaces; incluyendo la base de datos a utilizar y la definición de la arquitectura de la solución.

Capítulo V: Construcción, donde se realiza el modelado y la construcción del modelo de clasificación y aplicación web, incluyendo el cálculo del porcentaje de exactitud del modelo y su validación.

Finalmente, en las conclusiones se explica el logro de cada uno de los objetivos.

## **Capítulo I. PLANTEAMIENTO DEL ESTUDIO**

### **1.1 Planteamiento y formulación del problema**

La empresa CFCGROUP realiza capacitaciones a trabajadores de diversos rubros, los cuales son agrupados en el rubro de construcción, minería y gestión. La empresa contacta con sus clientes mediante llamadas, campañas de marketing, correos electrónicos y otros medios.

Actualmente la clasificación se realiza mediante un proceso manual, los resultados de la clasificación no son lo suficientemente aceptables por el tiempo empleado y la exactitud lograda, que en promedio es del 60%, limitando al área de ventas a tener mayor probabilidad de realizar una venta.

El área encargada de ventas de la empresa cuenta con una gran base de datos de clientes, los cuales son contactados mediante diversos medios para poder concretar una venta. Los datos de los clientes que registra la empresa se obtienen por diversas fuentes: de manera directa con las empresas, de manera indirecta con empresas intermediadoras, y con el público en general. De este conjunto de datos se tienen registros ya clasificados por el grupo de interés al que pertenecen, pero también se registran datos sin categoría asignada, los cuales son clasificados de forma manual.

En esta tesis se planea construir un software de clasificación de clientes potenciales utilizando Aprendizaje Automático, con el fin de mejorar la clasificación de los clientes y poder proporcionarles información acorde a sus intereses sobre servicios de capacitación que ofrece la empresa, basándose en datos históricos de personas que ya fueron capacitados por esta.

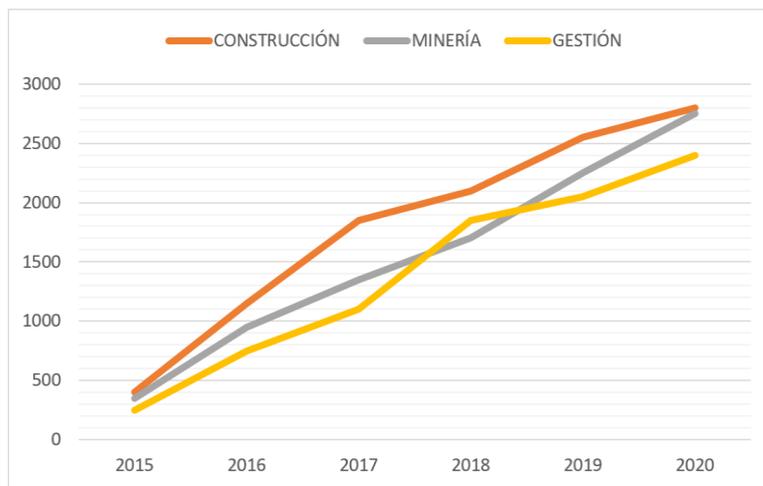
En la tabla 1, se resume el incremento anual de los datos de clientes sin clasificar desde el año 2015 al 2020.

AÑO	CANTIDAD DE CLIENTES TOTAL POR CATEGORÍA		
	CONSTRUCCIÓN	MINERÍA	GESTIÓN
2015	400	350	250
2016	1150	950	750
2017	1700	1350	1400
2018	2100	1900	1850
2019	2550	2250	2050
2020	2800	2750	2400

**Tabla 1. Incremento anual de clientes sin clasificar 2015-2020**

**Fuente: Elaboración propia**

Convirtiendo la tabla 1 se aprecia el crecimiento de los datos sin clasificar anualmente, como se muestra en la figura 1.



**Figura 1. Incremento anual de los clientes sin clasificar**

**Fuente: Elaboración propia**

En la tabla 2, se muestra el tiempo empleado en minutos con el proceso actual en los últimos 1000 registros clasificados.

Cantidad de registros	Tiempo en minutos
200	28.3
400	56.7
600	85
800	113.3
1000	141.7

**Tabla 2. Tiempo empleado con el proceso de clasificación actual**

**Fuente: Elaboración propia**

Se identificó que los datos de los clientes de la empresa CFCGROUP no están clasificados con una exactitud aceptable, lo cual trae problemas con el proceso de segmentación en las tareas de marketing que la empresa realiza. La empresa desea ofrecer sus servicios de capacitación a clientes potenciales que realmente estén interesados en un determinado programa de capacitación, es por ello que requiere de algún método de clasificación que permita dirigir el proceso de venta a un segmento de los clientes con mayor probabilidad de adquirirlos y, así, evitar el envío de información a clientes que puedan estar interesados en otro programa de capacitación ofrecido por la empresa.

Esto trae como consecuencia que clientes que no han sido clasificados correctamente sean contactados con información que no es de su interés, lo que llevaría a que el contacto con los clientes en su gran mayoría no llegue a concretarse de manera satisfactoria; lo cual ocasiona que estos, al no tener interés en los servicios ofertados, rompan contacto con la empresa de diversas formas, o recurran a la competencia en busca de un servicio de su interés, y, de esta manera, causar la reducción de la participación en el mercado de la empresa.

### **1.1.1 Formulación del problema**

#### **1.1.1.1 Problema General**

¿Cómo mejorar la clasificación de clientes para la empresa CFCGROUP Arequipa - 2021?

#### **1.1.1.2 Problemas Específicos**

¿Cómo mejorar la exactitud de la clasificación de clientes para la empresa CFCGROUP - Arequipa 2021?

¿Cómo reducir el tiempo de clasificación de clientes para la Empresa CFCGROUP - Arequipa 2021?

¿Cómo reducir las inconsistencias en los datos de los clientes clasificados para la empresa CFCGROUP - Arequipa 2021?

## **1.2 Objetivos**

### **1.2.1 Objetivo General**

Mejorar la clasificación de clientes para la empresa CFCGROUP - Arequipa – 2021 mediante una aplicación Web.

### **1.2.2 Objetivos específicos**

- Mejorar la exactitud de la clasificación de clientes para la empresa CFCGROUP - Arequipa 2021.
- Reducir el tiempo de clasificación de clientes para la empresa CFCGROUP - Arequipa 2021.
- Reducir las inconsistencias en los datos de los clientes para la empresa CFCGROUP - Arequipa 2021.

## **1.3 Justificación e importancia**

### **1.3.1 Justificación**

#### **1.3.1.1 Justificación Social:**

Esta solución es socialmente necesaria ya que al implementarse la solución planteada en esta tesis, la empresa podrá mejorar la eficacia con la que se contacta a sus clientes, utilizando técnicas de mercadeo mejor enfocadas, ofreciendo sus servicios, más acordes, con los intereses de sus clientes; y reduciendo el envío de información irrelevante a las preferencias de los clientes. La solución será de suma importancia para la oferta de servicios idóneos para los clientes, que podrán recibir información acorde con sus intereses.

### **1.3.1.2 Justificación Práctica:**

El aprovechamiento de las tecnologías de Aprendizaje Automático para la clasificación de datos, permitirá reducir la incertidumbre y el costo que está ligado a la toma de decisiones para la gestión de los clientes de la empresa. Utilizando los registros históricos que posee la empresa se podrá hacer uso del Aprendizaje Automático para categorizar a los clientes en base a los servicios que ofrece la empresa; de igual manera, mediante el empleo de tecnologías web para la construcción de la interfaz y lógica se podrá poner a disposición de manera remota la información resultante de este proceso.

### **1.3.2 Importancia:**

Esta tesis es importante porque de no implementarse se incrementaría la desvinculación del contacto con la empresa por parte de los clientes, puesto que, al no recibir información acorde con sus preferencias o intereses, los clientes optarían en mayor cantidad por desafiliarse de los diferentes medios de contacto que tiene con la empresa (correo electrónico, telefonía, mensajes y demás); al no ofrecerles los servicios idóneos los clientes potenciales perderán el interés, asimismo, podrían ser contactados por la competencia, de manera que reduciría la participación en el mercado de la empresa.

## **Capítulo II. MARCO TEÓRICO**

### **2.1 Antecedentes del Problema**

#### **2.1.1 Tesis Internacionales**

En la tesis titulada “Modelo de Machine Learning para la clasificación de estudiantes de acuerdo a su rendimiento académico en el centro de idiomas de la Universidad Nacional del Santa” elaborada por (1), se tuvo como objetivo principal, aumentar la efectividad del proceso de clasificación de los alumnos de un centro de idiomas; teniendo como problema el alto porcentaje de alumnos, en su mayoría nuevos provenientes de otras instituciones, que desaprobaban los cursos de idiomas. Aplicando Inteligencia Artificial se construyó un modelo de Aprendizaje Automático el cual utilizó la técnica de regresión logística, con el cual se logró una exactitud de 88.16% en la clasificación de estudiantes de acuerdo a su rendimiento académico, logrando que el número de clasificaciones acertadas se incrementara en un 82.08%. Del estudio en mención se destaca como aporte a la presente tesis: el empleo del algoritmo de regresión logística como una técnica de Aprendizaje Automático usado para la clasificación.

En la tesis titulada “Machine Learning para la estimación del riesgo de crédito en una cartera de consumo” elaborada por (2), se tuvo como objetivo principal comparar la precisión del modelo de regresión logística frente a otros modelos de Aprendizaje Automático, teniendo como problema la estimación del riesgo de crédito en una cartera de consumo. Para realizar la comparación se utilizaron los modelos de Bosque Aleatorio, Máquinas de vectores de soporte y Perceptrón Multicapa; se comparó la eficiencia para calcular la estimación de los clientes que van a entrar en mora;

obteniéndose como resultado un 83% de casos de mora, acertados con el algoritmo regresión logística, en comparación a los demás modelos mencionados anteriormente que obtuvieron menores resultados. Del estudio en mención se destaca como aporte a la presente tesis: el análisis y la comparación realizada a diferentes técnicas de clasificación para optar por el más exacto.

En la tesis titulada “Clasificación de clientes de la industria bancaria por métodos estadísticos y redes neuronales artificiales” elaborada por (3), se tuvo como objetivo abordar la problemática de la clasificación de clientes que pertenecen al sector bancario haciendo uso de tres métodos distintos de clasificación supervisada; teniendo como problema el clasificar un nuevo cliente dentro de una de las poblaciones determinadas, basándose en los valores de las variables aleatorias. Las técnicas estadísticas utilizadas fueron la regresión logística binaria y el análisis discriminante lineal, además, se utilizó un método de Aprendizaje Automático, que son las redes neuronales artificiales. La investigación tiene como resultado que la RNA necesita menos datos para el entrenamiento y converge más rápido hacia una exactitud aceptable, alcanzando una precisión de 90.71% con una muestra de entrenamiento de 30140 registros de clientes. Del estudio en mención se destaca como aporte a la presente tesis: el uso de redes neuronales artificiales como alternativa con mejores resultados, utilizando menos datos de entrenamiento.

El artículo titulado “Machine Learning en la detección de fraudes de comercio electrónico aplicado a los servicios bancarios” elaborado por (4), tiene el objetivo de elaborar un modelo para la identificación de fraudes en transacciones financieras; teniendo como problemática la cantidad limitada de transacciones fraudulentas existentes, que son el 0.1% de las transacciones realizadas por las instituciones bancarias, y el hecho de que procesar demasiadas transacciones legítimas podrían sesgar el resultado. Para esto se desarrolló un modelo haciendo uso de un conjunto de datos de prueba, que permitió predecir correctamente la mayoría de casos de transacciones fraudulentas; usando una muestra de 80.792 transacciones y utilizando el modelo de Random Forest se consiguió el nivel de exactitud de 96.14%. Del artículo mencionado se destaca como aporte a la presente tesis: el uso del algoritmo Random Forest como modelo para clasificación binaria.

En la tesis titulada “Modelo de elaboración de pronóstico de ventas mediante el uso de redes neuronales artificiales y SVR” elaborada por (5), se tuvo como objetivo pronosticar o prever el nivel de ventas para la supervivencia de una empresa; teniendo como

problema la incertidumbre que puede generar la demanda de productos y servicios en las empresas; se comparó el desempeño de los modelos tradicionales frente a sistemas mejor desarrollados, como son las redes neuronales artificiales y las máquinas de soporte vectorial o regresión de soporte vectorial. El análisis presentado en este estudio resalta la eficacia de utilizar redes neuronales artificiales para el pronóstico de series temporales, considerando tanto los factores externos como los internos de la organización. Del estudio en mención se destaca como aporte a la presente tesis: comparar el desempeño de modelos tradicionales frente a modelos más desarrollados como las redes neuronales.

En la tesis titulada “Modelo Machine Learning para la Clasificación de pacientes en términos del nivel asistencial requerido en una urgencia pediátrica con Área de Cuidados Mínimos” elaborada por (6), se tiene como objetivo construir un modelo de Aprendizaje Automático para la clasificación de pacientes que requieran urgencia pediátrica; se presenta como problema la atención en los departamentos de urgencias en diferentes clínicas, la demora en el tiempo de atención, la insatisfacción de los usuarios del servicio de salud y la ineficiente asignación de recursos. Se desarrolló una herramienta de Aprendizaje Automático para mejorar la toma de decisiones al momento de referir un paciente a su cita, consiguiendo como resultado una herramienta de soporte para los médicos, y, además, logrando reducir el tiempo de espera para recibir atención, mediante la clasificación de las citas según la especialidad a la que va referido. Del estudio en mención se destaca como aporte a la presente tesis: que el empleo de una herramienta de Aprendizaje Automático mejora la toma de decisiones.

La tesis titulada “Técnicas de Ingeniería informática e inteligencia artificial para clasificación: aplicaciones para el descubrimiento de fármacos y dianas moleculares” elaborada por (7), tiene como objetivo la implementación de una nueva herramienta informática para descubrir dianas moleculares y fármacos; teniendo como problema la rapidez con la que los fármacos actuales cambian en el entorno. Haciendo uso de técnicas de inteligencia artificial e ingeniería informática se construye una web de acceso libre para los científicos, y de esta manera, conseguir que los datos se publiquen en revistas internacionales online; logrando así el cálculo de los promedios e índices en las redes moleculares de fármacos y proteínas. Del estudio en mención se destaca como aporte a la presente tesis: el uso de una aplicación web para la publicación de los resultados, y de esa forma, la obtención de información en tiempo real y con disponibilidad inmediata.

La tesis titulada “Modelo de Machine Learning en la detección de sitios web Phishing” elaborada por (8) , tiene como objetivo clasificar los sitios web Phishing mediante algoritmos de Machine Learning; teniendo como problemática los limitados sistemas de detección de Phishing en los sitios web. Para lograr la clasificación se usó información de 11055 sitios web correctos y 2211 sitios web usados para Phishing, como entrenamiento del modelo de clasificación; logrando mejorar el rendimiento de la identificación correcta de los sitios web en un 4.40% y una mejora en el rendimiento global del 3.62%; obteniendo una precisión al clasificar adecuadamente los sitios web del 97.42%. Del estudio en mención se destaca como aporte a la presente tesis: el uso de algoritmos de Aprendizaje Automático para mejorar el rendimiento y la precisión en la clasificación de datos.

### **2.1.2 Tesis Nacionales**

En la tesis titulada “Clasificación de las prioridad de atención a Reclamos Presentados por clientes utilizando Machine Learning” elaborada por (9), se tiene como objetivo la evaluación del árbol de decisión- técnica de clasificación utilizado en la inteligencia artificial- cuya característica principal es su aporte visual a la toma de decisiones; teniendo como problemática la fuga de clientes ocasionada por la alta competencia en el sector de telecomunicaciones. La tesis busca establecer la idoneidad del árbol de decisión, el cual se utiliza como una herramienta de soporte para predecir la prioridad del reclamo, y de esta manera determinar si puede ser una herramienta de soporte para la predicción de la prioridad del reclamo; se consiguió una precisión de 89.79% utilizando la técnica de árbol de decisión. Del estudio en mención se destaca como aporte a la presente tesis: la evaluación del modelo de árbol de decisión para determinar su uso como herramienta válida para la toma de decisiones.

En la tesis titulada “Sistema Predictivo Basado en un modelo Credit Scoring de aprendizaje Automático para la Medición del Riesgo Crediticio en los Créditos PYME de la EDPYME alternativa S.A “ elaborada por (10), se tiene como problemática las deficientes herramientas tecnológicas y los métodos de referencia de evaluación de créditos a pequeñas y micro empresas, que determinan la medición del riesgo crediticio; se tiene como objetivo elaborar un modelo de Machine Learning para mejorar la medición del riesgo crediticio de los créditos pyme; se plantea, también, implementar un sistema predictivo inteligente que está basado en el puntaje crediticio. La evaluación del modelo con datos reales permitió comparar las distintas técnicas de Aprendizaje

Automático haciendo uso de algoritmos que soporten los datos para la fase de clasificación, consiguiendo resultados que indican una superioridad del algoritmo de bosques de decisión frente al algoritmo de red neural en un 0.86%. Del estudio en mención se destaca como aporte a la presente tesis: la comparación de diferentes técnicas de Aprendizaje Automático para la elección de la más óptima.

En la tesis titulada “Comparación de modelos de clasificación, regresión logística y árboles de clasificación para evaluar el rendimiento académico” elaborada por (11), se tiene como objetivo la comparación de dos técnicas de clasificación de Aprendizaje Automático, la Regresión Logística Binaria y Árboles de clasificación para evaluar el rendimiento académico; presentando como problemática el estado de éxito o fracaso como etiquetas que presentan muchas instituciones educativas para medir el desempeño de los estudiantes. Por lo cual se desarrollaron modelos que fueron medidos por cuatro indicadores: Curva ROC, Sensibilidad, Índice de GINI e Índice de Kappa en función del poder de clasificación y predicción de los modelos sobre el rendimiento académico. Evidenciando que los Árboles de clasificación consiguen mejores resultados sobre la clasificación y predicción. Para el análisis se usó datos de estudiantes universitarios del primer semestre. Según la evaluación de la clasificación de los modelos comparados se eligió la Técnica de Árboles de clasificación, siendo la más óptima y logrando resultados en Sensibilidad de 77,6%, AUC de 90,1%, Gini de 80,2% y Kappa de 0,589. Del estudio en mención se destaca como aporte a la presente tesis: la comparación entre las técnicas de Árboles de clasificación y Regresión Logística para la predicción y clasificación.

La tesis titulada “Análisis de las características que identifican a un usuario de practis premium; variables que deciden para convertirse de una cuenta freemium a premium” elaborada por (12), tiene como objetivo identificar las características que condicionan a los usuarios a migrar de una cuenta Freemium a una cuenta Premium de un software contable. El problema radica en la necesidad de identificar los motivos por los cuales los clientes no llegan a llevar a cabo la conversión de las cuentas Freemium a Premium en Perú; utilizando el modelo de árbol de decisión se logró identificar variables relevantes que fueron 15 y 1348 registros. Del estudio en mención se destaca como aporte a la presente tesis: la identificación de variables que influyen en el modelo presentado; y la utilización del modelo de árbol de decisión como técnica de Aprendizaje Automático, que permite la predicción del comportamiento de los usuarios.

La tesis titulada “Sistema de Gestión y clasificación automática de denuncias ambientales mediante aprendizaje máquina” elaborada por (13), tiene como objetivo reducir el tiempo empleado en la atención de denuncias ambientales, a través del desarrollo de un sistema de gestión y clasificación automática de las denuncias con el uso de algoritmos de Aprendizaje Automático; teniendo como problemática la información insuficiente, contradictora y mal clasificada de denuncias realizadas por los ciudadanos. Para esto se desarrolló un módulo de clasificación de denuncias ambientales, se construyó un aplicativo progresivo web, y se desarrolló un sistema de gestión de denuncias ambientales. Al implementarse el modelo web para los usuarios finales, se consiguió una precisión promedio del 85%. Del estudio en mención se destaca como aporte a la presente tesis: la implementación de un sistema web en conjunto con un modelo de Aprendizaje Automático donde se pueden visualizar los resultados obtenidos por el modelo.

La tesis titulada “Optimización del Proceso de gestión de flota para una empresa de transporte de carga por carretera usando Machine Learning, BI, GPS y SMS Gateway” elaborada por (14), tiene como objetivo plantear una solución de mejora al proceso de gestión de flota de una empresa que realiza transporte de carga de mercancías, en general, por carretera; teniendo como problema la falta de conocimiento del proceso de gestión de flota, ya que muchas de las actividades son realizadas de forma manual. Para la solución empresarial se utilizó el framework Zachman, que facilitó realizar el análisis del negocio, estructura y procesos, bajo el criterio del negocio como caja negra que necesita interpretarse. Esto permitió que se comprenda el propósito y los objetivos para los que fue creado el negocio, pues es allí donde se gestionan los recursos y donde se evidencia la problemática. Se plantea una solución mediante el uso de inteligencia artificial con el algoritmo de backpropagation, que optimiza el proceso de gestión de flota de la empresa de transporte. Del estudio en mención se destaca como aporte a la presente tesis: el uso del algoritmo del backpropagation en la optimización del proceso de gestión de flota.

La tesis titulada “Desarrollo de un sistema de proyección de costos y costeo unitario de importaciones con métodos predictivos basados en Machine Learning” elaborada por (15), tiene como objetivo la construcción, utilizando un modelo de Aprendizaje Automático, de un sistema de proyección de costos de importación con métodos predictivos; tiene como problemática el proceso manual realizado para evaluar las cotizaciones de los proveedores. Se plantea el desarrollo de una herramienta que hace

uso de la información publicada por la Sunat sobre las importaciones históricas para la predicción de seguros y costos, y de la información del mismo procedimiento para proporcionar datos de los costos de los productos importados colocados en almacén; suprimiendo así, los riesgos del factor humano al evitar transcribir la información; se obtiene como resultado una exactitud de al menos 90% del modelo a partir de la información pública derivada de la Sunat. Del estudio en mención de destaca como aporte a la presente tesis: el uso de datos de entidades públicas como datos históricos para ser utilizados en modelos de Aprendizaje Automático.

La tesis titulada “Mejora del Proceso de Toma de decisiones en las Ventas de Abarrotes de una Empresa de Chiclayo, a partir de un Sistema Informático basado en Herramientas OLAP” elaborado por (16), tiene como objetivo mejorar el proceso de toma de decisiones en las ventas, utilizando para ello un sistema informático web basado en herramientas OLAP; la problemática es la inconsistencia en los datos. Para ello se realizó un análisis de la información, el cual dio como resultado una gran cantidad de datos incompletos, los cuales requerían un estimado de 24 horas para realizar el conteo; mediante la implementación del sistema se logró como resultado la reducción en el tiempo hasta 14 minutos, también se redujo la inconsistencia en la información, la cual está disponible mediante el sistema web de manera inmediata; logrando que los responsables puedan tomar decisiones más acertadas en un menor tiempo y con información consistente. Del estudio en mención se destaca para la presente tesis: el uso de un sistema informático web para la reducción de inconsistencias en los datos, mediante una evaluación y registro de los datos completos evitando la pérdida de información relevante.

## 2.2 Bases Teóricas

### 2.2.1 Sistemas de Información:

Los sistemas de información son un conjunto de elementos interrelacionados que trabajan en grupo, interactuando entre sí con un objetivo en común; ayudan a procesar, gestionar y distribuir la información importante para procesos fundamentales. (17)

En la tabla 3, se describen los tipos de sistemas de información más importantes.

Tipo de Sistema	Descripción
<b>Sistemas de automatización de oficinas</b>	Es una red de distintas herramientas y personas necesarias para efectuar tareas administrativas.
<b>Sistemas de gestión del conocimiento</b>	Almacena y sustrae información que proporciona ayuda a los usuarios para optimizar los esfuerzos de colaboración y completar las tareas.
<b>Sistemas de información de gestión</b>	Utiliza diversos datos de transacciones para dar soporte a la administración, de esa manera optimizar la toma de decisiones.
<b>Sistemas de apoyo a la decisión</b>	Procesa información para ayudar en la toma de decisiones, recopila los datos necesarios para que la administración tome las medidas necesarias en el momento oportuno.
<b>Sistema de Soporte</b>	Ayuda a encontrar respuestas a preguntas poco comunes, para que se puedan tomar decisiones que ayuden a mejorar la perspectiva y el rendimiento.

**Tabla 3. Tipos de sistemas de información**

**Fuente: (18)**

### 2.2.2 Aplicación Web:

Una aplicación Web es un sitio Web que agrupa páginas con contenido; dependiendo la solicitud se aloja en un servidor web que puede ser accedido mediante el uso de un navegador web. El contenido de una página se determina cuando el usuario realiza la solicitud de una página del servidor Web. Dado que el contenido final de la página varía de una petición a otra en función de las acciones del usuario, este tipo de páginas se denominan página dinámica. (19)

### 2.2.2.1 Fundamentos de App web:

- **Http:** Es un protocolo el cual nos permite ejecutar una petición de información a los servidores web, los resultados suelen ser documentos HTML. Http es la base de cualquier intercambio de datos y recursos en la Web y, a su vez, un protocolo con la estructura cliente-servidor, esto quiere decir que una petición de datos inicia por el elemento que recibirá los datos resultantes (el cliente), usualmente un navegador Web o un dispositivo móvil. Así, una página web completa resulta de la unión de distintos subdocumentos recibidos, por ejemplo, un documento que determina el estilo de maquetación de la página web (CSS), el texto, las imágenes, vídeos, etc. (20).
- **HTML:** El lenguaje de marcado de hipertexto (HTML) es un lenguaje de marcado que forma parte de muchas de las páginas web que definen la estructura del contenido web. Un hipertexto hace referencia a la conexión de páginas web a través de enlaces. HTML hace uso de marcas para etiquetar el contenido que se mostrará en el navegador web. (21).
- **CSS:** (hojas de estilo en cascada) es un lenguaje que determina la apariencia de una página escrita en un lenguaje de marcado, separando el formato y el contenido. CSS determina cómo serán presentados los elementos de la página web. (22).

### 2.2.2.2 Lenguajes de Programación Web:

Es un código interpretado por un servidor web, usado principalmente para la creación de sitios web dinámicos que permiten crear aplicaciones de tipo cliente-servidor, en la cual el usuario realiza la petición y el servidor da respuesta. Ejemplos de lenguajes de programación web: PHP, Python, JavaScript, etc. (23).

- **JavaScript:** Es un lenguaje de programación que es utilizado para la creación de páginas web dinámicas. Una página web dinámica es la que integra efectos en los textos, animaciones, acciones que se activan al identificar un evento y ventanas con mensajes de aviso para el usuario. JavaScript es un lenguaje de programación interpretado por el navegador, por lo cual no es necesario compilar para ser ejecutado. Es decir, las aplicaciones que han sido escritos en JavaScript pueden ser ejecutadas directamente en el navegador web sin requerir procesos intermedios. (24).
- **PHP:** Es un lenguaje de programación de código abierto desarrollado específicamente para desarrollo web, cuya principal contribución es la conexión entre los servidores y las interfaces de usuario. La mayor ventaja

de usar PHP es su simplicidad para aprender a usarlo, además de la popularidad con la que cuenta, que posibilita encontrar respuesta rápida a problemas comunes. (25).

- **Laravel:** Es un framework de código abierto para el desarrollo de aplicación web utilizando el lenguaje de programación PHP. Su filosofía es desarrollar código PHP de manera simple y elegante, evitando un código espagueti o un código poco estructurado. Laravel proporciona una estructura con características que permiten a los desarrolladores enfocarse en problemas específicos. (26).

En la tabla 4, se describe las ventajas y desventajas del lenguaje PHP frente a Java y Python.

<b>Características</b>	<b>Java</b>	<b>Python</b>
<b>Mantenimiento y seguridad</b>	Java se considera más fácil de mantener y seguro, en comparación con PHP. Ha agregado características de seguridad integradas, mientras que PHP debe optar por otros marcos.	Python es fácil de mantener y tiene una amplia biblioteca estándar. PHP, por otro lado, es más difícil de mantener, especialmente cuando se hace más grande.
<b>Rendimiento y velocidad</b>	Java está pre compilado para el rendimiento, y PHP requiere tiempo para cumplir con el código de bytes en cada solicitud.	Después del lanzamiento de PHP 7, este lenguaje informático es más rápido que Python.
<b>Capacidad de aprendizaje:</b>	Java y PHP se consideran los lenguajes más sencillos, según los conocimientos y habilidades previos de los desarrolladores. Algunos creen que la biblioteca de Java es demasiado profunda.	Python tiene una curva de aprendizaje empinada, pero es fácil de entender con una curva de aprendizaje constante. Python es muy legible con una sintaxis sencilla.
<b>Rendimiento y velocidad</b>	Java está pre compilado para el rendimiento y PHP requiere tiempo para cumplir con el código.	PHP es perfecto para programación web. Python, por el contrario, es un lenguaje de programación de propósito general, ideal para proyectos a largo plazo.

**Tabla 4. Cuadro Comparativo Lenguaje PHP**

Fuente: (27)

### 2.2.3 Inteligencia Artificial:

El concepto de Inteligencia Artificial hace referencia a máquinas o sistemas que simulan la inteligencia humana para la realización de procesos o tareas, que pueden mejorar de manera iterativa a partir de la información que recopilan. La Inteligencia Artificial se enfoca sobre el proceso y el análisis de datos, mejorando el rendimiento y productividad de las empresas a través de la automatización de procesos o tareas recurrentes que requieren mucho esfuerzo humano. (28).

En la figura 2, se muestran los subcampos que engloba la inteligencia artificial.

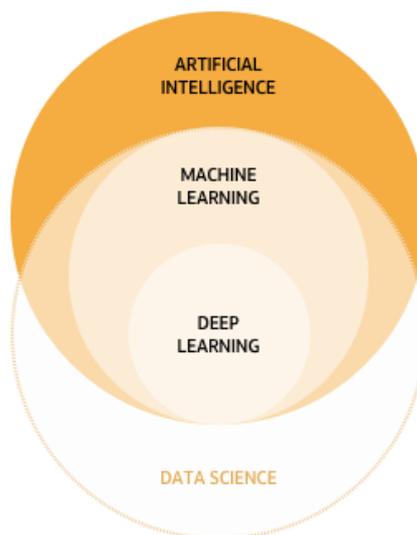


Figura 2. Sub campos de la inteligencia artificial

Fuente: (28)

#### 2.2.3.1 Aprendizaje Automático:

El Aprendizaje Automático realiza la tarea de analizar datos, identificar patrones, y hacer uso de los conocimientos obtenidos para finalizar mejor la tarea asignada. Toda tarea o proceso en las cuales se pueda identificar patrones de datos o reglas pueden ser automatizadas haciendo uso del Aprendizaje Automático, inclusive las tareas más complejas. (29).

- **Aprendizaje supervisado:**

El aprendizaje supervisado se caracteriza por tener un conjunto de datos que son etiquetados, los cuales permiten aprender a hacer una tarea donde los datos son utilizados para predecir un objetivo en un conjunto de datos nuevo. Este es

el modelo menos complicado, ya que trata de simular el aprendizaje humano. (29).

- **Aprendizaje No supervisado:**

El aprendizaje no supervisado se caracteriza por no contar con datos etiquetados, lo que busca es extraer del conjunto de datos, patrones o conocimientos previamente desconocidos; se diferencia del aprendizaje supervisado por no tener un conocimiento a priori. Existen diferentes formas en que los algoritmos de aprendizaje automático realizan esta tarea. (29).

En la tabla 5, se detalla las diferencias entre aprendizaje supervisado y no supervisado.

	Aprendizaje Supervisado	Aprendizaje No Supervisado
Casos de Uso	<ul style="list-style-type: none"> <li>• Problemas de clasificación, diagnósticos.</li> <li>• Detección de fraude de identidad).</li> <li>• Problemas de regresión (predicciones meteorológicas y de expectativa de vida).</li> </ul>	<ul style="list-style-type: none"> <li>• Problemas de <i>clustering</i>.</li> <li>• Agrupamientos de concurrencias.</li> <li>• Perfilado o <i>profiling</i>.</li> </ul>
Algoritmos	<ul style="list-style-type: none"> <li>• Árboles de decisión.</li> <li>• Clasificación de Naïve Bayes.</li> <li>• Regresión por mínimos cuadrados.</li> <li>• Regresión Logística.</li> <li>• Support Vector Machines (SVM).</li> <li>• Métodos "Ensemble" (Conjuntos de clasificadores).</li> </ul>	<ul style="list-style-type: none"> <li>• Algoritmos de <i>clustering</i>.</li> <li>• Análisis de componentes principales.</li> <li>• Descomposición en valores singulares.</li> <li>• Análisis de componentes principales.</li> </ul>

**Tabla 5. Cuadro comparativo: Aprendizaje supervisado y no supervisado**  
Fuente: (30)

### 2.2.3.2 Librerías Machine Learning:

- **Pytorch:** Es una biblioteca de código abierto usado para el Aprendizaje Automático. Se ejecuta de manera instantánea al ser imperativo, con la cual el usuario puede revisar si funciona correctamente antes de finalizar de escribir el código. Puede escribirse fracciones de código y verificarlo en tiempo real, es una implementación que utiliza el lenguaje de

programación Python para proporcionar compatibilidad, como una plataforma de Deep Learning. (31).

- **Tensorflow:** Es una biblioteca de software de código abierto para Aprendizaje Automático, que cuenta con un ecosistema de herramientas integrales, que ofrece varios niveles de abstracción que se adaptan a las necesidades y a la complejidad de los modelos a construir. (32).

En la tabla 5, se describe las diferencias entre los frameworks Pytorch y Tensorflow.

Pytorch	Tensorflow
<ul style="list-style-type: none"> <li>• Fue desarrollado por Facebook.</li> <li>• Fue hecho usando la biblioteca Torch.</li> <li>• Funciona en un concepto de gráfico dinámico.</li> <li>• Pytorch tiene menos funciones en comparación con Tensorflow.</li> <li>• Pytorch usa una API simple que ahorra todo el peso del modelo.</li> <li>• Es comparativamente menos favorable en implementaciones.</li> <li>• Es fácil de aprender y comprender.</li> <li>• Tiene un proceso computacional dinámico.</li> </ul>	<ul style="list-style-type: none"> <li>• Fue desarrollada por Google.</li> <li>• Se implementó en <i>theano</i>, que es una biblioteca Python.</li> <li>• Cree en un concepto de gráfico estático.</li> <li>• Tiene una funcionalidad de mayor nivel y proporciona un amplio espectro de opciones para trabajar.</li> <li>• Tiene la ventaja de que el gráfico completo se guarde como búfer de protocolo.</li> <li>• Es más compatible con implementaciones integradas y móviles.</li> <li>• Es comparativamente difícil de aprender.</li> <li>• Requiere el uso de una herramienta de depuración.</li> </ul>

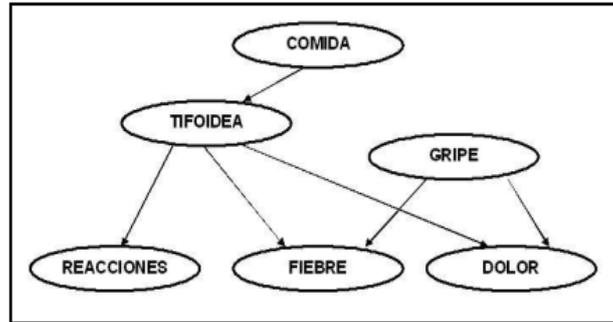
**Tabla 6. Cuadro comparativo Pytorch y Tensorflow**

Fuente: (33)

### 2.2.3.3 Algoritmos de clasificación

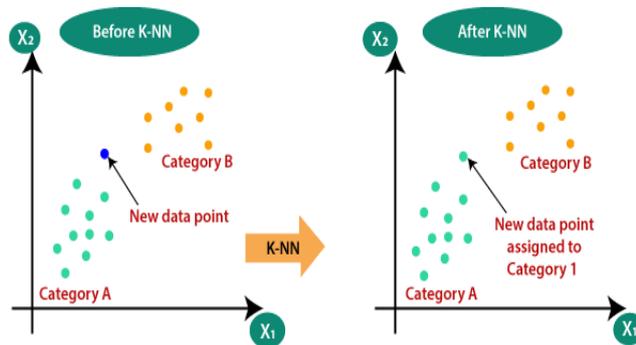
- **Redes Bayesianas:** Las redes bayesianas son un modelo caracterizado por representar gráficamente las relaciones de dependencia, en la cual cada nodo representa variables aleatorias; y sus arcos son la representación de relaciones de dependencia directa entre las variables, para afrontar problemas relacionados con la incertidumbre. (34).

En la figura 3, se muestra una representación gráfica del modelo de redes bayesianas.



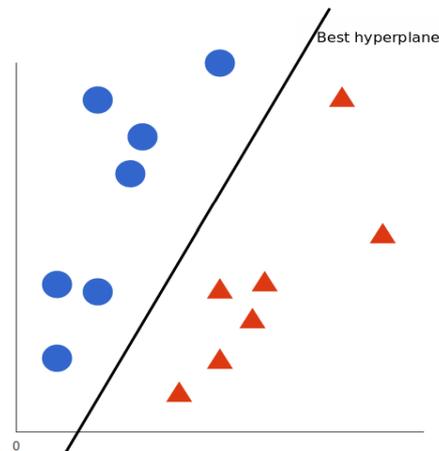
**Figura 3. Representación gráfica del modelo de redes bayesianas**  
Fuente : (34)

- **K Nearest Neighbor (KNN):** Es un algoritmo de aprendizaje supervisado, utilizado comúnmente para problemas de clasificación. A partir de un conjunto de datos inicial, su objetivo será el de clasificar correctamente todas las instancias al grupo que pertenezca, según tenga k vecinos más cerca de un determinado grupo o de otro; es un modelo que estima la probabilidad a posteriori de que un elemento pertenezca a una determinada clase. (35). En la figura 4, se muestra una representación gráfica del modelo de KNN.



**Figura 4. Representación gráfica del modelo KNN**  
Fuente: (35)

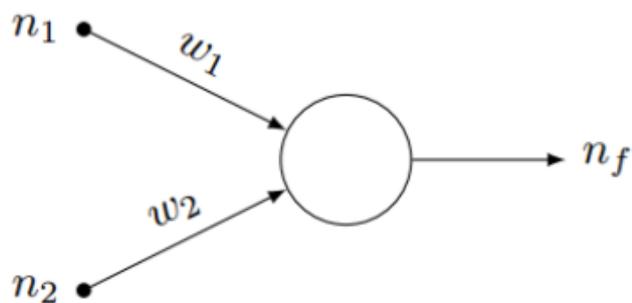
- **Vector Support Machines (SVM):** Es un algoritmo de aprendizaje automático supervisado, que puede ser utilizado para problemas de regresión o clasificación. A partir de dos o más categorías de datos etiquetados el algoritmo procede como un clasificador discriminativo, declarado expresamente por un hiperplano óptimo que separa todas las categorías. Los datos nuevos que luego se mapean en ese mismo espacio se pueden clasificar según el lado de la brecha en que se encuentran. (36). En la figura 5, se muestra una representación gráfica del modelo SVM.



**Figura 5. Representación gráfica del modelo SVM**  
**Fuente : (36)**

- **Redes Neuronales:** Las redes neuronales son un modelo que emula el modo de funcionamiento del cerebro humano. Las redes neuronales deben encontrar la composición que mejor se ajusta a ese proceso, conocido también como entrenamiento. Una red neuronal que ha pasado por el proceso de entrenamiento puede usarse para problemas de regresión o clasificación. Están compuestas por la capa de entrada, capa oculta donde se realizan los procesos de cálculo, y por la capa de salida, donde se obtienen los resultados procesados de la red neuronal. (37).

En la figura 6, se muestra una representación gráfica del modelo de redes neuronales.



**Figura 6. Representación gráfica del modelo de redes neuronales**  
**Fuente : (37)**

La distribución de neuronas dentro de la red se realiza estableciendo capas con un determinado número de neuronas para cada capa, se pueden distinguir tres tipos de capas. (38).

- **Capa de entrada:** Es la capa que recibe directamente la información externa a la red para ser procesada.

- **Capa oculta:** Las capas que se encuentran después de la capa de entrada son conocidas como capas ocultas, ya que no están expuestas directamente a la entrada, y pueden estar interconectadas de diversas maneras.

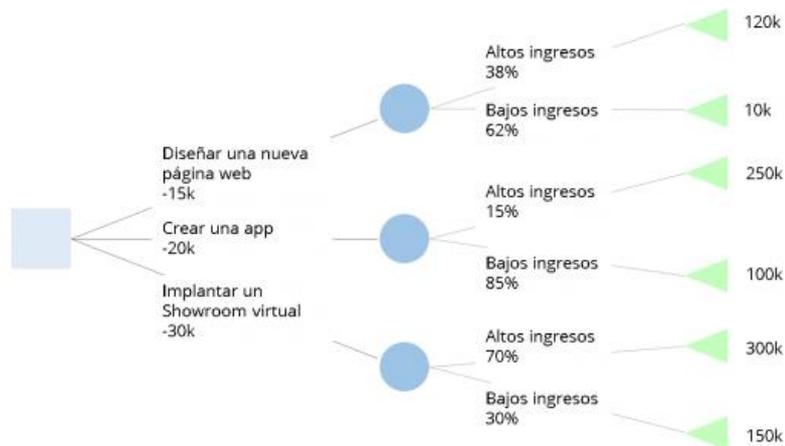
Diversos estudios han demostrado que una capa oculta es suficiente para la gran mayoría de problemas. Para determinar el número de neuronas en una capa oculta existen métodos aceptables, como el hecho de que el número de neuronas ocultas debe ser menos de dos veces el tamaño de la capa de entrada. (39).

- **Capa de salida:** Es la capa encargada de transmitir la información resultante al exterior.

Las neuronas artificiales cuentan con diversos estados de activación, los cuales determinan una salida de información a partir de los valores de entrada. (38).

- **Función lineal:** Los valores obtenidos mediante esta función de activación se fijan por encima o debajo de la zona en 1 o -1 respectivamente.
- **Función sigmoidea:** Los valores de salida obtenidos mediante esta función de activación están comprendidas en el rango de 0 a 1.
- **Función tangente hiperbólica:** Los valores de salida obtenidos mediante esta función de activación están comprendidas en el rango de -1 a 1.
- **Árboles de Decisión:** Un árbol de decisión es una especie de mapa en la que se muestran cada una de las posibles opciones de decisión y sus resultados. Es muy popular ya que puede representarse gráficamente, lo cual hace que sea muy fácil de entender; tiene una estructura donde cada nodo interno representa un atributo, cada rama representa las reglas de decisión, cada hoja representa el resultado y al nodo superior se le conoce como raíz.(40).

En la figura 7, se muestra una representación gráfica del modelo de árboles de decisión.



**Figura 7. Representación gráfica del modelo de árboles de decisión**  
**Fuente: (40)**

#### 2.2.3.4 Métricas de los modelos de ml

- **Exactitud:** Esta métrica representa el porcentaje de predicciones correctas por el modelo de clasificación, es una buena métrica cuando las clases tienen una proporción similar. (30).

En la figura 8, se muestra la fórmula para hallar la exactitud.

$$Accuracy = \frac{TN + TP}{TN + TP + FN + FP}$$

**Figura 8. Fórmula para hallar la exactitud**  
**Fuente: (30)**

- **Precisión:** Esta métrica indica cuántos resultados son realmente positivos de todas las predicciones positivas resultantes. Es la relación entre las predicciones positivas correctas y las predicciones positivas generales. (30).

En la figura 9, se muestra la fórmula para hallar la precisión.

$$Precision = \frac{TP}{TP + FP}$$

**Figura 9. Fórmula para hallar la precisión**  
**Fuente: (30)**

- **Sensibilidad:** Esta métrica indica de los valores realmente positivos, cuántos de estos se predicen como tales. Es la proporción de predicciones positivas que fueron correctas con relación al número total de positivos del conjunto de datos. (30)

En la figura 10, se muestra la fórmula para hallar la precisión.

$$Sensitivity = \frac{TP}{TP + FN}$$

**Figura 10. Fórmula para hallar la sensibilidad**  
Fuente: (30)

- **Matriz de confusión:** Si el modelo es de clasificación, una métrica muy utilizada es la matriz de confusión, que es la que indica para cada una de las posibles categorías o clases, su clasificación en cada una de las posibles opciones; cada columna de la matriz de confusión representa la cantidad de predicciones de cada clase, en tanto que cada fila representa las instancias en la clase real. La matriz facilita comprobar de manera sencilla si el modelo tiene tendencias a clasificar clases de manera no esperada, la diagonal representa los resultados que son clasificados de manera correcta, lo cual permite conocer el desempeño del modelo e identificar los resultados. (30).
  - **Verdaderos positivos**  
Son los casos en los que los datos reales utilizados indican que el valor es verdadero y el resultado obtenido también es verdadero.
  - **Verdaderos negativos**  
Son los casos en que los datos indican que el valor es negativo, y el modelo obtiene como resultado, verdadero.
  - **Falsos positivos**  
Son los casos de resultados obtenidos como verdaderos, pero que deberían ser negativos.
  - **Falsos negativos**  
Son los casos en que los datos reales indican que son verdaderos y el resultado es falso.

En la figura 11, se muestra un ejemplo de matriz de confusión.

		Predicted			$\Sigma$
		Iris-setosa	Iris-versicolor	Iris-virginica	
Actual	Iris-setosa	100.0 %	0.0 %	0.0 %	50
	Iris-versicolor	0.0 %	88.7 %	6.4 %	50
	Iris-virginica	0.0 %	11.3 %	93.6 %	50
$\Sigma$		50	53	47	150

Figura 11. Matriz de confusión

Fuente: (30)

## 2.3 Empresa

### 2.3.1 Datos generales de la institución

- **Nombre de la institución**  
CFCGROUP
- **Rubro o Giro del Negocio**  
Educación

#### 2.3.1.1 Breve Historia

CFCGROUP es creada con el propósito de brindar capacitación y asesoramiento a empresas y a personas que buscan desarrollarse dentro de su actividad laboral, potenciando y enriqueciendo sus conocimientos, adaptando sus programas educativos en función a los desafíos y exigencias del entorno, programas para las cuales existe gran demanda en el mercado laboral. Se brinda a los estudiantes excelente preparación académica, acompañada de una sólida formación en valores, para asegurar de esta forma una educación de calidad, para lo cual se está en constante capacitación de docentes y la totalidad de

nuestro equipo. CFC GROUP es una empresa con experiencia en el sector de capacitación, formación y asesoramiento de empresas y profesionales especializados en el sector minería y construcción.

### 2.3.1.2 Organigrama actual

La empresa CFCGROUP tiene una estructura casi plana, la cual se muestra en la figura 12.

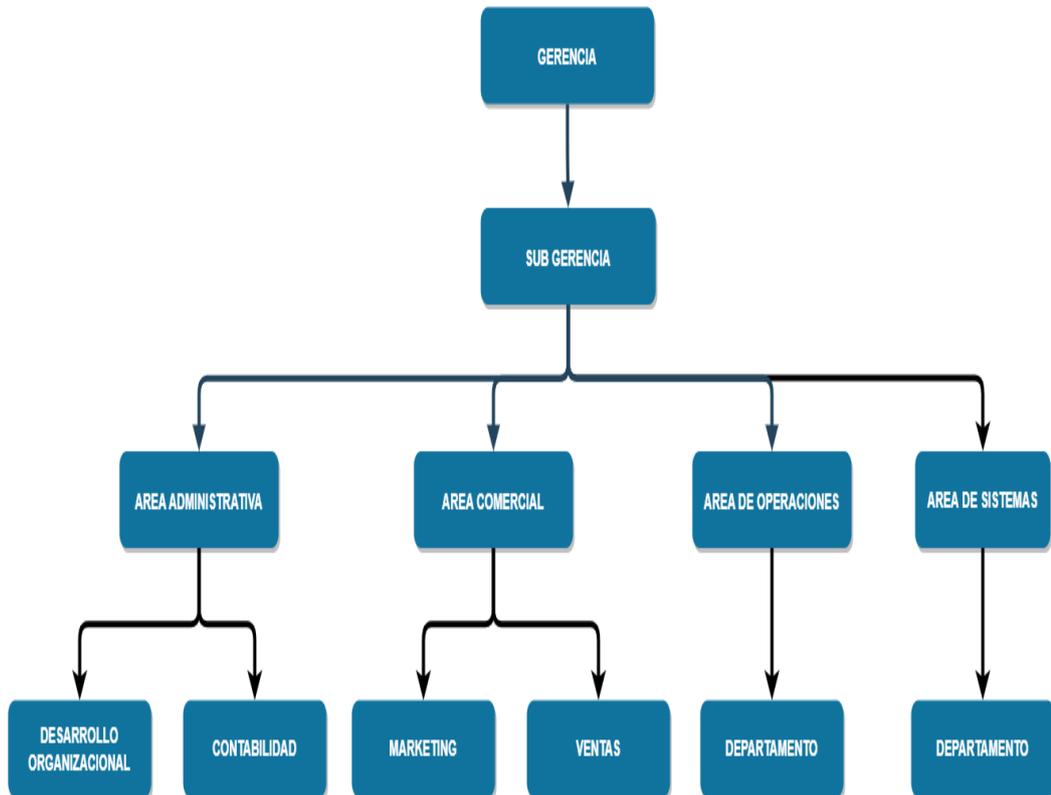
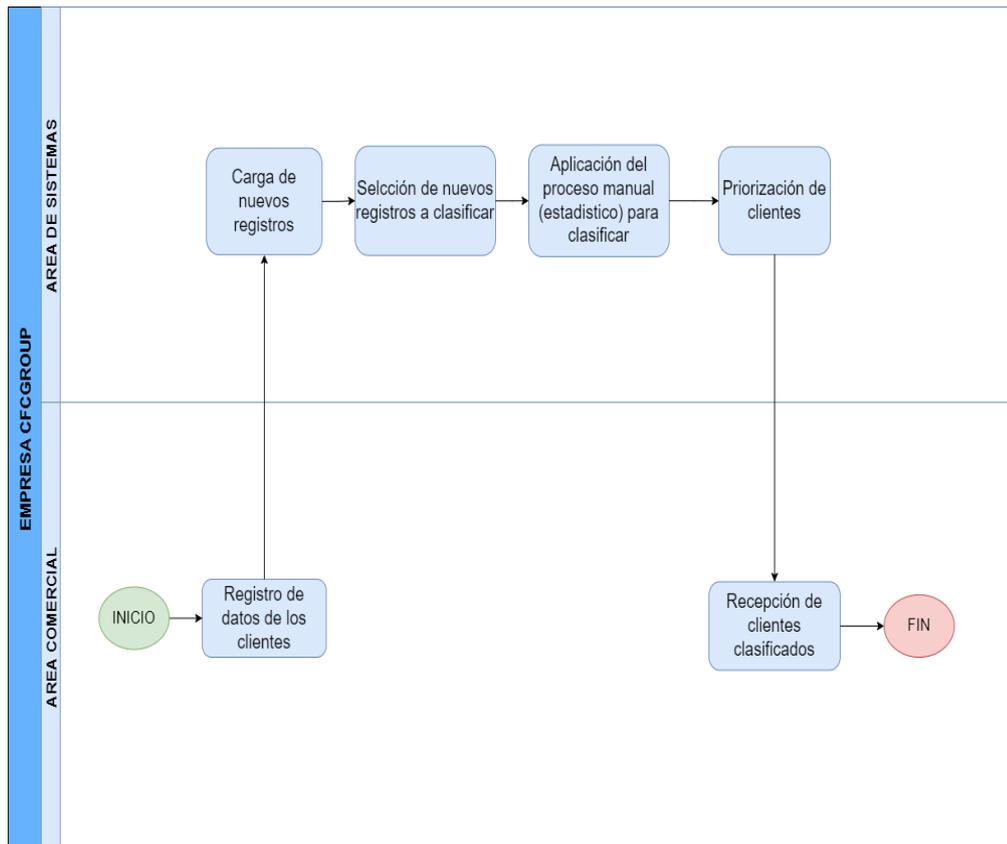


Figura 12. Organigrama empresa CFCGROUP  
Fuente (41)

### 2.3.1.3 Área de sistemas

El área de sistemas de la empresa CFCGROUP ayuda a la gerencia y a las demás áreas a desarrollar capacidades para la toma de decisiones; esta área tiene relación con los sistemas y las tecnologías de la información, haciendo hincapié en la importancia que tienen estos para la consecución de los objetivos de la empresa.

En la figura 13, se muestra el proceso manual realizado por la empresa para clasificar a sus clientes.



**Figura 13. Proceso de clasificación manual de clientes**

**Fuente: Elaboración propia**

## 2.4 SCRUM:

### 2.4.1 Metodología aplicada para el desarrollo de la solución:

Scrum es un marco de trabajo que permite el trabajo colaborativo, reúne un conjunto de buenas prácticas para el desarrollo de software. Se basa en lo empírico, lo que garantiza que el conocimiento obtenido proviene de la experiencia y de la toma de decisiones, teniendo como referencia lo que se conoce. Scrum emplea un enfoque incremental e iterativo para optimizar la predictibilidad. (42).

### 2.4.2 Roles de Scrum

El Equipo Scrum está compuesto por un Dueño de Producto (Product Owner), un Scrum Master y el Equipo de Desarrollo (Development Team). Los Equipos Scrum son auto organizados además de que cada uno de los integrantes desempeñan varias funciones. La característica de los equipos auto organizados es que optan por la mejor forma de llevar a cabo su trabajo. (42).

### **2.4.2.1 Product Owner**

El Product Owner es el responsable de garantizar el resultado óptimo realizado por el equipo de trabajo.

La manera en que se realiza la maximización del valor por parte del Product Owner es variable dependiendo de los individuos involucrados, las reglas de las organizaciones y lo equipos.

El Product Owner también es responsable de la gestión efectiva del Product Backlog, lo que incluye:

- Informar y desarrollar detalladamente los objetivos del producto.
- Elaborar e informar explícitamente los puntos a tratar en el Product Backlog.
- Poner en orden los elementos del Product Backlog.
- Garantizar de que el Product Backlog sea entendible y transparente.

El Product Owner puede ejecutar el trabajo mencionado anteriormente o también tiene la facultad de asignar esta gestión a los otros miembros del equipo. Indistintamente de ello, el Product Owner continúa siendo la persona responsable de que el trabajo se ejecute adecuadamente.

Para que el Product Owner realice sus tareas adecuadamente, todos los miembros interesados deben acatar y respetar sus decisiones. Las decisiones tomadas por el Product Owner se muestran claramente en el contenido del Product Backlog, y mediante el Increment pueden inspeccionarse en la Sprint Review. (42)

### **2.4.2.2 Scrum Master**

El Scrum Master es el encargado de realizar adecuadamente Scrum como se establece en la guía de esta. Lo consigue dando soporte a todos los miembros y a los interesados que participan en el desarrollo del producto, entendiendo los conceptos teóricos y la práctica sugerida por Scrum.

El Scrum Master es encargado de que el equipo de Scrum alcance la efectividad deseada. Esto lo consigue dando apoyo al equipo en el uso correcto de sus prácticas, establecido en la guía de Scrum.

Los Scrum Masters son líderes que guían y apoyan al Scrum Team y a la organización.

El Scrum Master apoya de diversas formas:

- Guían al equipo de Scrum a auto gestionarse y organizarse en las funciones requeridas.
- Ayudar al equipo de Scrum a focalizarse en realizar Increments con el valor suficiente para cumplir la Definición de Terminado que se determinó.

- Garantizar de que los eventos de Scrum se realicen, sean efectivos y se rijan dentro de los tiempos de referencia sugeridos en la Guía de Scrum. (42).

### **2.4.2.3 Desarrolladores**

Los integrantes del Scrum Team que se responsabilizan por realizar los elementos de un Increment que se llevará a cabo en los Sprint, son Desarrolladores.

Las competencias particulares que necesitan los Desarrolladores son extensas y dependen del escenario de trabajo requerido. Los Desarrolladores están encargados de:

- Realizar un plan para desarrollar el Sprint y el Sprint Backlog;
- Comprometerse con brindar calidad al regirse por la Definición de Terminado que se determinó.
- Ajustar su plan de trabajo diario para conseguir el Objetivo del Sprint.
- Ser responsables de cumplir sus tareas.

### **2.4.3 Eventos de Scrum**

El Sprint es un recipiente para los eventos de Scrum. Todo evento en Scrum ayuda a inspeccionar y ajustar los artefactos Scrum. Estos eventos están particularmente diseñados para tener claridad en el trabajo que se realizará. No realizar los eventos según lo determinado lleva a la pérdida de oportunidad para analizar el trabajo que se está realizando. Los eventos se usan en Scrum con el fin de disminuir la cantidad de reuniones no establecidas en la guía de Scrum. (42),

#### **2.4.3.1 El Sprint**

Los Sprints son el pilar de Scrum, son eventos que duran como máximo un mes, logrando predictibilidad. El Sprint siguiente inicia inmediatamente después del término del Sprint anterior.

El trabajo requerido para conseguir el Objetivo del Producto, como la planificación del Sprint, las reuniones diarias y la revisión del sprint, se realizan dentro del Sprint.

En el transcurso del Sprint:

- No se ejecutan cambios que retrasen o interrumpan el Objetivo del Sprint.
- No se disminuye la calidad.
- Se reajusta el Product Backlog según se requiera.

- El alcance se puede revisar y ajustar con el Product Owner, de esta manera se pueden conseguir resultados óptimos.

Cuando un Sprint tiene una dificultad muy alta o mucha complejidad para cumplirse, el Objetivo del Sprint puede invalidarse debido a que las posibilidades de terminar incrementan y con ello los riesgos; para resolver estos problemas se pueden realizar Sprints más cortos, dividir el problema en tareas más pequeñas que puedan ser realizadas, y con esto, lograr controlar el riesgo asociado a no terminar con un Sprint en el periodo de tiempo establecido.

Se puede cancelar un Spring si se determina que el objetivo de este es obsoleto. El único que tiene la facultad de establecer como cancelado un Sprint es el Product Owner. (42).

#### **2.4.3.2 Sprint Planning**

La Sprint Planning inicializa el Sprint al determinar el trabajo que se llevará a cabo durante este. A través del trabajo colaborativo el Scrum Team realiza la planificación del Sprint.

Para dialogar sobre los elementos con mayor relevancia del Product Backlog, el Product Owner debe garantizar que los participantes del equipo estén bien preparados sobre el trabajo a realizar, y cómo estos están relacionados con el Objetivo del Producto.

Para realizar el Sprint Planning se establece un tiempo de ocho horas como máximo para un Sprint con duración de un mes. Para los Sprint que tienen una duración menor a un mes, este evento también tiene una menor duración. (42).

#### **2.4.3.3 Daily Scrum**

El Daily Scrum tiene como objetivo revisar el avance hacia el Objetivo del Sprint, y con esa información, realizar el ajuste del Sprint Backlog si este lo requiere, ajustando las tareas planificadas.

Para los desarrolladores, el Daily Scrum es un evento con una duración máxima de 15 minutos, esto con el propósito de reducir el nivel de dificultad; este evento se realiza en el mismo lugar y a la misma hora todos los días que se encuentren planificados en el Sprint. Si el Product Owner o el Scrum Master están comprometidos constantemente con el trabajo que se realiza, estos también deberán participar en el Scrum Daily como si fueran Desarrolladores.

Los Desarrolladores pueden hacer uso de técnicas o recursos que consideren adecuados, siempre que su Daily Scrum se enfoque con el avance que acerque al Objetivo del Sprint, genere solidez y no tenga retrasos para el siguiente día de trabajo, de esta manera promover la autogestión. (42).

#### **2.4.3.4 Sprint Review**

El Sprint Review tiene como objetivo analizar el resultado obtenido en el Sprint, de esta manera, establecer los ajustes necesarios. El Scrum Team realiza la presentación de los resultados conseguidos a las personas interesadas, revisando el avance hacia el Objetivo del Producto.

El Scrum Team y los interesados realizan una revisión de lo conseguido en el Sprint y de los cambios que se han realizado, teniendo en cuenta esta información, los participantes determinan el siguiente paso a realizar. El Product Backlog puede reajustarse en este evento, el Sprint Review es más una presentación de resultados que solo una reunión. (42).

#### **2.4.3.5 Sprint Retrospective**

El Sprint Retrospective tiene como objetivo planificar maneras de incrementar la efectividad.

En este evento el Scrum Team revisa el desarrollo del último Sprint tomando en cuenta los procesos, las interacciones y la Definición de Terminado.

Se identifican las causas que ocasionaron problemas y se inspecciona su origen. El equipo revisa los aciertos y metas cumplidas durante el Sprint, los problemas con los que se toparon y cómo los resolvieron.

El Scrum Team registra los cambios más significativos con el objetivo de incrementar la efectividad. Las mejoras con mayor importancia se priorizan, asimismo estas mejoras se pueden registrar en el Sprint Backlog para desarrollarse en el siguiente Sprint.

El Sprint Retrospective da como finalizado un Sprint. Para un Sprint de un mes el Sprint Retrospective tiene una duración máxima de tres horas, y para Sprints menores a un mes, este evento también tiene una duración menor. (42).

#### **2.4.4 Artefactos de Scrum**

Los artefactos de Scrum tienen el objetivo de mostrar la información importante de manera clara, estos artefactos representan un valor.

Cada uno de los artefactos engloba un compromiso que garantiza que se pueda medir el avance del trabajo:

- Para el Product Backlog, es el Objetivo del Producto.
- Para el Sprint Backlog, es el Objetivo del Sprint.
- Para el Increment, es la Definición de Terminado.

Los compromisos anteriormente mencionados existen para asegurar el empirismo y fortalecer los valores mencionados en la guía de Scrum. (42).

##### **2.4.4.1 Product Backlog**

El Product Backlog es una lista con un orden específico, y con los puntos necesarios para añadir valor al producto.

El Scrum Team toma el Product Backlog como la única fuente de trabajo a realizar.

Los puntos en la lista del Product Backlog que el equipo de Scrum puede establecer como terminados durante un Sprint, se consideran listos para ser elegidos en los eventos del Sprint Planning, después de haber realizado el refinamiento.

El refinamiento del Product Backlog comprende el especificar detalladamente y el dividir los elementos de la lista para que sean más precisos. (42).

##### **2.4.4.2 Sprint Backlog**

El grupo de elementos especificados en el Product Backlog que han sido elegidos para el Sprint, el plan establecido para proporcionar el Increment y el objetivo del Sprint, son los que componen el Sprint Backlog.

El Sprint Backlog, creado exclusivamente por los Desarrolladores, es una representación visible del trabajo que van realizando y la planificación que tienen sobre sus tareas; por esta razón, el Sprint Backlog se va actualizando durante el Sprint para conseguir el objetivo planeado, este debe estar explícitamente detallado para ser inspeccionado en el Daily Scrum.

Durante el Sprint los desarrolladores trabajan enfocándose en el objetivo de este, si el resultado se muestra diferente de lo esperado pueden solicitar cambiar el alcance del Sprint Backlog con el Product Owner, para así continuar sin dificultades y de esta manera no afectar el objetivo del Sprint. (42).

### 2.4.4.3 Increment

Un Increment es un paso específico que lleva al objetivo del Producto, cada uno de los Increments es adicionado a los demás Increments con una revisión detallada, lo cual nos asegura que estos estén funcionando juntos de manera correcta; para que proporcionen valor estos deben ser utilizables.

En un Sprint pueden crearse muchos Increments, el conjunto de estos Increments es presentado en el Sprint Review, no obstante, los Increments también pueden ser entregados a los interesados antes de finalizar el Sprint.

Para que el trabajo realizado pueda ser incluido en los Increments deberá cumplir con la Definición del Terminado establecida. (42).

## 2.5 Definición de Términos Básicos

- **Servidor Web**

Un servidor web es una computadora que almacena los archivos que componen un sitio web como archivos html, css, imágenes, etc. Los cuales se entregan al dispositivo del usuario final. (20).

- **Aplicación Web**

Se denomina aplicación web a las aplicaciones que los usuarios pueden utilizar realizando peticiones a un Servidor web, mediante Internet o de una intranet desde un navegador. (19).

- **Modelo de Clasificación**

Los modelos de clasificación dividen la información de entrada y realizan cálculos predictivos para asignar los datos a categorías determinadas. (43).

- **Exactitud**

La exactitud es el número de elementos que han sido correctamente clasificados, divididos entre el total de elementos del conjunto de pruebas. (44).

- **Matriz confusión**

Una matriz de confusión es una representación matricial de los resultados obtenidos por la clasificación, es utilizado para describir el rendimiento del modelo. (45).

- **Base de datos**

Una base de datos es una recopilación de información o datos organizados y estructurados, que son almacenados de forma electrónica en un sistema informático. (46).

- **Framework**

Es un entorno de trabajo que tiene el objetivo de facilitar la labor de programación mediante la incorporación de una serie de características y funciones que aceleran el proceso de desarrollo. (47).

- **Librería**

Es un archivo importable que se utiliza para el desarrollo de un software. Está compuesta de un código y datos, con la finalidad de ser utilizada por otros programas de manera totalmente autónoma. (48).

- **Tasa de aprendizaje**

La tasa de aprendizaje es un valor constante que el algoritmo alcanza a converger con ponderaciones óptimas. (49).

- **Número de iteraciones**

Es la cantidad de recorridos que el algoritmo realiza en los datos de aprendizaje. Un número mayor de recorridos resulta en un modelo que se adapta mejor a los datos de entrenamiento.

- **Sobreajuste**

El sobreajuste ocurre cuando un modelo se ajusta exactamente a los datos de entrenamiento. Cuando esto sucede el modelo no puede funcionar con precisión para los datos no vistos. (50).

- **Función de pérdida**

La función de pérdida mide la divergencia entre la predicción de un algoritmo de aprendizaje automático y la salida. (51).

## Capítulo III. METODOLOGÍA

### 3.1 Metodología aplicada para el desarrollo de la solución

La metodología empleada en el presente proyecto es Scrum, una metodología empleada en la gestión de proyectos de desarrollo de software, que resalta por ser moderna, flexible y altamente utilizada en entornos corporativos. Scrum emplea un enfoque iterativo e incremental para optimizar la predictibilidad y el control del riesgo. Scrum define tres roles principales: Product Owner, Scrum Master y equipo de desarrollo. De la misma forma Scrum define tres artefactos principales: Product Backlog, Sprint Backlog y los incrementos.

En la tabla 7, se muestra las fases de desarrollo de la metodología Scrum.

FASES DE LA METODOLOGÍA SCRUM	
INICIACIÓN	
PROYECT CHARTER	Elaboración del documento
	Verificación del documento
	Entrega del documento
	Aprobación del documento
PLANIFICACIÓN	
	Elaboración del documento

PLAN DE DIRECCIÓN DEL PROYECTO	Verificación del documento
	Entrega del documento
	Aprobación del documento
	Elaboración de pila de producto
	Validación de la pila de producto
<b>EJECUCIÓN</b>	
PILA DE PRODUCTO	Identificación de Historias de Usuario
	Descripción de Historias de Usuario
	Priorización de Historias de Usuario
	Verificación de Historia de Usuario
PROTOTIPO	Diseño de interfaces
	Verificación de Diseños
INTERFAZ	Diseño de Interfaz
	Verificación de Diseño
BASE DE DATOS	Diseño de Base de Datos
	Verificación de Base de Datos
<b>PRODUCTO - SPRINT</b>	
SPRINT 1	H-2021-0001
	H-2021-0002
	H-2021-0003
	H-2021-0004
SPRINT 2	H-2021-0005
	H-2021-0006
	H-2021-0007

	H-2021-0008
SPRINT 3	H-2021-0009
	H-2021-0010
	H-2021-0011
	H-2021-0012
	H-2021-0013
SPRINT 4	H-2021-0014
	H-2021-0015
	H-2021-0016
	H-2021-0017
	H-2021-0018
<b>SEGUIMIENTO Y CONTROL DEL PROYECTO</b>	
SPRINT 1	Desarrollo de reuniones (Actas de reunión y Documentos)
SPRINT 2	
SPRINT 3	
SPRINT 4	
<b>CIERRA</b>	
INFORME FINAL	Elaboración de documento (Lecciones aprendidas, oportunidades)
	Verificación de Documentos
	Entrega de Documentación del Proyecto

**Tabla 7. Fases de desarrollo**

**Fuente: Elaboración propia**

## Capítulo IV. ANÁLISIS Y DISEÑO DE LA SOLUCIÓN

### 4.1 Designación de Roles

En la tabla 8, se describe la designación de roles de los integrantes que participarán en el proyecto.

Rol	Persona	Código
Product Owner	Santa Cruz Álvarez Diego Esteban	SCD
	Zapana Llanquiche Renato Boris	ZLR
Scrum Master	Santa Cruz Álvarez Diego Esteban	SCD
Team	Santa Cruz Álvarez Diego Esteban	SCD
	Zapana Llanquiche Renato Boris	ZLR

**Tabla 8. Asignación de roles**

Fuente: Elaboración propia

### 4.2 Alcance General

**A. Alcance del producto:** Para el desarrollo del proyecto se establece lo siguiente:

- La aplicación web permitirá al usuario mantener un registro y control general de los clientes potenciales e históricos.

- La aplicación web permitirá al usuario clasificar los clientes potenciales de la empresa.
- El sistema web generará reportes relacionados con la información de clasificación de clientes potenciales.

**B. Alcance del proyecto:** Durante los Sprints se realizaron las siguientes actividades.

- Informe semanal de avance.
- Pila de Producto.
- Iteraciones del producto.
- Iteraciones de seguimiento y control.
- Manual del usuario y del sistema.

### 4.3 Identificación de Historias de usuario

Siguiendo la metodología Scrum se enuncia la pila de requerimientos, los cuales se detallan mediante historias de usuarios en la tabla 9.

Identificador de historia (ID)	Enunciado de la historia
HU-2022-0001	Como usuario registrado, necesito visualizar la lista de clientes potenciales, con la finalidad de tener un control de los clientes potenciales registrados.
HU-2022-0002	Como usuario registrado, necesito registrar nuevos clientes potenciales mediante un formulario, con la finalidad agregar nuevos registros a la lista
HU-2022-0003	Como usuario registrado, necesito modificar clientes potenciales, con la finalidad hacer cambios a la información de los registros de la lista.
HU-2022-0004	Como usuario registrado, necesito eliminar clientes potenciales, con la finalidad de quitar registros con datos erróneos o duplicados a la información de los registros de la lista.
HU-2022-0005	Como usuario registrado, necesito clasificar los clientes potenciales, con la finalidad de conocer los intereses de los registros clasificados.
HU-2022-0006	Como usuario registrado, necesito visualizar las estadísticas de los clientes clasificados, con la finalidad de conocer la información resultante de la clasificación

HU-2022-0007	Como usuario administrador, necesito exportar la lista de los clientes clasificados, con la finalidad de proporcionar la información a las áreas que lo soliciten.
HU-2022-0008	Como usuario registrado, necesito filtrar la lista de los clientes clasificados por categoría, con la finalidad de visualizar la lista por cada categoría.
HU-2022-0009	Como usuario registrado, necesito importar clientes de entrenamiento, con la finalidad de utilizarlos para la clasificación.
HU-2022-0010	Como usuario registrado, necesito visualizar la lista de clientes de entrenamiento, con la finalidad de tener un control de los clientes potenciales registrados.
HU-2022-0011	Como usuario registrado, necesito registrar nuevos clientes de entrenamiento mediante un formulario, con la finalidad agregar nuevos registros a la lista
HU-2022-0012	Como usuario registrado, necesito modificar clientes de entrenamiento, con la finalidad hacer cambios a la información de los registros de la lista.
HU-2022-0013	Como usuario registrado, necesito eliminar clientes de entrenamiento, con la finalidad de quitar registros con datos erróneos o duplicados a la información de los registros de la lista.
HU-2022-0014	Como usuario registrado, necesito acceder al sistema mediante un inicio de sesión, con la finalidad de mantener segura la información del sistema.
HU-2022-0015	Como usuario registrado necesito cerrar sesión, con la finalidad de mantener segura la información del sistema.
HU-2022-0016	Como usuario administrador, requiero visualizar las fechas de las clasificaciones realizadas, con la finalidad de tender el control sobre el registro de clasificaciones.
HU-2022-0017	Como usuario administrador, requiero registrar nuevos usuarios del sistema, con la finalidad de dar acceso a las demás personas encargadas.
HU-2022-0018	Como usuario administrador, necesito entrenar el modelo de clasificación mediante un botón, con la finalidad de actualizar el modelo con nuevos registros.

**Tabla 9. Product backlog**

**Fuente: Elaboración propia**

#### 4.4 Validación de Pila de usuarios

Siguiendo la metodología Scrum se enuncia las historias de usuario con su respectiva observación, los cuales se detallan en la tabla 10.

Identificador de historia (ID)	Enunciado de la historia	Observación del Cliente o Sponsor
HU-2022-0001	Como usuario registrado, requiero visualizar la lista de clientes potenciales, con la finalidad de tener un control de los clientes potenciales registrados.	Aceptado
HU-2022-0002	Como usuario registrado, requiero registrar nuevos clientes potenciales mediante un formulario, con la finalidad agregar nuevos registros a la lista	Aceptado
HU-2022-0003	Como usuario registrado, necesito modificar clientes potenciales, con la finalidad hacer cambios a la información de los registros de la lista.	Aceptado
HU-2022-0004	Como usuario registrado, necesito eliminar clientes potenciales, con la finalidad de quitar registros con datos erróneos o duplicados a la información de los registros de la lista.	Aceptado
HU-2022-0005	Como usuario registrado, necesito clasificar los clientes potenciales, con la finalidad de conocer los intereses de los registros clasificados.	Mostrar los datos clasificados por cada categoría
HU-2022-0006	Como usuario registrado, necesito visualizar las estadísticas de los clientes clasificados, con la finalidad de conocer la información resultante de la clasificación	Aceptado
HU-2022-0007	Como usuario administrador, requiero exportar la lista de los clientes clasificados, con la finalidad de proporcionar la información a las áreas que lo soliciten.	Aceptado
HU-2022-0008	Como usuario registrado, requiero filtrar la lista de los clientes clasificados por categoría, con la finalidad de visualizar la lista por cada categoría.	Aceptado

HU-2022-0009	Como usuario registrado, requiero importar clientes de entrenamiento, con la finalidad de utilizarlos para la clasificación.	Aceptado
HU-2022-0010	Como usuario registrado, requiero visualizar la lista de clientes de entrenamiento, con la finalidad de tener un control de los clientes potenciales registrados.	Aceptado
HU-2022-0011	Como usuario registrado, requiero registrar nuevos clientes de entrenamiento mediante un formulario, con la finalidad agregar nuevos registros a la lista.	Aceptado
HU-2022-0012	Como usuario registrado, necesito modificar clientes de entrenamiento, con la finalidad hacer cambios a la información de los registros de la lista.	Aceptado
HU-2022-0013	Como usuario registrado, necesito eliminar clientes de entrenamiento, con la finalidad de quitar registros con datos erróneos o duplicados a la información de los registros de la lista.	Aceptado
HU-2022-0014	Como usuario registrado, necesito acceder al sistema mediante un inicio de sesión, con la finalidad de mantener segura la información del sistema.	Aceptado
HU-2022-0015	Como usuario registrado necesito cerrar sesión, con la finalidad de mantener segura la información del sistema.	Aceptado
HU-2022-0016	Como usuario administrador, requiero visualizar las fechas de las clasificaciones realizadas, con la finalidad de tender el control sobre el registro de clasificaciones.	Aceptado
HU-2022-0017	Como usuario administrador, requiero registrar nuevos usuarios del sistema, con la finalidad de dar acceso a las demás personas encargadas.	Aceptado
HU-2022-0018	Como usuario administrador, necesito entrenar el modelo de clasificación mediante un botón, con la finalidad de actualizar el modelo con nuevos registros.	Aceptado

**Tabla 10. Validación de Historias de usuario**

**Fuente: Elaboración propia**

## 4.5 Historias de Usuario

A continuación, se detallan las historias de usuarios de cada uno de los requerimientos planteados.

Historia de Usuario	
<b>Identificador</b>	HU-2022-0001
<b>Rol</b>	Como usuario registrado.
<b>Característica / Funcionalidad</b>	Necesito visualizar la lista de clientes potenciales.
<b>Resultado / Razón</b>	Con la finalidad de tener un control de los clientes potenciales registrados.
<b>Prioridad</b>	Alta
<b>Criterios de Aceptación</b>	El usuario deberá visualizar la lista de los clientes a través de una tabla.
	La tabla deberá mostrar la lista completa de todos los registros existentes.
	Se deberá mostrar la cantidad total de registros en la tabla.
<b>Restricciones / Validaciones</b>	Solo deberá mostrarse las columnas: nombre, teléfono, edad, sexo, departamento, estado civil, cargo y años de experiencia.

**Tabla 11. Historia de usuario 1**

**Fuente: Elaboración propia**

Historia de Usuario	
<b>Identificador</b>	HU-2022-0002
<b>Rol</b>	Como usuario registrado.
<b>Característica / Funcionalidad</b>	Necesito registrar nuevos clientes potenciales mediante un formulario.
<b>Resultado / Razón</b>	Con la finalidad de tener un control de los clientes potenciales registrados.
<b>Prioridad</b>	Alta
<b>Criterios de Aceptación</b>	El usuario deberá poder crear nuevos registros de clientes a través de un formulario.

	Se deberá validar la información registrada en campos del formulario.
<b>Restricciones / Validaciones</b>	No se admiten campos vacíos al registrar un cliente.
	Para todos los campos del formulario, se deberá validar la información ingresada.

**Tabla 12. Historia de usuario 2**

Fuente: Elaboración propia

Historia de Usuario	
<b>Identificador</b>	HU-2022-0003
<b>Rol</b>	Como usuario registrado.
<b>Característica / Funcionalidad</b>	Necesito modificar clientes potenciales.
<b>Resultado / Razón</b>	Con la finalidad hacer cambios a la información de los registros de la lista.
<b>Prioridad</b>	Alta
<b>Criterios de Aceptación</b>	El usuario deberá poder modificar la lista de los clientes a través de un formulario que permitan cambiar datos.
<b>Restricciones / Validaciones</b>	No se admiten campos vacíos al modificar un registro.
	Todos los campos del formulario deberán validar la información modificada.

**Tabla 13. Historia de usuario 3**

Fuente: Elaboración propia

Historia de Usuario	
<b>Identificador</b>	HU-2022-0004
<b>Rol</b>	Como usuario registrado.
<b>Característica / Funcionalidad</b>	Necesito eliminar clientes potenciales.

<b>Resultado / Razón</b>	Con la finalidad de quitar registros con datos erróneos o duplicados a la información de los registros de la lista.
<b>Prioridad</b>	Alta
<b>Criterios de Aceptación</b>	El sistema deberá mostrar un botón para eliminar el registro en cada fila.
	El sistema deberá mostrar en la tabla los datos actualizados.
<b>Restricciones / Validaciones</b>	El sistema deberá mostrar mensaje de confirmación al intentar eliminar el registro.

**Tabla 14. Historia de usuario 4**

**Fuente: Elaboración propia**

Historia de Usuario	
<b>Identificador</b>	HU-2022-0005
<b>Rol</b>	Como usuario registrado.
<b>Característica / Funcionalidad</b>	Necesito clasificar los clientes potenciales.
<b>Resultado / Razón</b>	Con la finalidad de conocer los intereses de la registros clasificados.
<b>Prioridad</b>	Alta
<b>Criterios de Aceptación</b>	El usuario deberá poder clasificar los clientes potenciales a través de un botón.
<b>Restricciones / Validaciones</b>	El botón para clasificar deberá mostrar los resultados en una tabla.

**Tabla 15. Historia de usuario 5**

**Fuente: Elaboración propia**

Historia de Usuario	
<b>Identificador</b>	HU-2022-0006

<b>Rol</b>	Como usuario registrado.
<b>Característica / Funcionalidad</b>	Necesito visualizar las estadísticas de los clientes clasificados.
<b>Resultado / Razón</b>	Con la finalidad de conocer la información resultante de la clasificación.
<b>Prioridad</b>	Alta
<b>Criterios de Aceptación</b>	El usuario deberá visualizar la lista de los clientes clasificados a través de una tabla.
<b>Restricciones / Validaciones</b>	El sistema deberá mostrar mediante gráficos estadísticos los resultados.

**Tabla 16. Historia de usuario 6**

Fuente: Elaboración propia

Historia de Usuario	
<b>Identificador</b>	HU-2022-0007
<b>Rol</b>	Como usuario administrador.
<b>Característica / Funcionalidad</b>	Requiero exportar la lista de los clientes clasificados.
<b>Resultado / Razón</b>	Con la finalidad de proporcionar la información a las áreas que lo soliciten.
<b>Prioridad</b>	Alta
<b>Criterios de Aceptación</b>	El usuario deberá poder generar reportes o informes de los clientes clasificados en formato “.xlsx”.
<b>Restricciones / Validaciones</b>	Los datos a exportar solo serán los que se muestran en la tabla.

**Tabla 17. Historia de usuario 7**

Fuente: Elaboración propia

Historia de Usuario	
<b>Identificador</b>	HU-2022-0008
<b>Rol</b>	Como usuario registrado.
<b>Característica / Funcionalidad</b>	Requiero filtrar la lista de los clientes clasificados por categoría.
<b>Resultado / Razón</b>	Con la finalidad de visualizar la lista por cada categoría.
<b>Prioridad</b>	Alta
<b>Criterios de Aceptación</b>	El usuario deberá visualizar los datos filtrados en una tabla.
	La tabla deberá mostrar la lista completa de todos los registros existentes separados por categorías.
<b>Restricciones / Validaciones</b>	El filtro para clasificar clientes deberá contener solo las categorías resultantes de la clasificación.

**Tabla 18. Historia de usuario 8**

Fuente: Elaboración propia

Historia de Usuario	
<b>Identificador</b>	HU-2022-0009
<b>Rol</b>	Como usuario registrado.
<b>Característica / Funcionalidad</b>	Requiero importar clientes de entrenamiento.
<b>Resultado / Razón</b>	Con la finalidad de utilizarlos para la clasificación.
<b>Prioridad</b>	Alta
<b>Criterios de Aceptación</b>	El usuario deberá poder importar clientes de entrenamientos en formato “.xlsx”.
<b>Restricciones / Validaciones</b>	Solo se aceptarán los datos de importación que coincidan con la tabla de clientes de entrenamiento.

**Tabla 19. Historia de usuario 9**

Fuente: Elaboración propia

Historia de Usuario	
<b>Identificador</b>	HU-2022-0010
<b>Rol</b>	Como usuario registrado.
<b>Característica / Funcionalidad</b>	Requiero visualizar la lista de clientes de entrenamiento.
<b>Resultado / Razón</b>	Con la finalidad de tener un control de los clientes potenciales registrados.
<b>Prioridad</b>	Alta
<b>Criterios de Aceptación</b>	El usuario deberá visualizar la lista de los clientes potenciales a través de una tabla.
	La tabla deberá mostrar la lista completa de todos los registros existentes.
<b>Restricciones / Validaciones</b>	Solo deberá mostrarse las columnas: nombre, teléfono, edad, sexo, departamento, estado civil, cargo, años de experiencia y categoría.

**Tabla 20. Historia de usuario 10**

Fuente: Elaboración propia

Historia de Usuario	
<b>Identificador</b>	HU-2022-0011
<b>Rol</b>	Como usuario registrado.
<b>Característica / Funcionalidad</b>	Requiero registrar nuevos clientes de entrenamiento mediante un formulario.
<b>Resultado / Razón</b>	Con la finalidad agregar nuevos registros a la lista.
<b>Prioridad</b>	Alta
<b>Criterios de Aceptación</b>	El usuario deberá poder crear nuevos registros de clientes de entrenamiento a través de un formulario.
<b>Restricciones / Validaciones</b>	No se admiten campos vacíos al registrar un cliente.
	Para todos los campos del formulario se deberá validar la información ingresada.

**Tabla 21. Historia de usuario 11**

Fuente: Elaboración propia

Historia de Usuario	
<b>Identificador</b>	HU-2022-0012
<b>Rol</b>	Como usuario registrado.
<b>Característica / Funcionalidad</b>	Necesito modificar clientes de entrenamiento.
<b>Resultado / Razón</b>	Con la finalidad hacer cambios a la información de los registros de la lista.
<b>Prioridad</b>	Alta
<b>Criterios de Aceptación</b>	El usuario deberá poder modificar la lista de los clientes de entrenamiento a través de un formulario que permitan cambiar datos.
<b>Restricciones / Validaciones</b>	No se admiten campos vacíos al modificar un registro.
	Todos los campos del formulario deberán validar la información modificada.

**Tabla 22. Historia de usuario 12**

**Fuente: Elaboración propia**

Historia de Usuario	
<b>Identificador</b>	HU-2022-0013
<b>Rol</b>	Como usuario registrado.
<b>Característica / Funcionalidad</b>	Necesito eliminar clientes de entrenamiento.
<b>Resultado / Razón</b>	Con la finalidad de quitar registros con datos erróneos o duplicados a la información de los registros de la lista.
<b>Prioridad</b>	Alta
<b>Criterios de Aceptación</b>	El sistema deberá mostrar un botón para eliminar el registro en cada fila.
	El sistema deberá mostrar en la tabla los datos actualizados.

<b>Restricciones / Validaciones</b>	El sistema deberá mostrar un mensaje de confirmación al intentar eliminar el registro.
-------------------------------------	--

**Tabla 23. Historia de usuario 13**

Fuente: Elaboración propia

Historia de Usuario	
<b>Identificador</b>	HU-2022-0014
<b>Rol</b>	Como usuario registrado.
<b>Característica / Funcionalidad</b>	Necesito acceder al sistema mediante un inicio de sesión.
<b>Resultado / Razón</b>	Con la finalidad de mantener segura la información del sistema.
<b>Prioridad</b>	Alta
<b>Criterios de Aceptación</b>	El usuario deberá ingresar al sistema mediante un formulario con usuario y contraseña.
<b>Restricciones / Validaciones</b>	El sistema deberá poder validar la información ingresada.

**Tabla 24. Historia de usuario 14**

Fuente: Elaboración propia

Historia de Usuario	
<b>Identificador</b>	HU-2022-0015
<b>Rol</b>	Como usuario registrado.
<b>Característica / Funcionalidad</b>	Necesito cerrar sesión.
<b>Resultado / Razón</b>	Con la finalidad de mantener segura la información del sistema.
<b>Prioridad</b>	Alta

<b>Criterios de Aceptación</b>	El usuario deberá poder cerrar sesión del sistema mediante un botón.
<b>Restricciones / Validaciones</b>	El sistema deberá redirigir al usuario al inicio de sesión.

**Tabla 25. Historia de usuario 15**

Fuente: Elaboración propia

Historia de Usuario	
<b>Identificador</b>	HU-2022-0016
<b>Rol</b>	Como usuario administrador.
<b>Característica / Funcionalidad</b>	Necesito visualizar la lista de clientes potenciales.
<b>Resultado / Razón</b>	Requiero visualizar las fechas de las clasificaciones realizadas.
<b>Prioridad</b>	Alta
<b>Criterios de Aceptación</b>	El usuario deberá visualizar el historial de las clasificaciones realizadas en una tabla.
<b>Restricciones / Validaciones</b>	La tabla solo deberá mostrar la fecha y la cantidad de registros utilizados.

**Tabla 26. Historia de usuario 16**

Fuente: Elaboración propia

Historia de Usuario	
<b>Identificador</b>	HU-2022-0017
<b>Rol</b>	Como usuario Administrador.
<b>Característica / Funcionalidad</b>	Requiero registrar nuevos usuarios del sistema.
<b>Resultado / Razón</b>	Con la finalidad de dar acceso a las demás personas encargadas.
<b>Prioridad</b>	Alta

<b>Criterios de Aceptación</b>	El usuario deberá poder crear nuevos registros de Usuarios del sistema a través de un formulario.
<b>Restricciones / Validaciones</b>	No se admiten campos vacíos al registrar un cliente.
	Para todos los campos del formulario, se deberá validar la información ingresada.

**Tabla 27. Historia de usuario 17**

Fuente: Elaboración propia

Historia de Usuario	
<b>Identificador</b>	HU-2022-0018
<b>Rol</b>	Como usuario administrador.
<b>Característica / Funcionalidad</b>	Necesito entrenar el modelo de clasificación mediante un botón.
<b>Resultado / Razón</b>	Con la finalidad de actualizar el modelo con nuevos registros.
<b>Prioridad</b>	Alta
<b>Criterios de Aceptación</b>	El usuario deberá poder realizar el entrenamiento mediante un botón.
<b>Restricciones / Validaciones</b>	Debe mostrar un mensaje de confirmación antes de realizar el entrenamiento.

**Tabla 28. Historia de usuario 18**

Fuente: Elaboración propia

## 4.6 Planificación de Sprints

### 4.6.1 Historias de usuarios de los usuarios registrados

En la tabla 29, se visualiza las historias de usuario que requieren que el usuario esté registrado e identificado por ID de historia.

Identificador de historia (ID)	Estado	Dimensión / Esfuerzo	Sprint	Prioridad	Responsable
HU-2022-0001	Planificado	3 días	1	4	SCD
HU-2022-0002	Planificado	3 días	1	4	SCD
HU-2022-0003	Planificado	4 días	1	4	ZLR
HU-2022-0004	Planificado	4 días	1	4	ZLR
HU-2022-0005	Planificado	3 días	2	3	SCD
HU-2022-0006	Planificado	3 días	2	4	SCD
HU-2022-0008	Planificado	5 días	2	4	ZLR
HU-2022-0009	Planificado	4 días	3	3	SCD
HU-2022-0010	Planificado	4 días	3	4	SCD
HU-2022-0011	Planificado	4 días	3	4	SCD
HU-2022-0012	Planificado	3 días	3	4	SCD
HU-2022-0013	Planificado	5 días	4	4	SCD
HU-2022-0014	Planificado	3 días	4	5	SCD
HU-2022-0015	Planificado	5 días	4	4	SCD

**Tabla 29. Planificación de Historias de Usuario de los usuarios registrados**

Fuente: Elaboración propia

#### 4.6.2 Historias de usuarios de usuario Administrador

En la tabla 30, se visualiza las historias de usuario que requieren que el usuario sea administrador identificado por ID de historia.

Identificador de historia (ID)	Estado	Dimensión / Esfuerzo	Sprint	Prioridad	Responsable
HU-2022-0007	Planificado	5 días	2	4	ZLR
HU-2022-0016	Planificado	3 días	4	4	ZLR
HU-2022-0017	Planificado	3 días	4	5	SCD
HU-2022-0018	Planificado	3 días	4	5	ZLR

**Tabla 30. Planificación de Historias de Usuario del usuario administrador**

Fuente: Elaboración propia

Código	Estado
1	Insignificante
2	Baja
3	Media
4	Alta
5	Urgente

**Tabla 31. Leyenda de prioridad**

Fuente: Elaboración propia

#### 4.7 Arquitectura de solución

En la figura 14, se muestra la arquitectura de la aplicación web. La solución propuesta es una aplicación web que hace uso de la biblioteca de aprendizaje automático tensorflow.js para la clasificación de clientes potenciales, al igual que hace uso del Framework Laravel para el desarrollo de la aplicación web, junto con la base de datos Mysql para el almacenamiento y consulta de los datos de la aplicación.

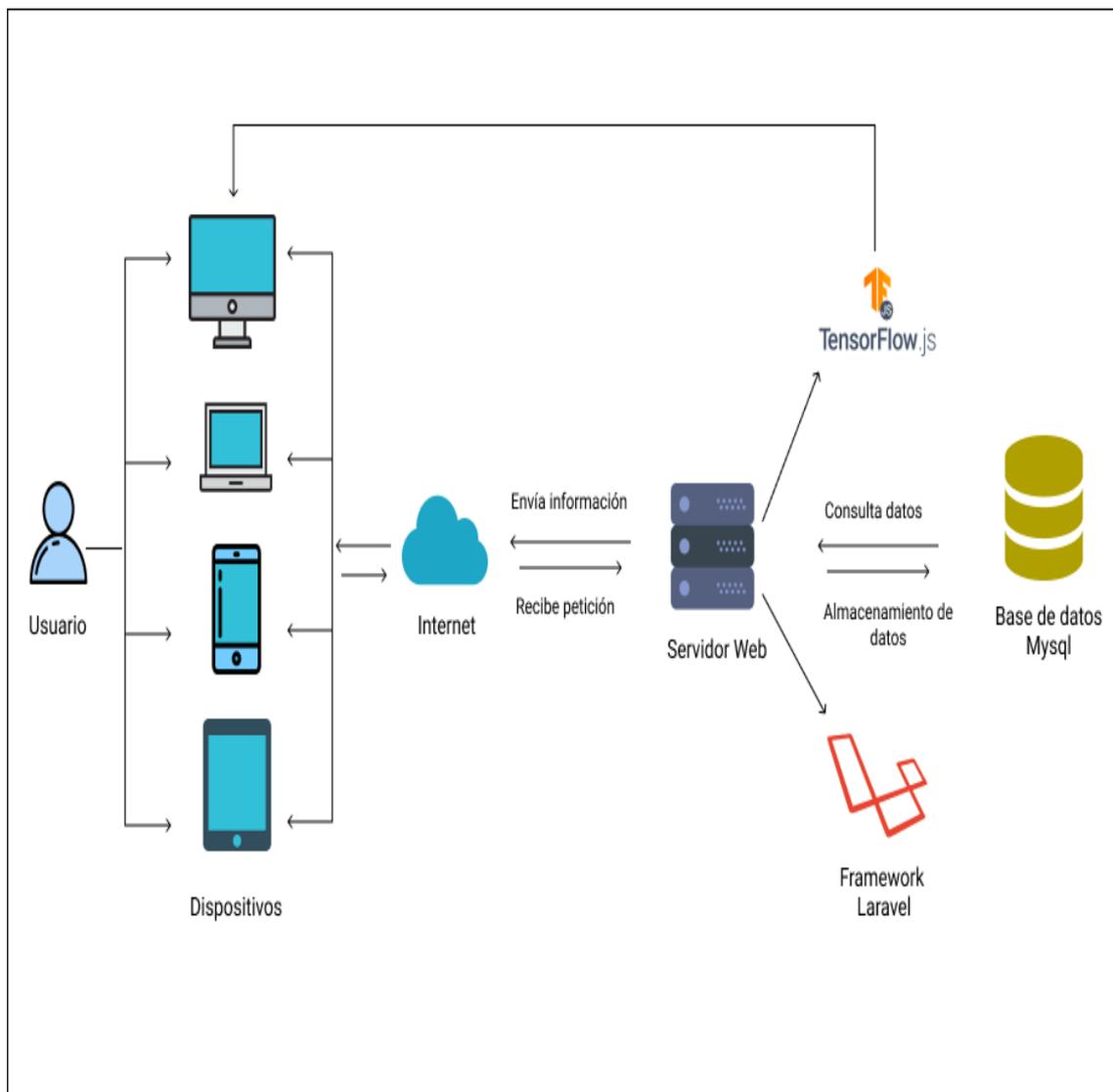


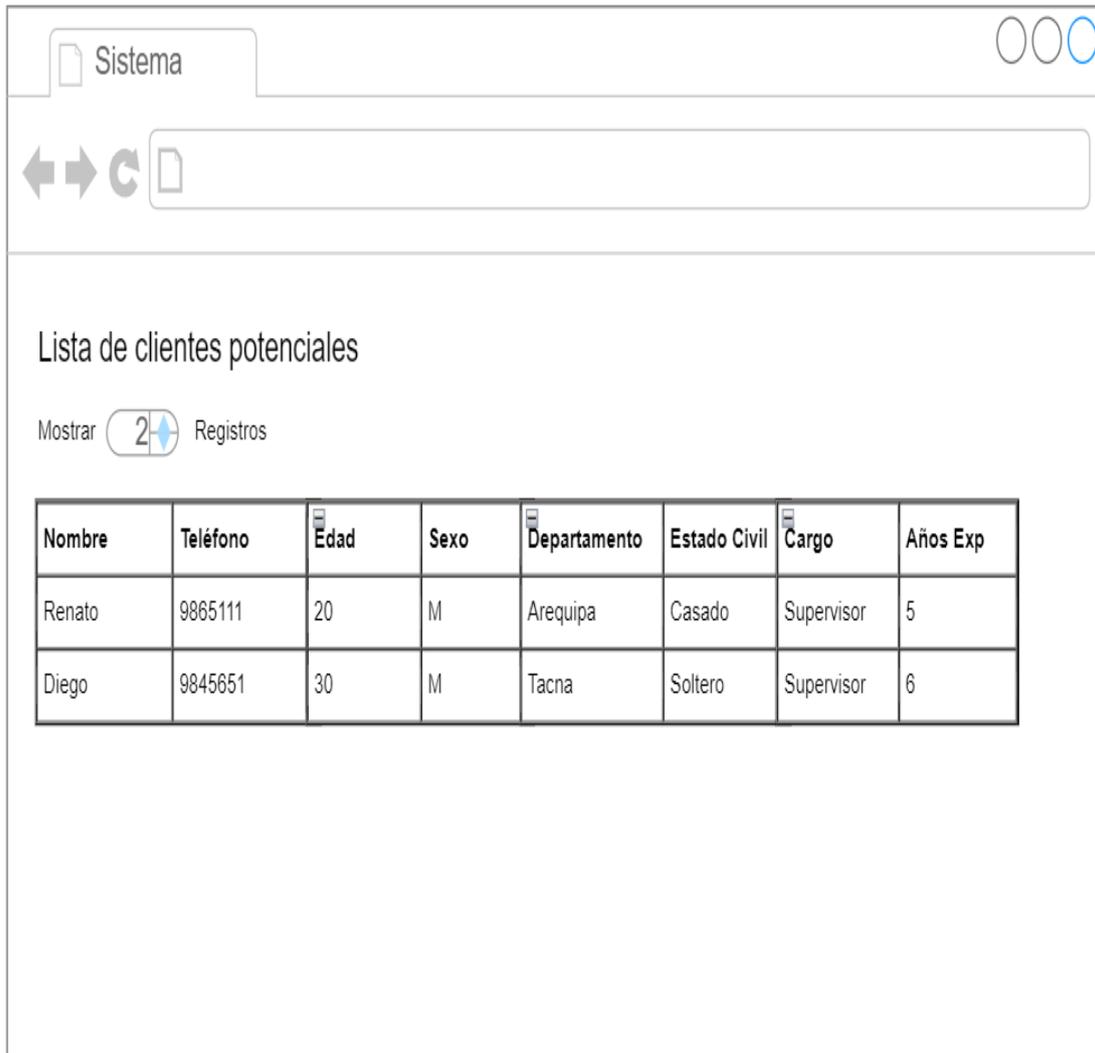
Figura 14. Arquitectura de la aplicación web

Fuente: Elaboración propia

## 4.8 Diseño de interfaces

Se elaboraron los mockups (diseños o modelos de interfaces) de la solución propuesta, se muestran a continuación.

Historia de usuario 1: Como usuario registrado, requiero visualizar la lista de clientes potenciales, con la finalidad de tener un control de los clientes potenciales registrados.



**Figura 15. Interfaz Historia de usuario 1**  
**Fuente: Elaboración propia**

Historia de usuario 2: Como usuario registrado, requiero registrar nuevos clientes potenciales mediante un formulario, con la finalidad de agregar nuevos registros a la lista.

The image shows a web browser window with a single tab labeled 'Sistema'. The address bar contains navigation icons (back, forward, refresh) and a document icon. The main content area displays a form titled 'Nuevo Cliente Potencial'. The form includes the following fields and controls:

- Nombre: Text input field.
- Telefono: Text input field.
- Email: Text input field.
- Fecha de nacimiento: Text input field.
- Sexo: Radio buttons for 'M' (Male) and 'F' (Female).
- Departamento: Dropdown menu.
- Cargo: Dropdown menu.
- Años experiencia: Text input field.
- Registrar Cliente: A prominent blue button at the bottom of the form.

**Figura 16. Interfaz Historia de usuario 2**

**Fuente: Elaboración propia**

Historia de usuario 3: Como usuario registrado, necesito modificar clientes potenciales, con la finalidad de hacer cambios a la información de los registros de la lista.

Sistema

← → ↻ 📄

### Modificar Cliente Potencial

Nombre

Telefono

Email

Fecha de nacimiento

M  F

Departamento

Cargo

Años experiencia

**Figura 17. Interfaz Historia de usuario 3**  
**Fuente: Elaboración propia**

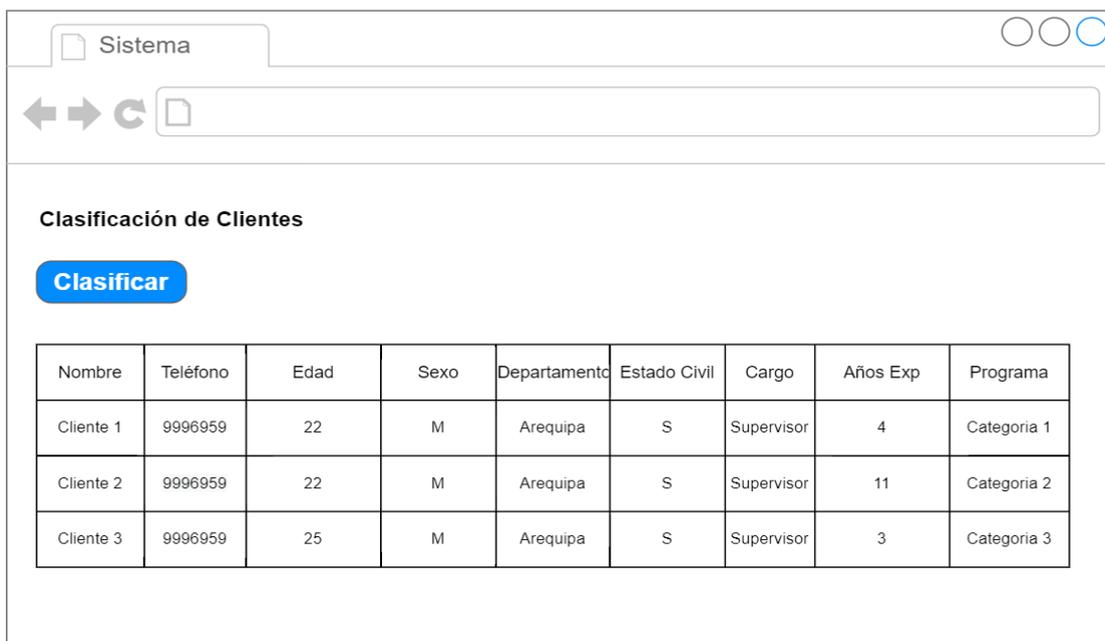
Historia de usuario 4: Como usuario registrado, necesito eliminar clientes potenciales, con la finalidad de quitar registros con datos erróneos o duplicados a la información de los registros de la lista.



**Figura 18. Interfaz Historia de usuario 4**

**Fuente: Elaboración propia**

Historia de usuario 5: Como usuario registrado, necesito clasificar los clientes potenciales, con la finalidad de conocer los intereses de los registros clasificados.



**Figura 19. Interfaz Historia de usuario 5**

**Fuente: Elaboración propia**

Historia de usuario 6: Como usuario registrado, necesito visualizar las estadísticas de los clientes clasificados, con la finalidad de conocer la información resultante de la clasificación.



**Figura 20. Interfaz Historia de usuario 6**

**Fuente: Elaboración propia**

Historia de usuario 7: Como usuario administrador, requiero exportar la lista de los clientes clasificados, con la finalidad de proporcionar la información a las áreas que lo soliciten.

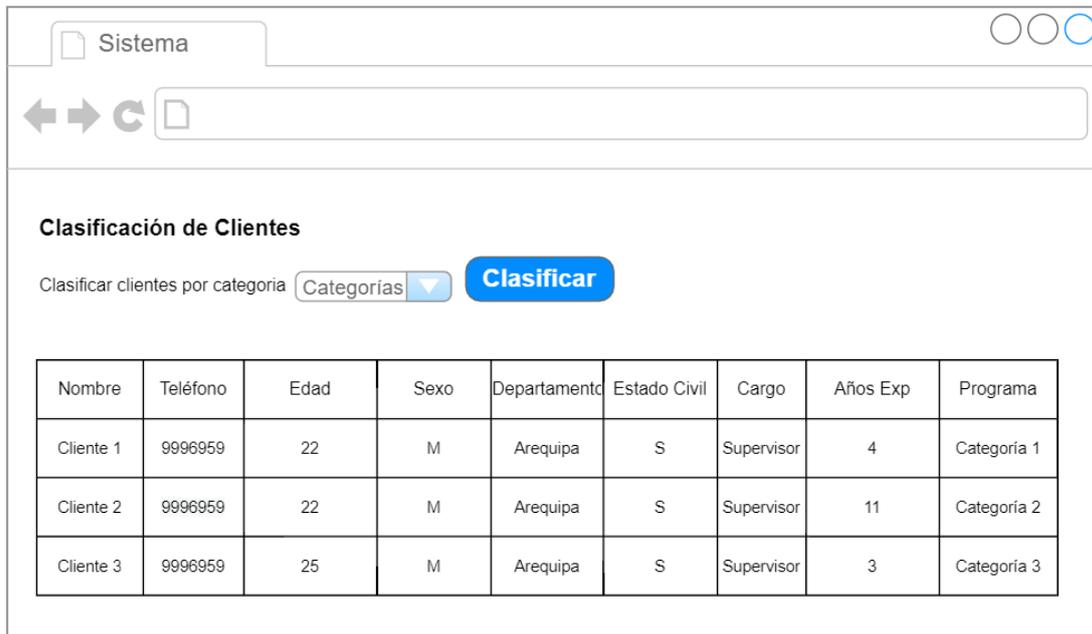
The figure shows a table with 9 columns and 4 rows. The columns are: Nombre, Teléfono, Edad, Sexo, Departamentc, Estado Civil, Cargo, Años Exp, and Programa. The rows represent three different clients. There is a green 'Exportar' button in the top right corner of the table area.

Nombre	Teléfono	Edad	Sexo	Departamentc	Estado Civil	Cargo	Años Exp	Programa
Cliente 1	9996959	22	M	Arequipa	S	Supervisor	4	Categoría 1
Cliente 2	9996959	22	M	Arequipa	S	Supervisor	11	Categoría 2
Cliente 3	9996959	25	M	Arequipa	S	Supervisor	3	Categoría 3

**Figura 21. Interfaz Historia de usuario 7**

**Fuente: Elaboración propia**

Historia de usuario 8: Como usuario registrado, requiero filtrar la lista de los clientes clasificados por categoría, con la finalidad de visualizar la lista por cada categoría.



**Figura 22. Interfaz Historia de usuario 8**

**Fuente: Elaboración propia**

Historia de usuario 9: Como usuario registrado, requiero importar clientes de entrenamiento, con la finalidad de utilizarlos para la clasificación.



**Figura 23. Interfaz Historia de usuario 9**

**Fuente: Elaboración propia**

Historia de usuario 10: Como usuario registrado, requiero visualizar la lista de clientes de entrenamiento, con la finalidad de tener un control de los clientes potenciales registrados.



**Figura 24. Interfaz Historia de usuario 10**

**Fuente: Elaboración propia**

Historia de usuario 11: Como usuario registrado, requiero registrar nuevos clientes de entrenamiento mediante un formulario, con la finalidad de agregar nuevos registros a la lista.

The image shows a web browser window with a single tab titled 'Sistema'. The address bar is empty. The main content area displays a form titled 'Nuevo Cliente de entrenamiento'. The form contains the following elements:

- Text input field: Nombre
- Text input field: Telefono
- Text input field: Email
- Text input field: Fecha de nacimiento
- Radio buttons: M (Male) and F (Female)
- Dropdown menu: Departamento
- Dropdown menu: Cargo
- Text input field: Años experiencia
- Dropdown menu: Programa
- Blue button: Actualizar registro

**Figura 25. Interfaz Historia de usuario 11**

**Fuente: Elaboración propia**

Historia de usuario 12: Como usuario registrado, necesito modificar clientes de entrenamiento, con la finalidad de hacer cambios a la información de los registros de la lista.

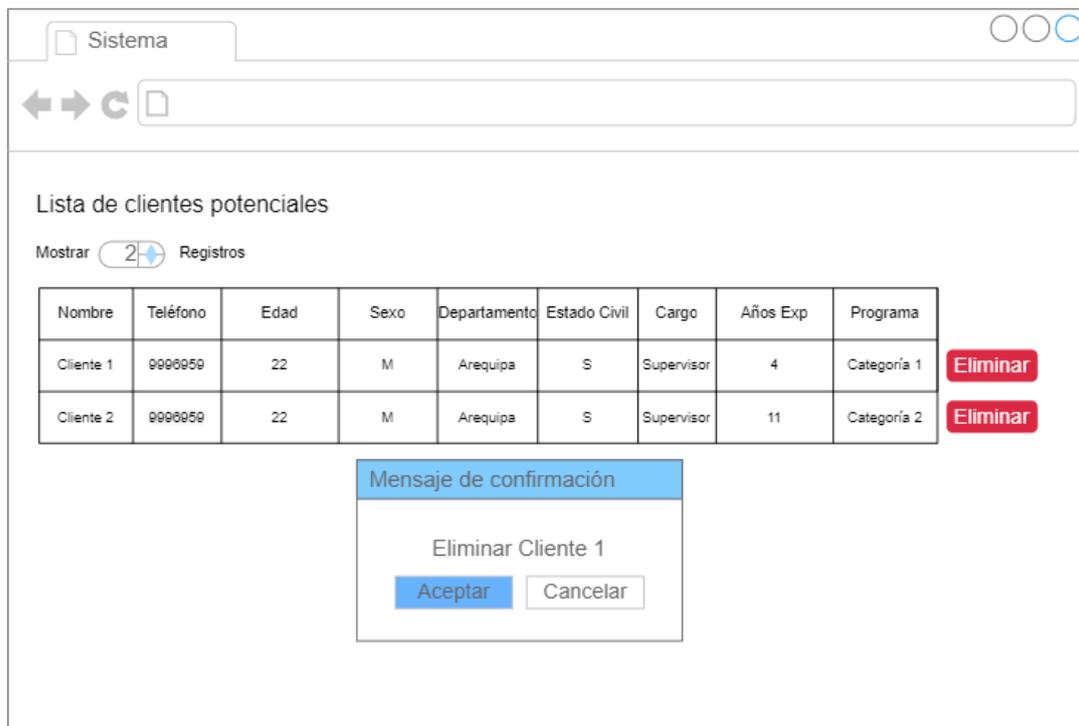
The image shows a web browser window with a single tab titled 'Sistema'. The browser's address bar is empty. The main content area displays a form titled 'Modificar Cliente de entrenamiento'. The form contains the following elements:

- Nombre: Text input field.
- Telefono: Text input field.
- Email: Text input field.
- Fecha de nacimiento: Text input field.
- Gender: Radio buttons for 'M' (Male) and 'F' (Female).
- Departamento: Dropdown menu.
- Cargo: Dropdown menu.
- Años experiencia: Text input field.
- Programa: Dropdown menu.
- Actualizar registro: A prominent blue button at the bottom of the form.

**Figura 26. Interfaz Historia de usuario 12**

**Fuente: Elaboración propia**

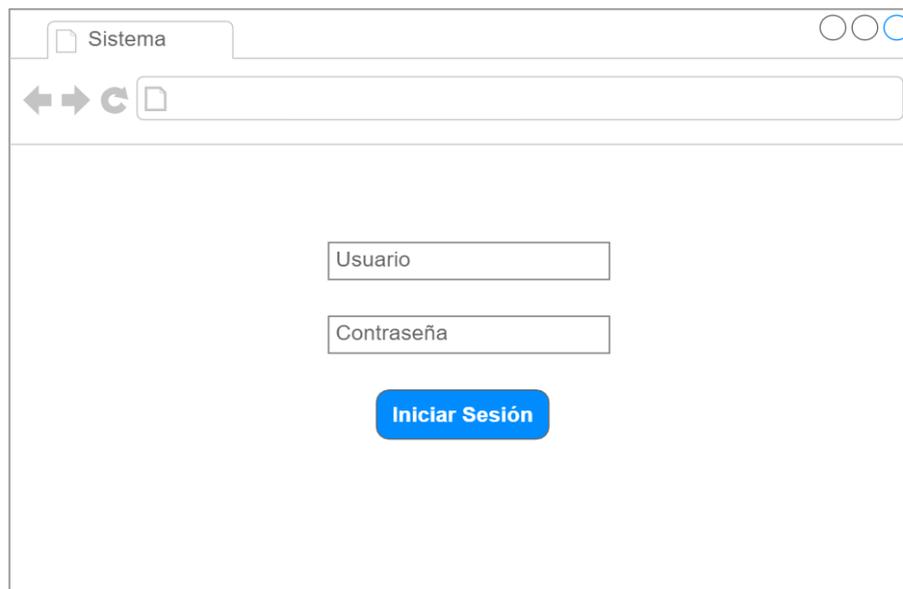
Historia de usuario 13: Como usuario registrado, necesito eliminar clientes de entrenamiento, con la finalidad de quitar registros con datos erróneos o duplicados a la información de los registros de la lista.



**Figura 27. Interfaz Historia de usuario 13**

**Fuente: Elaboración propia**

Historia de usuario 14: Como usuario registrado, necesito acceder al sistema mediante un inicio de sesión, con la finalidad de mantener segura la información del sistema.



**Figura 28. Interfaz Historia de usuario 14**

**Fuente: Elaboración propia**

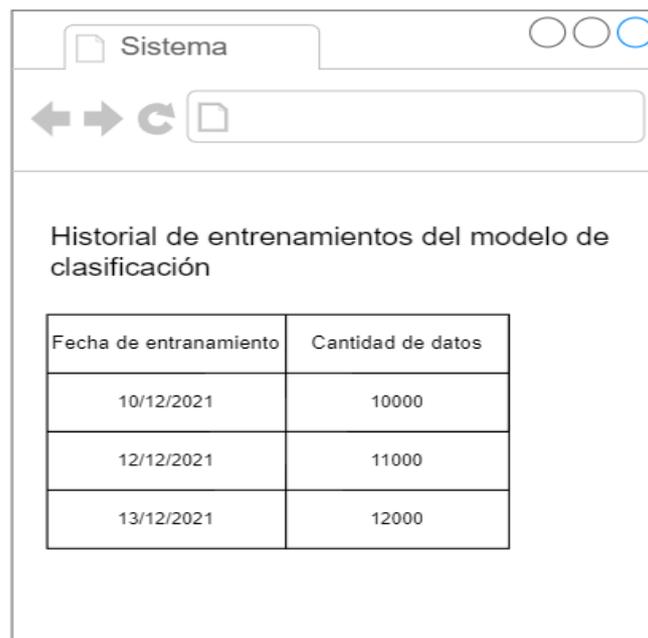
Historia de usuario 15: Como usuario registrado necesito cerrar sesión, con la finalidad de mantener segura la información del sistema.



**Figura 29. Interfaz Historia de usuario 15**

**Fuente: Elaboración propia**

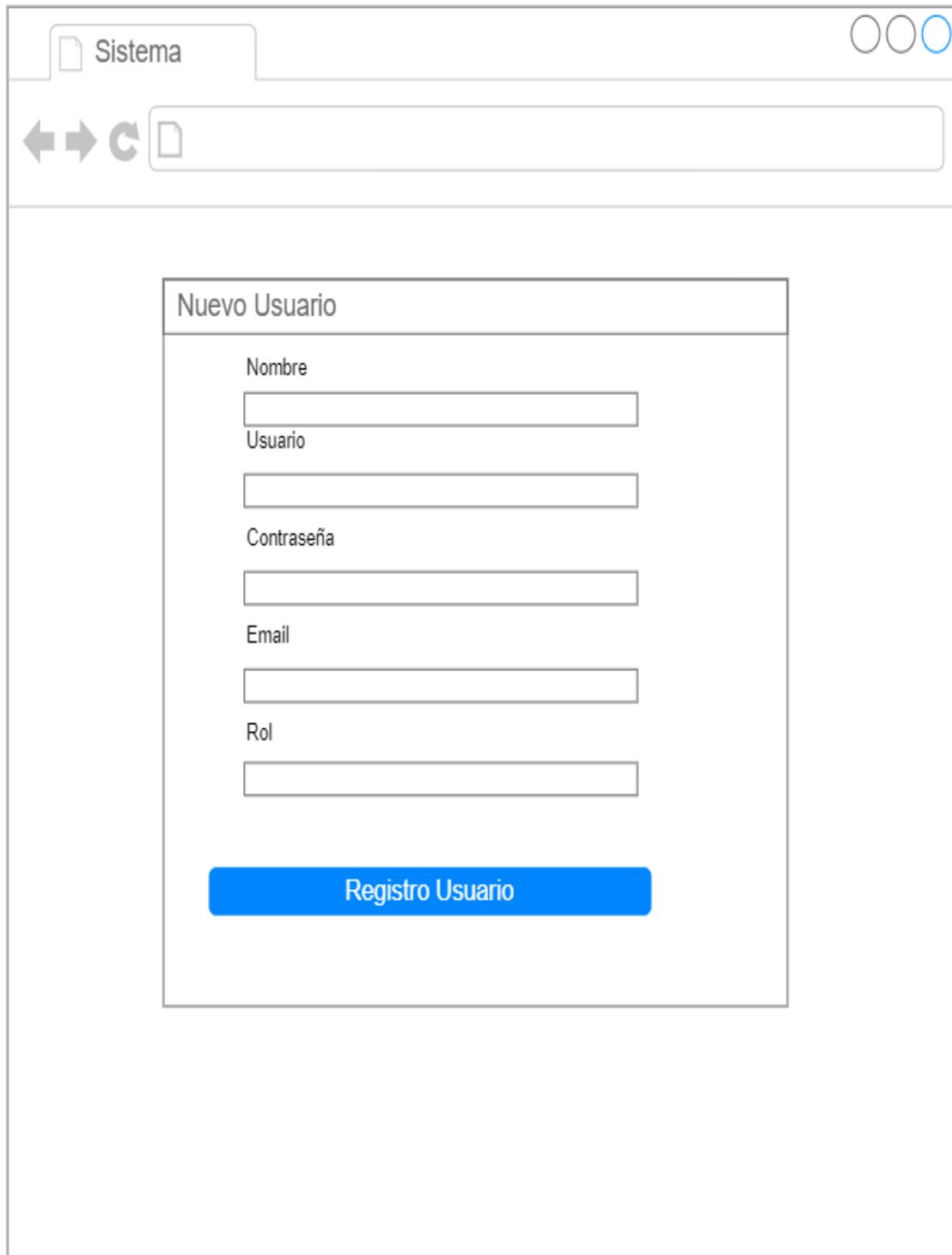
Historia de usuario 16: Como usuario administrador, requiero visualizar las fechas de las clasificaciones realizadas, con la finalidad de tender control sobre el registro de clasificaciones.



**Figura 30. Interfaz Historia de usuario 16**

**Fuente: Elaboración propia**

Historia de usuario 17: Como usuario administrador, requiero registrar nuevos usuarios del sistema, con la finalidad de dar acceso a las demás personas encargadas.



The image shows a web browser window with a single tab titled 'Sistema'. The address bar contains navigation icons (back, forward, refresh) and a document icon. The main content area displays a form titled 'Nuevo Usuario'. The form contains five text input fields labeled 'Nombre', 'Usuario', 'Contraseña', 'Email', and 'Rol'. Below these fields is a prominent blue button with the text 'Registro Usuario'.

**Figura 31. Interfaz Historia de usuario 17**  
**Fuente: Elaboración propia**

Historia de usuario 18: Como usuario administrador, necesito entrenar el modelo de clasificación mediante un botón, con la finalidad de actualizar el modelo con nuevos registros.



**Figura 32. Interfaz Historia de usuario 18**

**Fuente: Elaboración propia**

#### **4.9 Validación de diseño de interfaces**

Después de realizar el diseño de interfaces en *mockups* se realizó una validación como se muestra en la tabla 32, donde se comparó la utilización y la secuencia de las interfaces diseñadas y relacionadas con las historias de usuario.

Identificador de historia (ID)	Enunciado de la historia	Secuencia de interfaces
HU-2022-0001	Como usuario registrado, necesito visualizar la lista de clientes potenciales, con la finalidad de tener un control de los clientes potenciales registrados.	I14->Menú-I1
HU-2022-0002	Como usuario registrado, necesito registrar nuevos clientes potenciales mediante un formulario, con la finalidad de agregar nuevos registros a la lista	I14->Menú-I1->I2
HU-2022-0003	Como usuario registrado, necesito modificar clientes potenciales, con la finalidad de hacer cambios a la información de los registros de la lista.	I14->Menú-I1->I3
HU-2022-0004	Como usuario registrado, necesito eliminar clientes potenciales, con la finalidad de quitar registros con datos erróneos o duplicados a la información de los registros de la lista.	I14->Menú-I1->I4
HU-2022-0005	Como usuario registrado, necesito clasificar los clientes potenciales, con la finalidad de conocer los intereses de los registros clasificados.	I14->Menú->I5
HU-2022-0006	Como usuario registrado, necesito visualizar las estadísticas de los clientes clasificados, con la finalidad de conocer la información resultante de la clasificación.	I14->Menú->I6
HU-2022-0007	Como usuario administrador, necesito exportar la lista de los clientes clasificados, con la finalidad de proporcionar la información a las áreas que lo soliciten.	I14->Menú->I5->I7
HU-2022-0008	Como usuario registrado, necesito filtrar la lista de los clientes clasificados por categoría, con la finalidad de visualizar la lista por cada categoría.	I14->Menú->I5->I8
HU-2022-0009	Como usuario registrado, necesito importar clientes de entrenamiento, con la finalidad de utilizarlos para la clasificación.	I14->Menú->I10->I9
HU-2022-0010	Como usuario registrado, necesito visualizar la lista de clientes de entrenamiento, con la finalidad de tener un control de los clientes potenciales registrados.	I14->Menú->I10

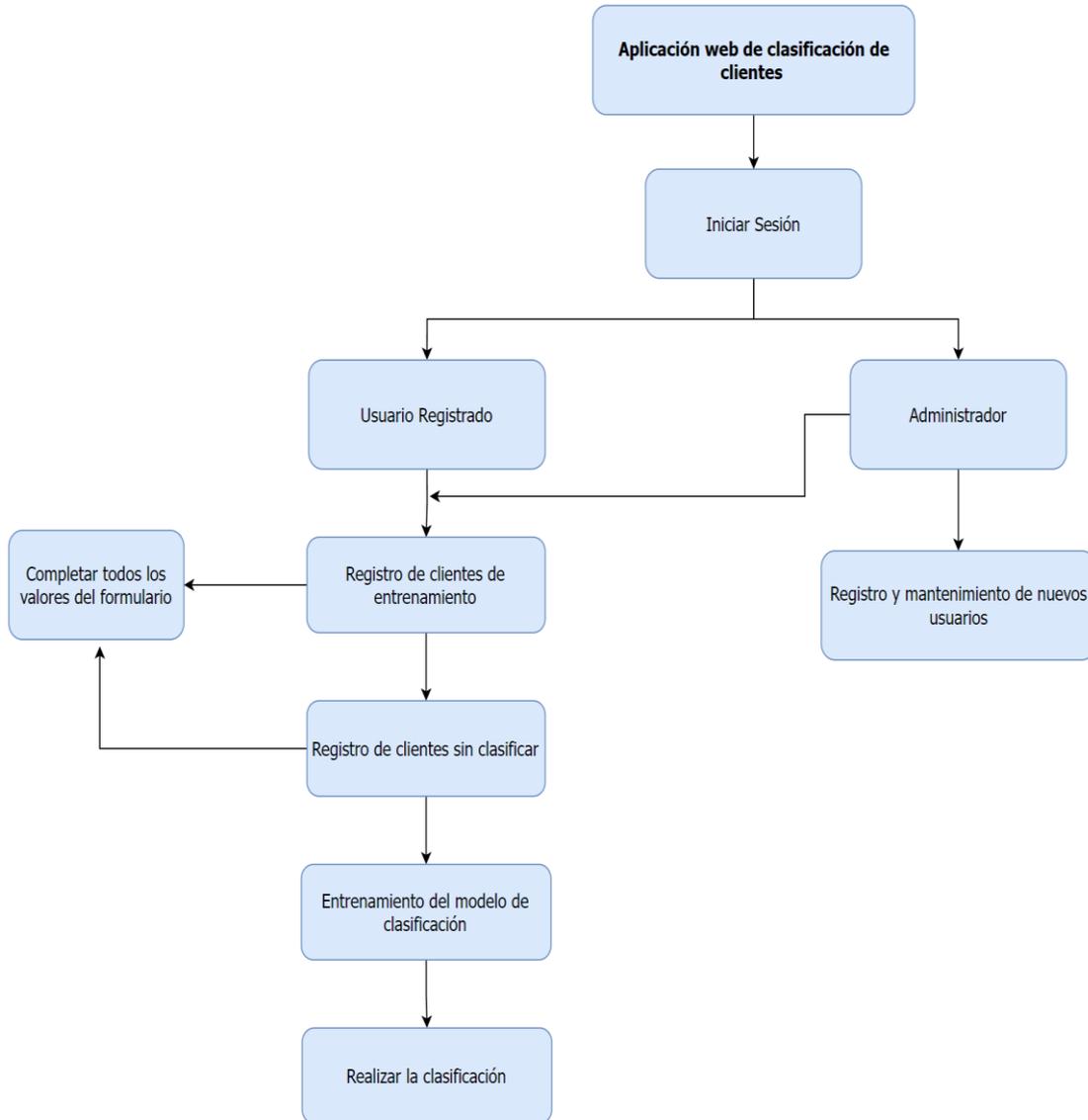
HU-2022-0011	Como usuario registrado, necesito registrar nuevos clientes de entrenamiento mediante un formulario, con la finalidad de agregar nuevos registros a la lista.	I14->Menú->I10->I11
HU-2022-0012	Como usuario registrado, necesito modificar clientes de entrenamiento, con la finalidad de hacer cambios a la información de los registros de la lista.	I14->Menú->I10->I12
HU-2022-0013	Como usuario registrado, necesito eliminar clientes de entrenamiento, con la finalidad de quitar registros con datos erróneos o duplicados a la información de los registros de la lista.	I14->Menú->I10->I13
HU-2022-0014	Como usuario registrado, necesito acceder al sistema mediante un inicio de sesión, con la finalidad de mantener segura la información del sistema.	I14
HU-2022-0015	Como usuario registrado necesito cerrar sesión , con la finalidad de mantener segura la información del sistema.	I14->I15
HU-2022-0016	Como usuario administrador, requiero visualizar las fechas de las clasificaciones realizadas, con la finalidad de tender el control sobre el registro de clasificaciones.	I14->Menú->I18->I16
HU-2022-0017	Como usuario administrador, requiero registrar nuevos usuarios del sistema, con la finalidad de dar acceso a las demás personas encargadas.	I14->Menú->I17
HU-2022-0018	Como usuario administrador, necesito entrenar el modelo de clasificación mediante un botón, con la finalidad de actualizar el modelo con nuevos registros.	I14->Menú->I18

**Tabla 32. Diccionario de datos**

**Fuente: Elaboración propia**

#### **4.10 Diagrama de navegación**

En la figura 33, se muestra el diagrama de navegación del prototipo de la aplicación web que representan la secuencia de navegación.



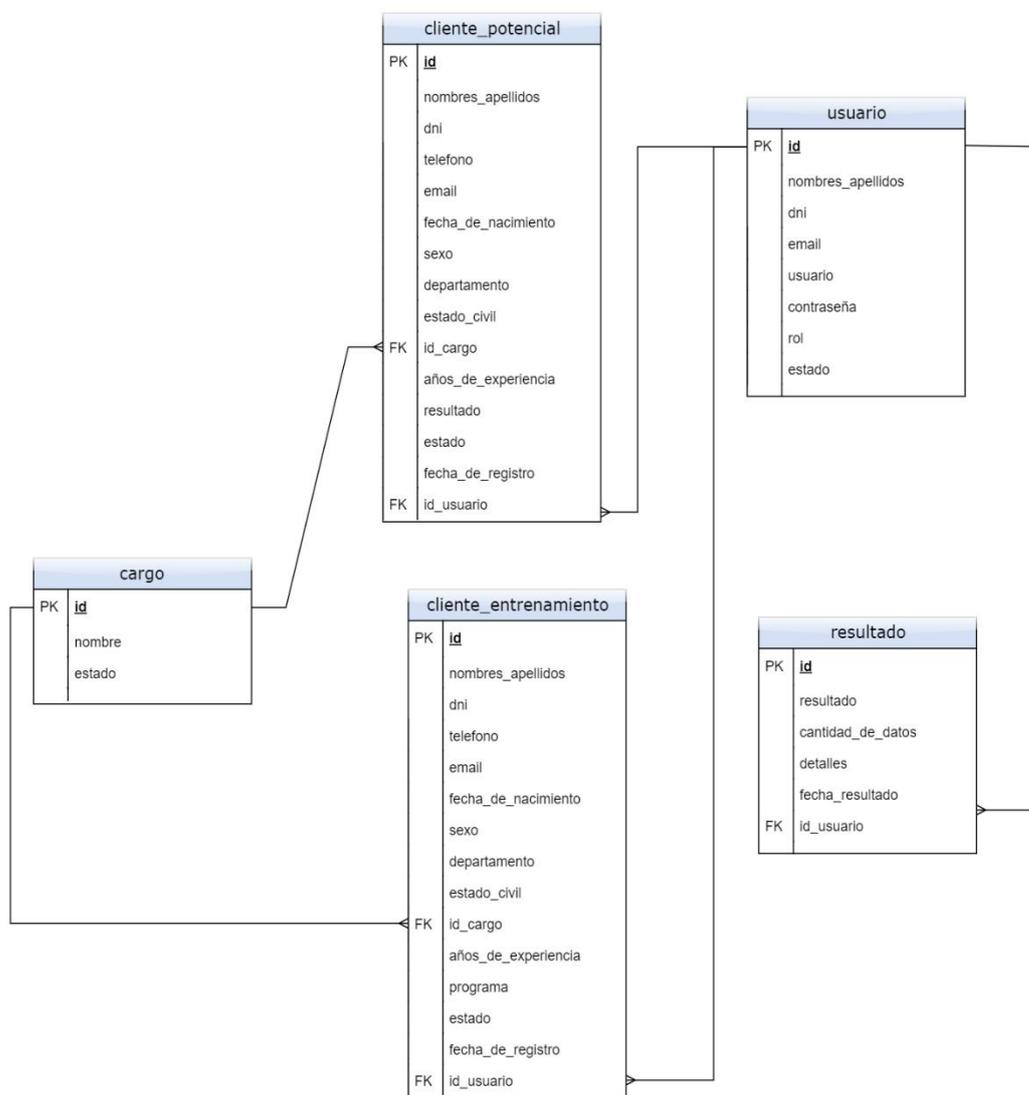
**Figura 33. Diagrama de navegación**

**Fuente: Elaboración propia**

#### 4.11 Diseño de base de datos

La propuesta desarrollada requirió una base de datos para hacer posible la clasificación; se requirió una tabla de clientes de entrenamiento donde se almacenan los registros históricos que serán necesarios para el entrenamiento del modelo de clasificación; una tabla de clientes potenciales, en los cuales se registrarán nuevos clientes que son clasificados con el modelo; una tabla de cargo, en la que se registrarán los cargos de los clientes; una tabla de resultado, en la cual se almacenarán los resultados obtenidos

de la clasificación; y una tabla de usuarios, en la cual se registrarán los usuarios con acceso al sistema, como se muestra en la figura 34.



**Figura 34. Diagrama de base de datos**

**Fuente: Elaboración propia**

#### 4.11.1 Diccionario de datos

En la tabla 33, se describe las tablas utilizadas en el proyecto, también se detallan las características de cada uno de los campos utilizados en las tablas.

Tabla: Usuario			
Campo	Tamaño	Tipo de Dato	Descripción
id	32	int	Id del usuario
nombres_apellidos	40	varchar()	Nombres del usuario

dni	8	varchar()	Documento de identidad del usuario
email	40	varchar()	Correo electrónico del usuario
usuario	15	varchar()	Usuario de ingreso al sistema
contraseña	15	varchar()	1. Contraseña de ingreso al sistema
rol	30	varchar()	Tipo de Rol del usuario
estado	32	int	Estado del cliente potencial

**Tabla: cliente\_entrenamiento**

<b>Campo</b>	<b>Tamaño</b>	<b>Tipo de Dato</b>	<b>Descripción</b>
id	32	int	Id del cliente de entrenamiento
nombres_apellidos	50	varchar()	Nombres y apellidos del cliente entrenamiento
dni	8	varchar()	Documento identidad del cliente entrenamiento
telefono	30	varchar()	Teléfono del cliente de entrenamiento
email	40	varchar()	Email del cliente de entrenamiento
fecha_de_nacimiento	---	date	Edad del cliente de entrenamiento
sexo	10	varchar()	Sexo del cliente de entrenamiento
departamento	20	varchar()	Departamento donde vive cliente entrenamiento
estado_civil	10	varchar()	Estado civil del cliente entrenamiento
id_cargo	32	int	Id del cargo

años experiencia	10	varchar()	Años de experiencia laboral del cliente entrenamiento
programa	40	varchar()	Programa del Cliente Entrenamiento
estado	32	int	Estado del Cliente Entrenamiento (activo e inactivo)
fecha_de_registro	--	date	Fecha de registro del cliente potencial
id_usuario	32	int	Id del usuario

**Tabla: cliente\_potencial**

<b>Campo</b>	<b>Tamaño</b>	<b>Tipo de Dato</b>	<b>Descripción</b>
id	32	int	Id del Cliente Potencial
nombres_apellidos	50	varchar()	Nombre , Apellido del Cliente potencial
dni	8	varchar()	Dni del Cliente Potencial
teléfono	30	varchar()	Número de teléfono del cliente potencial
email	40	varchar()	Correo electrónico del cliente potencial
fecha_de_nacimiento	---	date	Edad del cliente potencial
sexo	10	varchar()	Sexo del cliente potencial
departamento	20	varchar()	Departamento del cliente potencial
estado_civil	10	varchar()	Estado civil del cliente potencial
id_cargo	32	int	Id del cargo
años_de_experiencia	10	varchar()	Años de experiencia del cliente potencial
resultado	20	varchar()	Resultados del cliente potencial

estado	20	int ()	Estado del cliente potencial (activo e inactivo)
fecha_de_registro	--	date	Fecha de registro del cliente potencial
id_usuario	32	int	Id del usuario
<b>Tabla: cargo</b>			
<b>Campo</b>	<b>Tamaño</b>	<b>Tipo de Dato</b>	<b>Descripción</b>
id	32	int	Id del cargo
nombre	50	varchar()	Nombre del cargo
estado	20	varchar()	Estado del cargo
<b>Tabla: Resultado</b>			
<b>Campo</b>	<b>Tamaño</b>	<b>Tipo de Dato</b>	<b>Descripción</b>
id	32	int	Id de los Resultados
descripcion	200	varchar()	Descripción de los Resultados
detalles	200	varchar()	Detalles de los Resultados
cantidad_de_datos	---	text	Cantidad de datos de entrenamiento
fecha_resultado	---	date	Fecha del Resultado
id_usuario	32	int	Id del Usuario

**Tabla 33. Diccionario de datos**

Fuente: Elaboración propia

#### 4.11.2 Validación de base de datos básica

En la tabla 34, se muestra la asociación básica entre las historias de usuario y tablas, esto nos permite validar la existencia de un origen o destino de datos.

Identificador de historia (ID)	Enunciado de la historia	Tabla
HU-2022-0001	Como usuario registrado, requiero visualizar la lista de clientes potenciales, con la finalidad de tener un control de los clientes potenciales registrados.	cliente_potencial
HU-2022-0002	Como usuario registrado, requiero registrar nuevos clientes potenciales mediante un formulario, con la finalidad agregar nuevos registros a la lista	cliente_potencial, cargo
HU-2022-0003	Como usuario registrado, necesito modificar clientes potenciales, con la finalidad de hacer cambios a la información de los registros de la lista.	cliente_potencial
HU-2022-0004	Como usuario registrado, necesito eliminar clientes potenciales, con la finalidad de quitar registros con datos erróneos o duplicados a la información de los registros de la lista.	cliente_potencial
HU-2022-0005	Como usuario registrado, necesito clasificar los clientes potenciales, con la finalidad de conocer los intereses de los registros clasificados.	resultado
HU-2022-0006	Como usuario registrado, necesito visualizar las estadísticas de los clientes clasificados, con la finalidad de conocer la información resultante de la clasificación	resultado
HU-2022-0007	Como usuario administrador, requiero exportar la lista de los clientes clasificados, con la finalidad de proporcionar la información a las áreas que lo soliciten.	cliente_potencial
HU-2022-0008	Como usuario registrado, requiero filtrar la lista de los clientes clasificados por categoría, con la finalidad de visualizar la lista por cada categoría.	cliente_potencial

HU-2022-0009	Como usuario registrado, requiero importar clientes de entrenamiento, con la finalidad de utilizarlos para la clasificación.	cliente_entrenamiento
HU-2022-0010	Como usuario registrado, requiero visualizar la lista de clientes de entrenamiento, con la finalidad de tener un control de los clientes potenciales registrados.	cliente_entrenamiento
HU-2022-0011	Como usuario registrado, requiero registrar nuevos clientes de entrenamiento mediante un formulario, con la finalidad de agregar nuevos registros a la lista	cliente_entrenamiento ,cargo
HU-2022-0012	Como usuario registrado, necesito modificar clientes de entrenamiento, con la finalidad de hacer cambios a la información de los registros de la lista.	cliente_entrenamiento
HU-2022-0013	Como usuario registrado, necesito eliminar clientes de entrenamiento, con la finalidad de quitar registros con datos erróneos o duplicados a la información de los registros de la lista.	cliente_entrenamiento
HU-2022-0014	Como usuario registrado, necesito acceder al sistema mediante un inicio de sesión, con la finalidad de mantener segura la información del sistema.	usuario
HU-2022-0016	Como usuario administrador, requiero visualizar las fechas de las clasificaciones realizadas, con la finalidad de tender el control sobre el registro de clasificaciones.	resultado

**Tabla 34. Validación de la base de datos**

**Fuente: Elaboración propia**

## Capítulo V. CONSTRUCCIÓN

### 5.1 Construcción de la aplicación web

#### 5.1.1 Pantallas de la Aplicación Web

Historia de usuario 1: Como usuario registrado, requiero visualizar la lista de clientes potenciales, con la finalidad de tener un control de los clientes potenciales registrados.

SISTEMA DE CLASIFICACIÓN diego Cerrar Sesión

Usuarios

Clientes Entrenamiento

Clientes Potenciales

Clasificación

Reportes

Mantenimiento de clientes potenciales

[Nuevo cliente](#)

Mostrar 10 registros Buscar:

Nombre	Teléfono	Email	Edad	Sexo	Departamento	Estado civil	Cargo	Años exp		
ÁLVAREZ JARA VÍCTOR EDGARDO	992524708	Ayj_sac@hotmail.com	22	M	AREQUIPA	S	SUPERVISOR	1	<a href="#">Editar</a>	<a href="#">Eliminar</a>
ZUÑIGA HUANCA DANIEL RENZO	989597266	Danzunigahuanca@gmail.com	26	M	CUSCO	S	SUPERVISOR	17	<a href="#">Editar</a>	<a href="#">Eliminar</a>
ZEVALLS SOTO CONSTANCIO HECTOR	914677925	zevallos.hzs@gmail.com	40	M	AREQUIPA	C	SUPERVISOR	12	<a href="#">Editar</a>	<a href="#">Eliminar</a>
ZEVALLS CRUZ MABEL LIZ	962443584	belle_31_mabel@hotmail.com	51	M	AREQUIPA	S	SUPERVISOR	4	<a href="#">Editar</a>	<a href="#">Eliminar</a>
ZETA FLORES BERNARDINO	981644595	Zeta1986flores@gmail.com	18	M	AREQUIPA	S	SUPERVISOR	15	<a href="#">Editar</a>	<a href="#">Eliminar</a>
ZAVALETA SEBASTIAN LUIS FERNANDO	990380407	luis.zavaletasebastian@gmail.com	28	M	AREQUIPA	S	SUPERVISOR	6	<a href="#">Editar</a>	<a href="#">Eliminar</a>
ZARATE SALAS MIGUEL RENATO	948161020	Miguelrenatoic@outlook.com	40	M	AREQUIPA	S	SUPERVISOR	12	<a href="#">Editar</a>	<a href="#">Eliminar</a>
ZAPATA VIZCARRA JOMNER	928916870	jomnerxd312@gmail.com	20	M	CUSCO	C	SUPERVISOR	29	<a href="#">Editar</a>	<a href="#">Eliminar</a>
ZAPANA MORALES ROLANDO BRAULIO	944092980	zapanamoralessrolando@gmail.com	22	M	CUSCO	S	SUPERVISOR	14	<a href="#">Editar</a>	<a href="#">Eliminar</a>
ZAMUDIO YAUJRI FRANS CESAR	948552032	operador-fzy@outlook.es	44	M	AREQUIPA	S	SUPERVISOR	19	<a href="#">Editar</a>	<a href="#">Eliminar</a>

Figura 35. Interfaz de aplicación web Historia de usuario 1

Fuente: Elaboración propia

Historia de usuario 2: Como usuario registrado, requiero registrar nuevos clientes potenciales mediante un formulario, con la finalidad de agregar nuevos registros a la lista.

The image shows a web application interface with a dark blue sidebar on the left containing menu items: 'Usuarios', 'Clientes', 'Clasificación', and 'Entrenamiento'. The main content area is partially obscured by a modal window titled 'Nuevo cliente' with a close button (X) in the top right corner. The form contains the following fields and values:

- Nombre: ZUÑIGA DANIEL RENZO
- Teléfono: 923421323
- Email: daniel.zren@gmail.com
- Fecha de nacimiento: 11/04/1988 (with a calendar icon)
- Sexo:  M  F
- Departamento: Tacna (dropdown menu)
- Estado civil: Casado (dropdown menu)
- Cargo: Supervisor (dropdown menu)
- Años de experiencia: 5

At the bottom of the form is a blue button labeled 'Registrar cliente'. Below the modal, a portion of a table is visible with the following text: SEBASTIAN, LUIS, FERNANDO.

**Figura 36. Interfaz de aplicación web Historia de usuario 2**  
**Fuente: Elaboración propia**

Historia de usuario 3: Como usuario registrado, necesito modificar clientes potenciales, con la finalidad de hacer cambios a la información de los registros de la lista.

SISTEMA DE  
CLA

Modificar cliente

Nombre  
ZUÑIGA DANIEL RENZO

Teléfono  
923421323

Email  
daniel.zren@gmail.com

Fecha de nacimiento  
11/04/1988

Sexo  
 M  F

Departamento  
Tacna

Estado civil  
Casado

Cargo  
Supervisor

Años de experiencia  
5

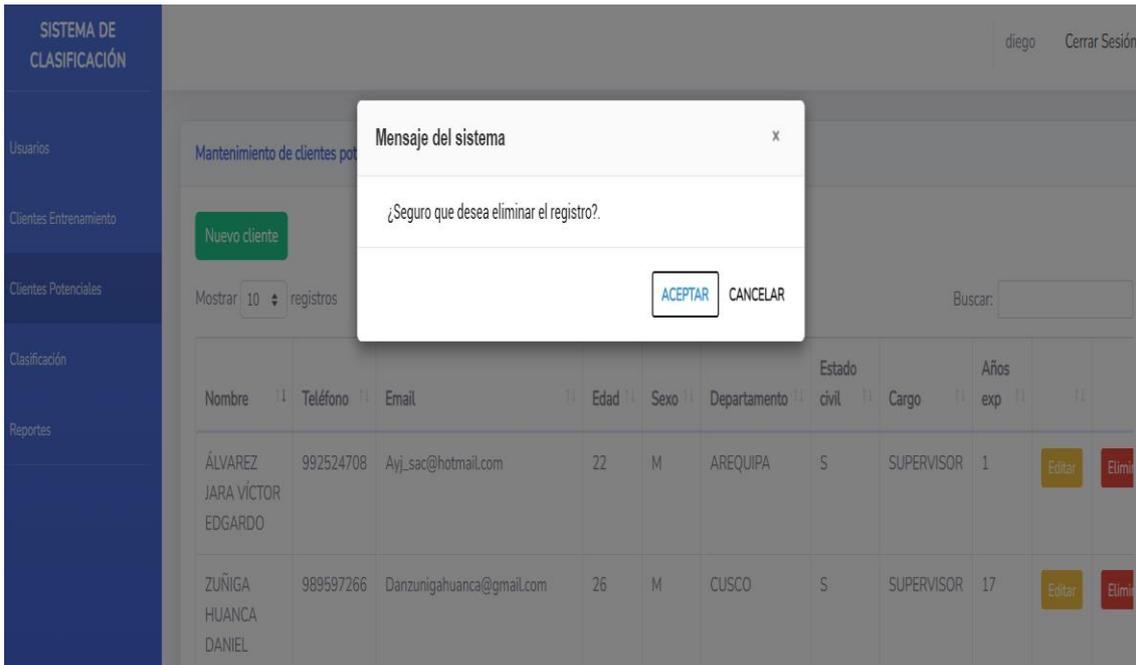
Actualizar registro

SEBASTIAN LUIS FERNANDO			
ZARATE	948161020	Miguelrenatoic@outlook.com	40

**Figura 37. Interfaz de aplicación web Historia de usuario 3**

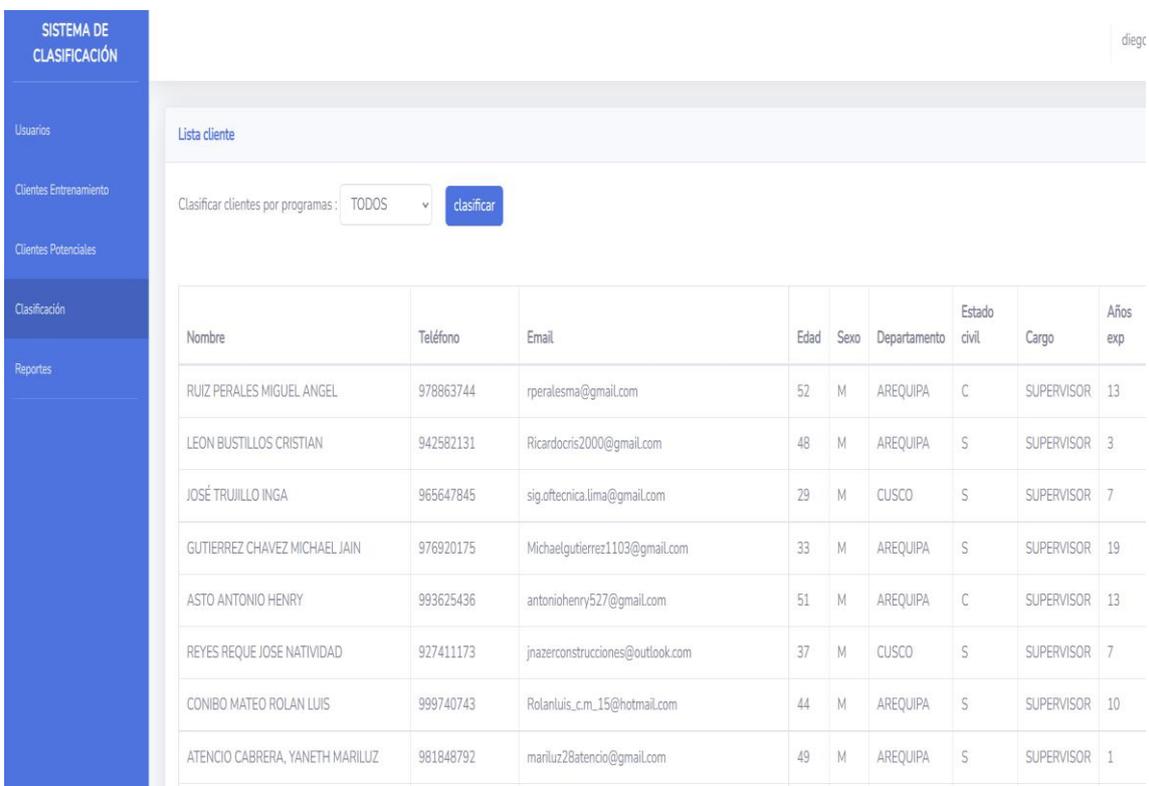
**Fuente: Elaboración propia**

Historia de usuario 4: Como usuario registrado, necesito eliminar clientes potenciales, con la finalidad de quitar registros con datos erróneos o duplicados a la información de los registros de la lista.



**Figura 38. Interfaz de aplicación web Historia de usuario 4**  
**Fuente: Elaboración propia**

Historia de usuario 5: Como usuario registrado, necesito clasificar los clientes potenciales, con la finalidad de conocer los intereses de los registros clasificados.



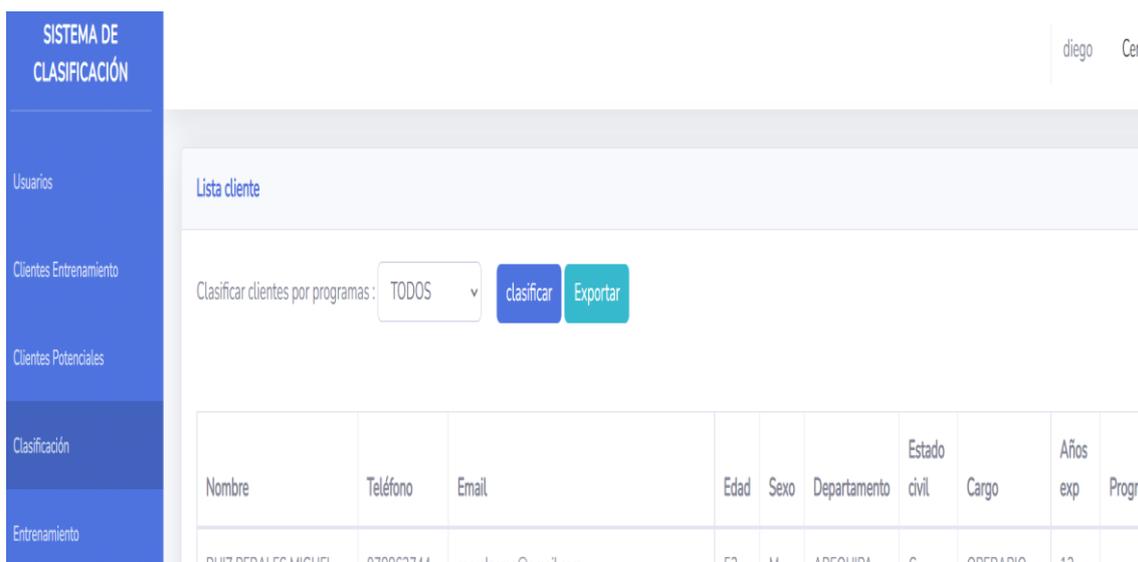
**Figura 39. Interfaz de aplicación web Historia de usuario 5**  
**Fuente: Elaboración propia**

Historia de usuario 6: Como usuario registrado, necesito visualizar las estadísticas de los clientes clasificados, con la finalidad de conocer la información resultante de la clasificación.



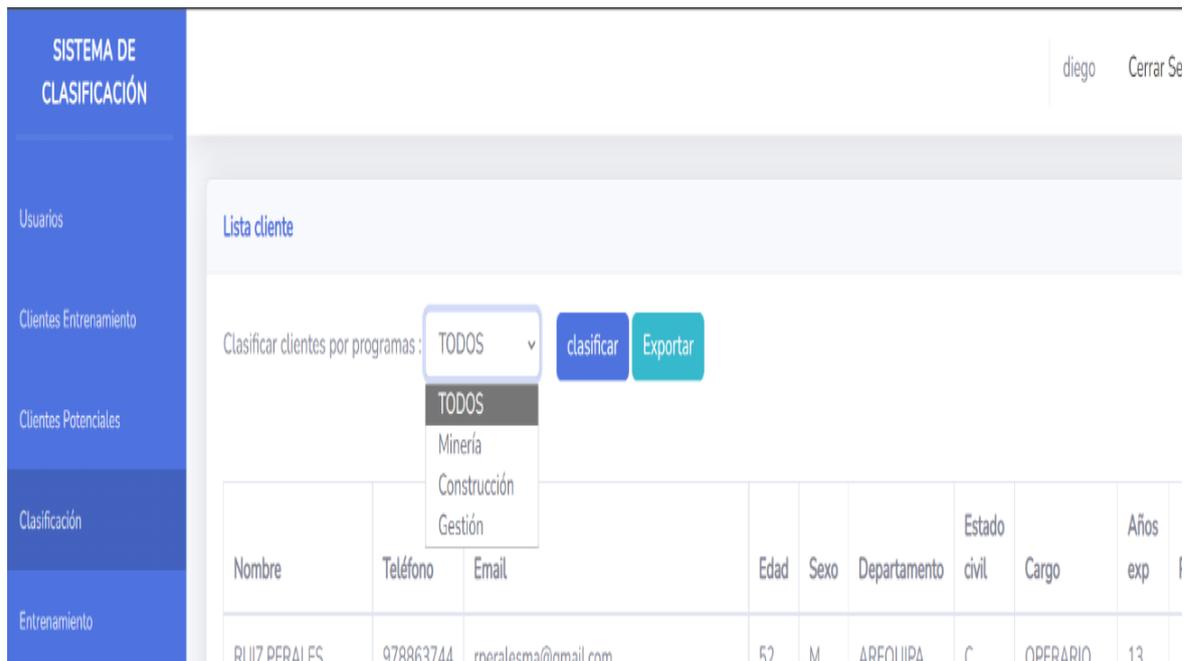
**Figura 40. Interfaz de aplicación web Historia de usuario 6**  
**Fuente: Elaboración propia**

Historia de usuario 7: Como usuario administrador, requiero exportar la lista de los clientes clasificados, con la finalidad de proporcionar la información a las áreas que lo soliciten.



**Figura 41. Interfaz de aplicación web Historia de usuario 7**  
**Fuente: Elaboración propia**

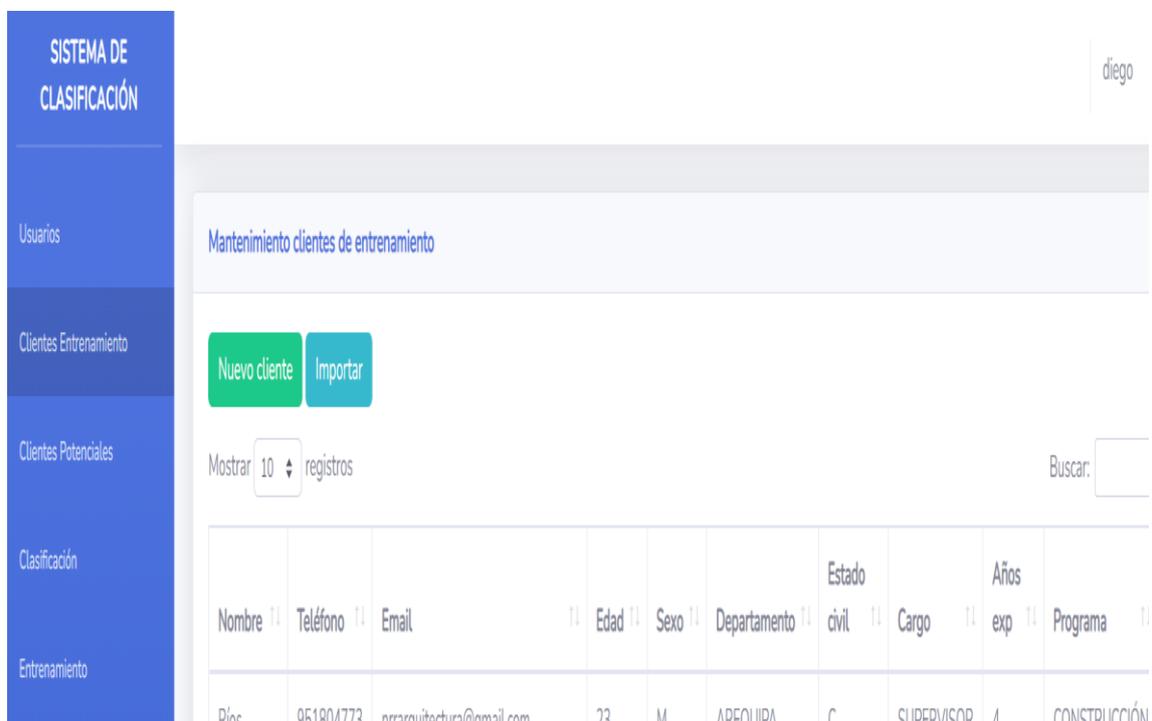
Historia de usuario 8: Como usuario registrado, requiero filtrar la lista de los clientes clasificados por categoría, con la finalidad de visualizar la lista por cada categoría.



**Figura 42. Interfaz de aplicación web Historia de usuario 8**

**Fuente: Elaboración propia**

Historia de usuario 9: Como usuario registrado, requiero importar clientes de entrenamiento, con la finalidad de utilizarlos para la clasificación.



**Figura 43. Interfaz de aplicación web Historia de usuario 9**

**Fuente: Elaboración propia**

Historia de usuario 10: Como usuario registrado, requiero visualizar la lista de clientes de entrenamiento, con la finalidad de tener un control de los clientes de entrenamiento registrados.

SISTEMA DE CLASIFICACIÓN diego Cerrar Sesión

Usuarios **Mantenimiento clientes de entrenamiento**

Clientes Entrenamiento **Nuevo cliente**

Clientes Potenciales

Clasificación

Entrenamiento

Mostrar 10 registros Buscar:

Nombre	Teléfono	Email	Edad	Sexo	Departamento	Estado civil	Cargo	Años exp	Programa		
Quispe Cabrera Roger Franklin	954074760	rogerfranklin.quispe89@gmail.com	48	M	CUSCO	S	SUPERVISOR	11	MINERÍA	Editar	Eliminar
Flores Bedon Eberth Ivan	937510352	eberthi07@gmail.com	47	M	AREQUIPA	S	OPERARIO	24	CONSTRUCCIÓN	Editar	Eliminar
Ñaupari torres Ronald Aldo	934148788	Nauparitorres131@gmail.com	47	M	CUSCO	S	OPERARIO	18	CONSTRUCCIÓN	Editar	Eliminar
Márquez Tucno Julinho	935499573	jmarquestucno0704@gmail.com	27	M	AREQUIPA	S	SUPERVISOR	6	MINERÍA	Editar	Eliminar
Mereglido Pedro Cervantes Ortiz	955453574	pedro.cervantes270476@gmail.com	49	M	AREQUIPA	S	SUPERVISOR	6	MINERÍA	Editar	Eliminar
Jimenez chavez Jesus angel	990485303	Angel_8510@hotmail.com	27	M	CUSCO	C	JEFE DE OBRA	22	CONSTRUCCIÓN	Editar	Eliminar
LLontop Bustamante Luis Enrique	953014974	lavin_angel@hotmail.com	19	M	AREQUIPA	S	ASISTENTE	29	GESTIÓN	Editar	Eliminar
RODRIGUEZ SORIA ANGEL MARADONA	925562432	Rodrimixangelito@gmail.com	40	M	CUSCO	S	SUPERVISOR	24	MINERÍA	Editar	Eliminar
Guzmán Palomino Edinson	951106206	guzmanpalominoedy@gmail.com	26	M	AREQUIPA	S	OPERARIO	6	CONSTRUCCIÓN	Editar	Eliminar
Jean Carlos Caballero Quispe	973250090	ing.jean.ccp@gmail.com	23	M	CUSCO	S	ASISTENTE	26	GESTIÓN	Editar	Eliminar

**Figura 44. Interfaz de aplicación web Historia de usuario 10**  
**Fuente: Elaboración propia**

Historia de usuario 11: Como usuario registrado, requiero registrar nuevos clientes de entrenamiento mediante un formulario, con la finalidad de agregar nuevos registros a la lista.

The image shows a modal window titled "Nuevo cliente de entrenamiento" (New training client) with a close button (X) in the top right corner. The form contains the following fields and values:

- Nombre: ÁLVAREZ VICTOR EDUARDO
- Teléfono: 924351239
- Email: victor.eal@gmail.com
- Fecha de nacimiento: 11/02/1982 (with a calendar icon)
- Sexo:  M  F
- Departamento: Arequipa (dropdown menu)
- Estado civil: Casado (dropdown menu)
- Cargo: Operario (dropdown menu)
- Años de experiencia: 4
- Programa: CONSTRUCCION (dropdown menu)

At the bottom of the form is a blue button labeled "Registrar cliente".

**Figura 45. Interfaz de aplicación web Historia de usuario 11**

**Fuente: Elaboración propia**

Historia de usuario 12: Como usuario registrado, necesito modificar clientes de entrenamiento, con la finalidad de hacer cambios a la información de los registros de la lista.

STEMA DE  
SIFICACIÓN

Modificar cliente de entrenamiento

Nombre  
ÁLVAREZ VICTOR EDUARDO

Teléfono  
924351239

Email  
victor.eal@gmail.com

Fecha de nacimiento  
11/02/1982

Sexo  
 M  F

Departamento  
Arequipa

Estado civil  
Casado

Cargo  
Operario

Años de experiencia  
4

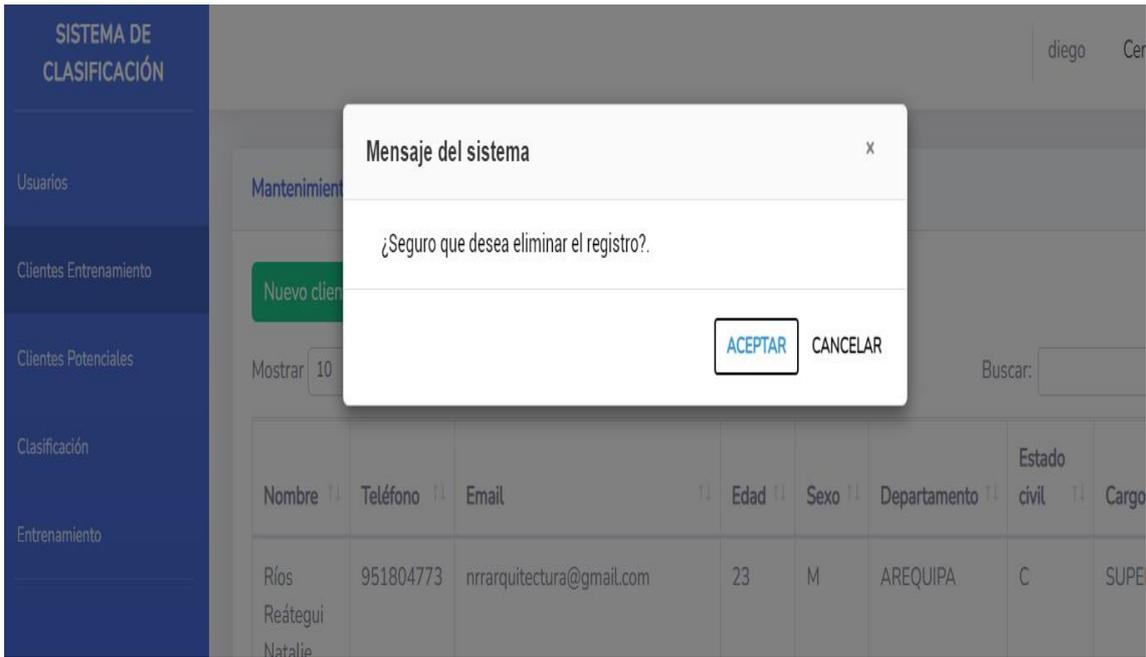
Programa  
CONSTRUCCION

Actualizar registro

**Figura 46. Interfaz de aplicación web Historia de usuario 12**

**Fuente: Elaboración propia**

Historia de usuario 13: Como usuario registrado, necesito eliminar clientes de entrenamiento, con la finalidad de quitar registros con datos erróneos o duplicados a la información de los registros de la lista.



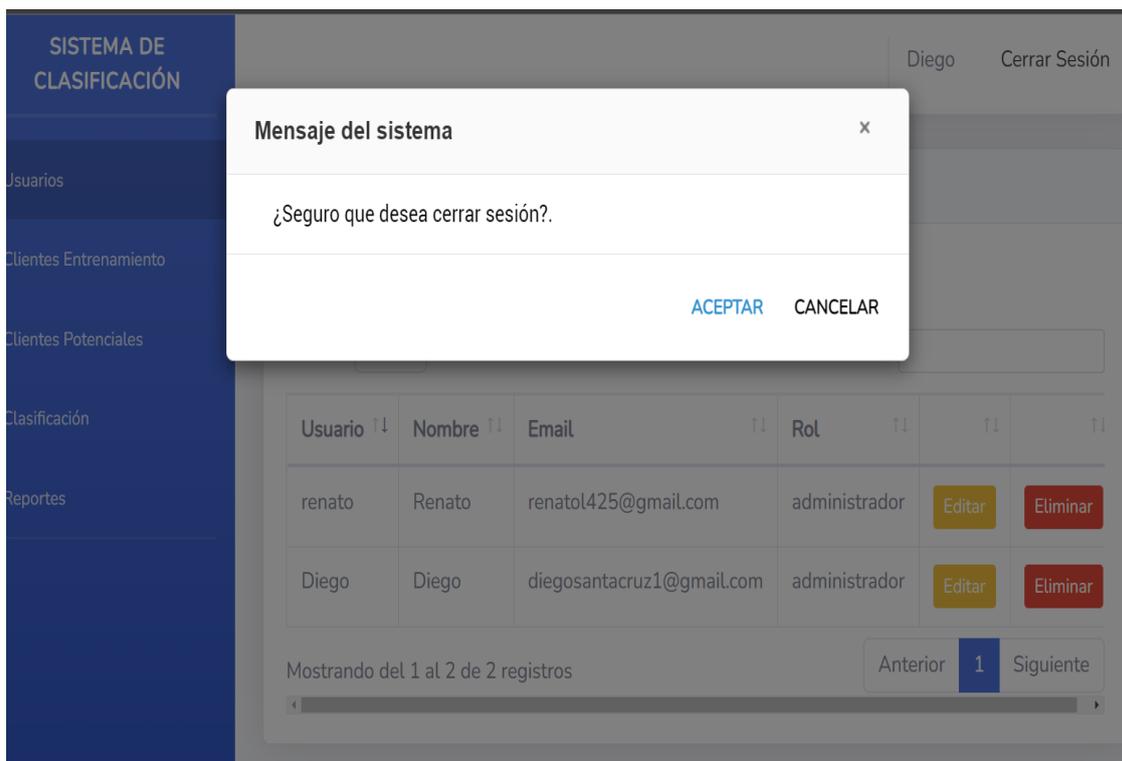
**Figura 47. Interfaz de aplicación web Historia de usuario 13**  
**Fuente: Elaboración propia**

Historia de usuario 14: Como usuario registrado, necesito acceder al sistema mediante un inicio de sesión, con la finalidad de mantener segura la información del sistema.



**Figura 48. Interfaz de aplicación web Historia de usuario 14**  
**Fuente: Elaboración propia**

Historia de usuario 15: Como usuario registrado necesito cerrar sesión, con la finalidad de mantener segura la información del sistema.



**Figura 49. Interfaz de aplicación web Historia de usuario 15**  
**Fuente: Elaboración propia**

Historia de usuario 16: Como usuario administrador, requiero visualizar las fechas de las clasificaciones realizadas, con la finalidad de tender el control sobre el registro de clasificaciones.



**Figura 50. Interfaz de aplicación web Historia de usuario 16**  
**Fuente: Elaboración propia**

Historia de usuario 17: Como usuario administrador, requiero registrar nuevos usuarios del sistema, con la finalidad de dar acceso a las demás personas encargadas.

SISTEMA DE CLASIFICACIÓN

Usuarios

Cientes Entrenamiento

Cientes Potenciales

Clasificación

Entrenamiento

Nuevo Usuario

Nombre

Renato Zapana

Usuario

renato

Contraseña

.....

Email

renatzll@gmail.com

Rol

Administrador

Registrar usuario

Figura 51. Interfaz de aplicación web Historia de usuario 17

Fuente: Elaboración propia

Historia de usuario 18: Como usuario administrador, necesito entrenar el modelo de clasificación mediante un botón, con la finalidad de actualizar el modelo con nuevos registros.

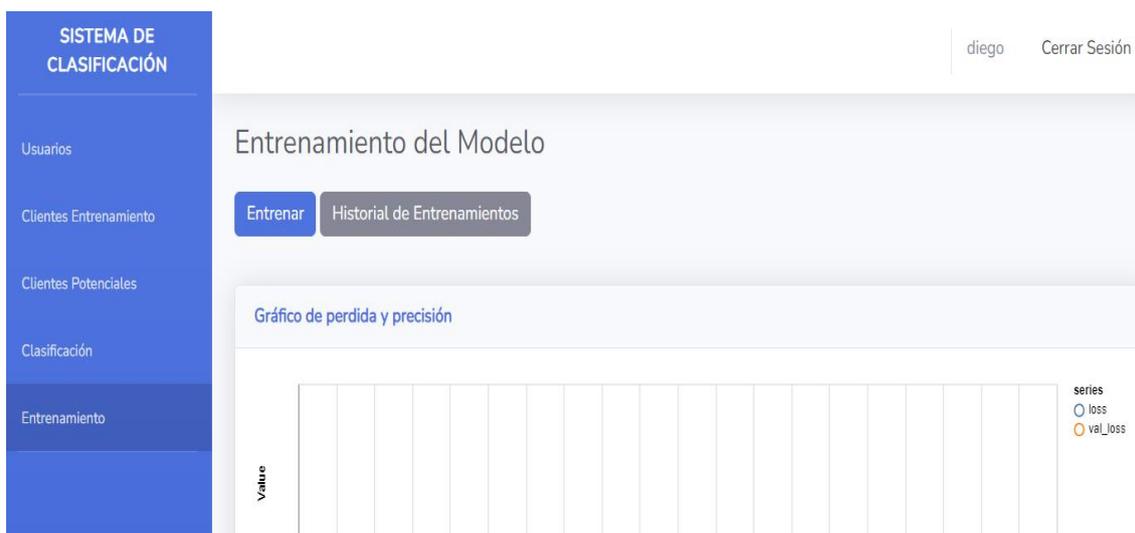


Figura 52. Interfaz de aplicación web Historia de usuario 18

Fuente: Elaboración propia

### 5.1.2 Proceso de clasificación en la aplicación web

Para la construcción de la aplicación web y la vinculación con el modelo de clasificación, se elaboró una función que permitirá almacenar los datos de los clientes de entrenamiento, los cuales servirán para el entrenamiento del modelo, como se muestra en la figura 53.

```
public function store(Request $request)
{
    $nuevo_cliente = Cliente::create([
        'nombre' => $request->nombre,
        'telefono' => $request->telefono,
        'email' => $request->email,
        'fecha_de_nacimiento' => $request->fecha_de_nacimiento,
        'sexo' => $request->sexo,
        'departamento' => $request->departamento,i
        'estado_civil' => $request->estado_civil,
        'cargo' => $request-> $cargo,
        'años_experiencia' => $request->años_experiencia,
        'programa' => $request->programa,
    ]);

    return $nuevo_cliente;
}
...
```

**Figura 53. Registro de clientes a la base de datos**

**Fuente: Elaboración propia**

Teniendo los datos de los clientes de entrenamiento almacenados en la base de datos del sistema, se realiza una función con el objetivo de obtener todos los datos de los registros de entrenamiento almacenados, como se muestra en la figura 54.

```
//Obtener todos los clientes de entrenamientos activos
public function getAll(){
    return Entrenamiento::orderBy('id','DESC')->get();
}
```

**Figura 54. Obtención de los registros de la base de datos**

**Fuente: Elaboración propia**

Los datos de clientes de entrenamiento que se encuentran en la base de datos son expuestos por la aplicación web mediante rutas, por las cuales se podrán solicitar todos los registros, como se muestra en la figura 55.

```
//Rutas para la obtención de los clientes
Route::get('/datos_entrenamiento','ClienteController@getClientesEntrenamiento');
Route::get('/datos_validacion','ClienteController@getClientesPotenciales');
```

**Figura 55. Rutas web para la obtención de datos**

**Fuente: Elaboración propia**

Los datos obtenidos mediante las rutas definidas por el sistema web son almacenados en variables, las cuales se usarán más adelante para el tratamiento de los datos, como se muestra en la figura 56.

```
//Obtección de datos de entrenamiento y validación
fetchJSONFile('/datos_entrenamiento', function(data){
    clientes_entrenamiento = data;
}),
fetchJSONFile('/datos_validacion', function(data){
    clientes_validacion = data;
})
```

**Figura 56. Obtención de datos de clientes**

**Fuente: Elaboración propia**

Los datos obtenidos son normalizados para que puedan ser usados en el modelo de clasificación, también se descarta la información irrelevante para el entrenamiento, como se muestra en la figura 57.

```
datos = datos.map(col=>
    {
        return{
            edad : normalize(col.fecha_de_nacimiento),
            sexo : normalize(col.sexo),
            departamento: normalize(col.departamento),
            estado_civil:normalize(col.estado_civil),
            cargo:normalize(col.cargo),
            años_exp:normalize(col.años_exp),
            programa:normalize(col.programa)
        }
    });
```

**Figura 57. Selección de variables para el modelo**

**Fuente: Elaboración propia**

## 5.2 Construcción del modelo de clasificación

### 5.2.1 Procesamiento de datos

Para la selección de las características de la construcción del modelo, se examinaron las particularidades de los datos de los clientes que se muestran en el Anexo 4, con el fin de seleccionar la información relevante para el modelo de clasificación.

#### 5.2.1.1 Limpieza de datos

En ese proceso se excluyeron características de datos sin valor para el modelo, los cuales son: nombres, DNI, teléfono y email; además, se verificó la duplicidad de los datos eliminando los registros duplicados y filtrando los registros únicos por DNI.

### 5.2.2 Selección del modelo de clasificación

Para la selección del modelo de clasificación se evaluaron cuatro algoritmos de Aprendizaje Automático, los cuales fueron: Regresión logística, Máquinas de vectores de soporte, Redes neuronales y K vecinos más próximos, con el fin de descubrir cuál de ellos obtiene mejor exactitud al clasificar los datos.

Como se puede observar en la figura 58, se evidencian los valores de exactitud en los modelos de clasificación evaluados: un 77,40% utilizando el algoritmo de Regresión logística, un 85,40% con Maquinas de vectores de soporte, un 88,42% con redes neuronales, y 78,83%, con k vecinos más próximos; de esta manera podemos considerar el algoritmo de redes neuronales para la construcción de nuestro modelo de clasificación, por ser el más exacto entre los modelos evaluados.

```
datos = pd.read_csv("datos_clientes.csv")
df=pd.DataFrame(datos,columns=['edad','sexo','departamento','estado_civil','cargo','años_exp','programa'])
#Selección de Las columnas utilizadas
X=df[['edad','sexo','departamento','estado_civil','cargo','años_exp']]
#Selección de la columna utilizada para determinar la clase perteneciente
y=df['programa']
#Declaración de variables para el entrenamiento y el tamaño de muestra utilizada como validación
X_train, X_test, y_train, y_test = train_test_split(X,y,test_size=0.1)
#Creación de Los algoritmos de clasificación utilizando valores por defecto
algoritmoLR = LogisticRegression(solver='lbfgs')
algoritmoLR.fit(X_train,y_train)
algoritmoSVC = SVC()
algoritmoSVC.fit(X_train,y_train)
algoritmoKN = KNeighborsClassifier()
algoritmoKN.fit(X_train,y_train)
RedNeuronal = MLPClassifier(hidden_layer_sizes=(1))
RedNeuronal.fit(X_train,y_train)
RedNeuronal.predict(X_test)

print("Exactitud : ")
print("Máquina de vectores de soporte :",algoritmoSVC.score(X_train, y_train))
print("K Vecinos más proximos :",algoritmoKN.score(X_train, y_train))
print("Red Neuronal:",RedNeuronal.score(X_train, y_train))
print("Regresión Logística :",algoritmoLR.score(X_train, y_train))

Exactitud :
Máquina de vectores de soporte : 0.8540740740740741
K Vecinos más proximos : 0.7883333333333333
Red Neuronal : 0.8842592592592593
Regresión Logística : 0.7740740740740741
```

Figura 58. Evaluación de los algoritmos de clasificación

Fuente: Elaboración propia

### 5.2.3 Diseño del Modelo de la red neuronal

Para el modelo de clasificación de redes neuronales se tomaron las variables obtenidas de los clientes las cuales son: edad, sexo, departamento, estado civil, cargo y años de experiencia, variables utilizadas como parámetros para la capa de entrada.

Existen diversas maneras para determinar el número de neuronas en la capa oculta, la sugerida es que “el número de neuronas en la capa oculta debería ser menor a dos veces la cantidad de nodos de la capa de entrada” (39).

Se realizaron pruebas rápidas de la precisión obtenida, usando un número de neuronas menores a dos veces la cantidad de nodos de la capa de entrada.

Nro. de neuronas	Precisión obtenida
1	0.7157
2	0.8422
3	0.8318
4	0.8518
5	0.7918
6	0.8569
7	0.8719
8	0.8945
9	0.9012
10	0.9074
11	0.8994

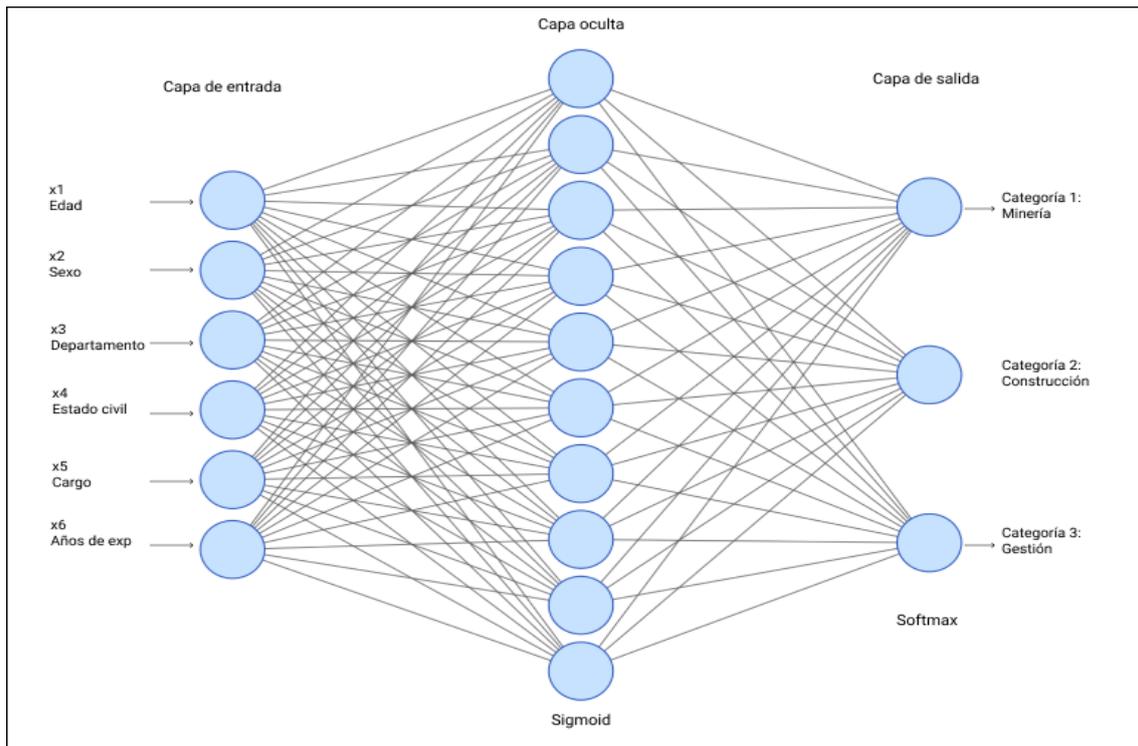
**Tabla 35. Selección del número de neuronas en la capa oculta**

**Fuente: Elaboración propia**

Basándonos en los resultados de las pruebas rápidas de la tabla 35, se seleccionaron 10 neuronas en la capa oculta para el modelo, debido a que obtuvo un mejor resultado. Según (39) con una capa oculta es suficiente para solucionar la mayoría de problemas de clasificación y regresión; a razón de que dos capas a más se utilizan para problemas de Deep Learning que requieren diferentes procesos en cada capa.

La función de activación utilizada es Sigmoid, se usa para problemas de clasificación donde se realiza el cálculo y se establecen los valores que resultan en tres nodos; además, cuenta la capa de salida que utiliza la función Softmax, para normalizar los tres valores resultantes y que estos sumen uno; los resultados después del entrenamiento del modelo tendrán como resultado un tensor con tres valores entre cero y uno, en el que el valor mayor corresponderá al resultado más probable de la clasificación.

En la figura 59, se muestra de manera gráfica los nodos utilizados en la capa de entrada, la capa oculta y la capa de salida del modelo de clasificación.



**Figura 59. Diseño del modelo de clasificación**

Fuente: Elaboración propia

#### 5.2.4 Construcción del modelo de la red neuronal

Para la construcción del modelo se establecieron 250 épocas, se usó el 10% del conjunto de datos de entrenamiento de 6000 registros para el proceso de validación, una tasa de aprendizaje del 0.001 que es lo recomendado para el optimizador usado, y se estableció el mezclado del conjunto de datos, como se muestra en la figura 60.

```

174 | //épocas
175 | epochs:250,
176 | //división de validación
177 | validationSplit:0.10,
178 | //tasa de aprendizaje
179 | learningRate : 0.001,
180 | //mezclar
181 | shuffle: true,

```

**Figura 60. Parámetros principales**

Fuente: Elaboración propia

Después de definir los parámetros, se procede a definir la creación de un modelo de tipo secuencial que establece que cada capa estará conectada a la siguiente de manera secuencial, como se muestra en la figura 61.

```
139 // Creación del modelo
140 const model = tf.sequential();
```

**Figura 61. Creación del modelo**

**Fuente: Elaboración propia**

Establecemos una capa oculta que recibe los parámetros de los seis nodos de entrada de la capa anterior, se establecen diez nodos y la función de activación Sigmoid que es utilizada para la clasificación; y cuenta con valores entre 0 y 1, como se muestra en la figura 62.

```
143 model.add(tf.layers.dense(
144     {   inputShape: 6,
145         activation: 'sigmoid',
146         units: 10
147     }
148 ));
```

**Figura 62. Definición de capa oculta y función de activación**

**Fuente: Elaboración propia**

Establecemos la capa de salida que recibe los parámetros de los diez nodos de la capa anterior, se establecen tres nodos de salida que corresponden a las tres categorías que se desea categorizar; a su vez se establece la función de activación Softmax que sirve para normalizar los tres resultados que sumados resultan uno, como se muestra en la figura 63.

```
150 model.add(tf.layers.dense(
151     {
152         inputShape: 10,
153         units: 3,
154         activation: 'softmax'
155     }
156 ));
```

**Figura 63. Definición de la capa de salida**

**Fuente: Elaboración propia**

Definimos la función de pérdida utilizada para la clasificación y para el optimizador Adam en la figura 64, los cuales nos permitirán conocer los valores de pérdida en cada época de entrenamiento, siendo los valores que tienden a cero más óptimos para el modelo.

```

161 |     loss: "categoricalCrossentropy",
162 |     optimizer: tf.train.adam(),

```

**Figura 64. Definición de la función de pérdida**  
**Fuente: Elaboración propia**

Definimos las variables de los clientes utilizados para el entrenamiento del modelo, los cuales son: edad, sexo, departamento, estado civil, cargo y años de experiencia; también definimos las tres categorías resultantes como salida del modelo, las cuales tendrán valores entre cero y uno, como se muestra en figura 65.

```

101 | // datos de entrenamiento
102 | const Datos_entrenamiento = tf.tensor2d(clientes.map(item=> [
103 |   item.edad, item.sexo, item.departamento, item.estado_civil,item.cargo,item.años_exp
104 | ]
105 | ),[clientes.length,6])
115 | // datos salidas de entrenamiento
116 | const Salidas_entrenamiento = tf.tensor2d(clientes.map(item => [
117 |   item.programa === 'minería' ? 1 : 0,
118 |   item.programa === 'construcción' ? 1 : 0,
119 |   item.programa === 'gestión' ? 1 : 0
120 | ]), [clientes.length,3])
173 | await model.fit(Datos_entrenamiento, Salidas_entrenamiento

```

**Figura 65. Parámetros de entrenamiento**  
**Fuente: Elaboración propia**

Comenzamos a entrenar el modelo, cuyos valores resultantes correspondientes a Accuracy expresan su exactitud en porcentaje, tomando valores entre 0 y 1, donde 1 corresponde a una exactitud del 100% y 0 a una exactitud del 0%; los valores resultantes de Loss expresan la pérdida o error obtenido, donde los valores que son más cercanos a 0 indican una tasa de error más óptima.

Durante el entrenamiento del modelo, el Accuracy o la exactitud comienzan con un valor de 0.3035 que aumenta al 0.65; mientras que el valor de pérdida inicia con un valor de 1.1765 y disminuye a 0.8156 en la época 7, como se muestra en la figura 66.

1	{val_loss: 1.1649079322814941, val_acc: 0.2083333283662796, loss: 1.1764700412750244, acc: 0.30351850390434265}
2	{val_loss: 1.1061203479766846, val_acc: 0.27166664600372314, loss: 1.0108678340911865, acc: 0.48629629611968994}
3	{val_loss: 1.0715311765670776, val_acc: 0.4449999928474426, loss: 0.9466447830200195, acc: 0.6133333444595337}
4	{val_loss: 1.0598138570785522, val_acc: 0.4583333134651184, loss: 0.9017797708511353, acc: 0.6424074172973633}
5	{val_loss: 1.0539131164550781, val_acc: 0.47333329916000366, loss: 0.863782525062561, acc: 0.6538888812065125}
6	{val_loss: 1.045447587966919, val_acc: 0.4699999690055847, loss: 0.8358423709869385, acc: 0.6587036848068237}
7	{val_loss: 1.03036630153656, val_acc: 0.4816666543483734, loss: 0.8156617283821106, acc: 0.6568518280982971}

**Figura 66. Época 1 al 7 de modelo de clasificación**  
**Fuente: Elaboración propia**

Entre la época 40 a la 50, la exactitud logra aumentar del 0.7409 a 0.7709 y la pérdida disminuye de 0.5820 a 0.5433, lo cual indica que el entrenamiento mejora el modelo en su exactitud y pérdida, como se muestra en la figura 67.

40	{val_loss: 0.7961792349815369, val_acc: 0.6649999618530273, loss: 0.5820228457450867, acc: 0.7409259080886841}
41	{val_loss: 0.7894876003265381, val_acc: 0.6649999618530273, loss: 0.577198326587677, acc: 0.742222249507904}
42	{val_loss: 0.7824771404266357, val_acc: 0.6649999618530273, loss: 0.5729440450668335, acc: 0.7462962865829468}
43	{val_loss: 0.7989624738693237, val_acc: 0.6716666221618652, loss: 0.5687954425811768, acc: 0.7470370531082153}
44	{val_loss: 0.770180881023407, val_acc: 0.6666666269302368, loss: 0.565529465675354, acc: 0.7535185217857361}
45	{val_loss: 0.7697482109069824, val_acc: 0.6649999618530273, loss: 0.5611283779144287, acc: 0.7540740966796875}
46	{val_loss: 0.7664527893066406, val_acc: 0.6649999618530273, loss: 0.5570082068443298, acc: 0.762592613697052}
47	{val_loss: 0.7562165260314941, val_acc: 0.6733332872390747, loss: 0.5538886785507202, acc: 0.7638888955116272}
48	{val_loss: 0.7629719972610474, val_acc: 0.6716666221618652, loss: 0.5503528118133545, acc: 0.7675926089286804}
49	{val_loss: 0.7533764243125916, val_acc: 0.6683332920074463, loss: 0.5466959476470947, acc: 0.7688888907432556}
50	{val_loss: 0.7432652115821838, val_acc: 0.6716666221618652, loss: 0.5433539152145386, acc: 0.7709259390830994}

**Figura 67. Época 40 a la 50 de modelo de clasificación**  
**Fuente: Elaboración propia**

Entre la época 100 a la 110, la exactitud logra aumentar del 0.8168 a 0.8248 y la pérdida disminuye de 0.4425 a 0.4272, lo cual indica que el entrenamiento mejora el modelo en su exactitud y pérdida, como se muestra en la figura 68.

100	▶ {val_loss: 0.6593796610832214, val_acc: 0.6716666221618652, loss: 0.44258472323417664, acc: 0.8168518543243408}
101	▶ {val_loss: 0.6622201800346375, val_acc: 0.6799999475479126, loss: 0.441461443901062, acc: 0.8190740942955017}
102	▶ {val_loss: 0.6740028858184814, val_acc: 0.68833327293396, loss: 0.4395568072795868, acc: 0.817037045955658}
103	▶ {val_loss: 0.6569051742553711, val_acc: 0.6799999475479126, loss: 0.43798959255218506, acc: 0.817037045955658}
104	▶ {val_loss: 0.6670365929603577, val_acc: 0.68833327293396, loss: 0.43645092844963074, acc: 0.8192592859268188}
105	▶ {val_loss: 0.6616788506507874, val_acc: 0.691666626930237, loss: 0.4348355531692505, acc: 0.8185185194015503}
106	▶ {val_loss: 0.6502805352210999, val_acc: 0.6749999523162842, loss: 0.4335159361362457, acc: 0.820555567741394}
107	▶ {val_loss: 0.684890866279602, val_acc: 0.691666626930237, loss: 0.4317071735858917, acc: 0.8198148012161255}
108	▶ {val_loss: 0.6718714237213135, val_acc: 0.691666626930237, loss: 0.4303228557109833, acc: 0.8244444131851196}
109	▶ {val_loss: 0.6590127348899841, val_acc: 0.691666626930237, loss: 0.429372102022171, acc: 0.8251851797103882}
110	▶ {val_loss: 0.6678299307823181, val_acc: 0.68833327293396, loss: 0.42722660303115845, acc: 0.8248147964477539}

**Figura 68. Época 100 a la 110 de modelo de clasificación**

**Fuente: Elaboración propia**

Entre la época 140 a la 150, la exactitud logra aumentar del 0.8703 a 0.8751 y la pérdida disminuye de 0.3792 a 0.3608, lo cual indica que el entrenamiento mejora el modelo en su exactitud y pérdida, como se muestra en la figura 69.

140	▶ {val_loss: 0.6716602444648743, val_acc: 0.6933333277702332, loss: 0.3792484700679779, acc: 0.8703703880310059}
141	▶ {val_loss: 0.6606472134590149, val_acc: 0.6966666579246521, loss: 0.37751713395118713, acc: 0.871666669845581}
142	▶ {val_loss: 0.6739771962165833, val_acc: 0.691666626930237, loss: 0.37565115094184875, acc: 0.87111110949516296}
143	▶ {val_loss: 0.694830060005188, val_acc: 0.6699999570846558, loss: 0.3737725615501404, acc: 0.8703703880310059}
144	▶ {val_loss: 0.6788262128829956, val_acc: 0.691666626930237, loss: 0.372193843126297, acc: 0.8729629516601562}
145	▶ {val_loss: 0.6830242276191711, val_acc: 0.684999942779541, loss: 0.3701310455799103, acc: 0.8742592334747314}
146	▶ {val_loss: 0.6740808486938477, val_acc: 0.689999976158142, loss: 0.36817795038223267, acc: 0.8737037181854248}
147	▶ {val_loss: 0.684646487236023, val_acc: 0.694999928474426, loss: 0.3664405047893524, acc: 0.8746296167373657}
148	▶ {val_loss: 0.6896522641181946, val_acc: 0.699999988079071, loss: 0.36414530873298645, acc: 0.872592568397522}
149	▶ {val_loss: 0.6846791505813599, val_acc: 0.6966666579246521, loss: 0.36237066984176636, acc: 0.8735185265541077}
150	▶ {val_loss: 0.6855882406234741, val_acc: 0.694999928474426, loss: 0.36083006858825684, acc: 0.8751851916313171}

**Figura 69. Época 140 al 150 de modelo de clasificación**

**Fuente: Elaboración propia**

Entre la época 190 a la 200, la exactitud logra aumentar del 0.8981 a 0.8975 y la pérdida disminuye de 0.2953 a 0.2829, lo cual indica que el entrenamiento mejora el modelo en su exactitud y pérdida, pero de manera mínima, como se muestra en la figura 70.

190 ▶ {val\_loss: 0.688384473323822, val\_acc: 0.7016666531562805, loss: 0.2953701317310333, acc: 0.8981481194496155}

191 ▶ {val\_loss: 0.6797755360603333, val\_acc: 0.70333331823349, loss: 0.29424697160720825, acc: 0.8987036943435669}

192 ▶ {val\_loss: 0.6734408736228943, val\_acc: 0.70333331823349, loss: 0.29288357496261597, acc: 0.8994444608688354}

193 ▶ {val\_loss: 0.6752996444702148, val\_acc: 0.70333331823349, loss: 0.29181551933288574, acc: 0.8987036943435669}

194 ▶ {val\_loss: 0.681756854057312, val\_acc: 0.7016666531562805, loss: 0.29036393761634827, acc: 0.8979629874229431}

195 ▶ {val\_loss: 0.6818003058433533, val\_acc: 0.7016666531562805, loss: 0.28922080993652344, acc: 0.8983333110809326}

196 ▶ {val\_loss: 0.7007868885993958, val\_acc: 0.699999988079071, loss: 0.28812775015830994, acc: 0.8981481194496155}

197 ▶ {val\_loss: 0.6940481066703796, val\_acc: 0.699999988079071, loss: 0.2872082591056824, acc: 0.8959259390830994}

198 ▶ {val\_loss: 0.6748968362808228, val\_acc: 0.7066666483879089, loss: 0.28549724817276, acc: 0.8949999809265137}

199 ▶ {val\_loss: 0.6892614960670471, val\_acc: 0.7049999833106995, loss: 0.28459036350250244, acc: 0.8974074125289917}

200 ▶ {val\_loss: 0.706031084060669, val\_acc: 0.7049999833106995, loss: 0.28299379348754883, acc: 0.8975926041603088}

**Figura 70. Época 190 a la 200 de modelo de clasificación**

**Fuente: Elaboración propia**

Entre la época 240 a la 250, la exactitud logra aumentar del 0.896275 a 0.9072 y la pérdida disminuye de 0.2425 a 0.2424, donde finaliza el entrenamiento. Se puede observar que los valores de exactitud de validación tienen una disminución mínima, lo cual indica que un mayor entrenamiento del modelo podría causar un sobreajuste, con estos resultados se tienen ya los valores establecidos para realizar la clasificación de nuevos clientes, como se muestra en la figura 71.

240 ▶ {val\_loss: 0.7605658173561096, val\_acc: 0.699999988079071, loss: 0.3118516206741333, acc: 0.8996296525001526}

241 ▶ {val\_loss: 0.7572177648544312, val\_acc: 0.6966666579246521, loss: 0.3105561435222626, acc: 0.9020370244979858}

242 ▶ {val\_loss: 0.7802460789680481, val\_acc: 0.6966666579246521, loss: 0.3092767000198364, acc: 0.9025925993919373}

243 ▶ {val\_loss: 0.7659571766853333, val\_acc: 0.6916666626930237, loss: 0.30876463651657104, acc: 0.899814784526825}

244 ▶ {val\_loss: 0.7696533799171448, val\_acc: 0.6916666626930237, loss: 0.307741641998291, acc: 0.9044444561004639}

245 ▶ {val\_loss: 0.7725109457969666, val\_acc: 0.6949999928474426, loss: 0.30634307861328125, acc: 0.9040740728378296}

246 ▶ {val\_loss: 0.772609293460846, val\_acc: 0.68833327293396, loss: 0.30533885955810547, acc: 0.9072222113609314}

247 ▶ {val\_loss: 0.7799980044364929, val\_acc: 0.68833327293396, loss: 0.3041500151157379, acc: 0.9053703546524048}

248 ▶ {val\_loss: 0.7827581763267517, val\_acc: 0.6899999976158142, loss: 0.30340754985809326, acc: 0.9075925946235657}

249 ▶ {val\_loss: 0.7841905355453491, val\_acc: 0.6866666078567505, loss: 0.3019312620162964, acc: 0.9055555462837219}

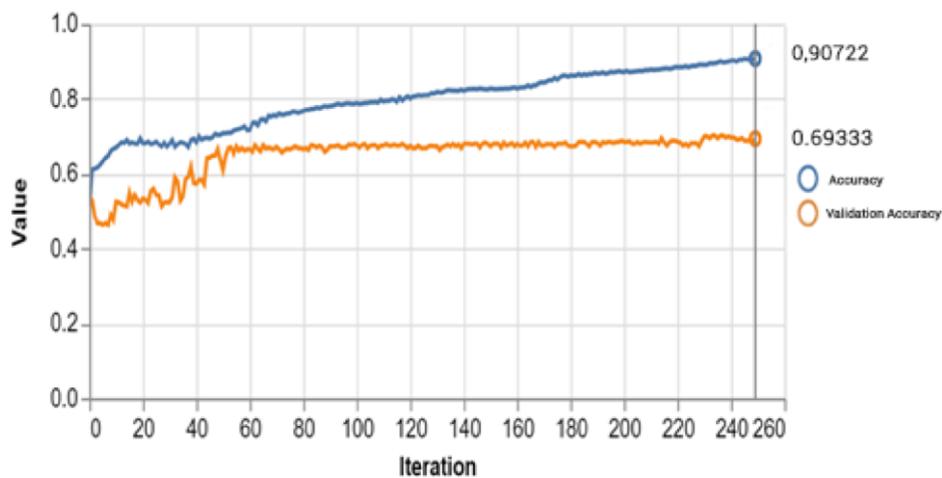
250 ▶ {val\_loss: 0.7856701016426086, val\_acc: 0.693333327702332, loss: 0.3011634051799774, acc: 0.9072222113609314}

**Figura 71. Época 240 a la 250 de modelo de clasificación**

**Fuente: Elaboración propia**

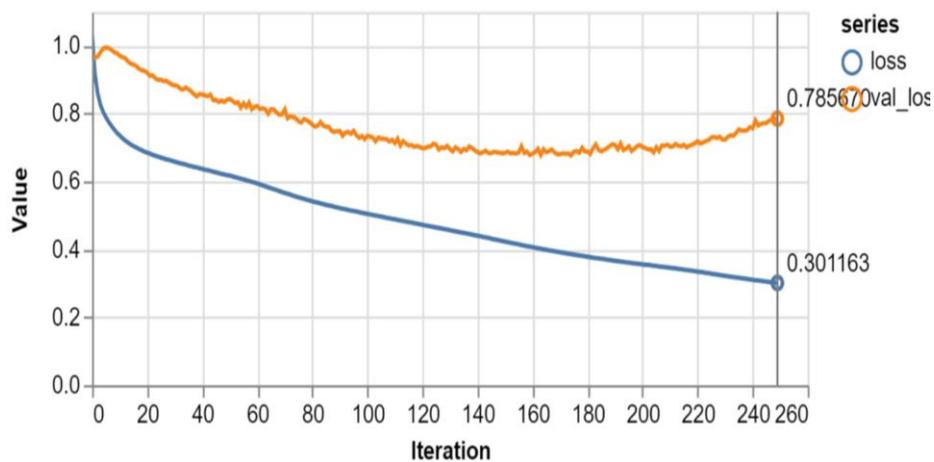
### 5.2.5 Resultados y Validación del Modelo

En la figura 72, se muestra el Accuracy en cada iteración, la cual expresa la exactitud del modelo de clasificación durante el entrenamiento, resultando valores entre 0 y 1, donde 0 es 0% y 1 representa el 100%. El resultado del entrenamiento, en este caso, alcanzó un 0,90722. De la misma manera se muestra el Validation Accuracy, que obtuvo un resultado de 0.69333 que representa la exactitud de la división usada para la validación del modelo, el porcentaje de datos usados para la validación fue del 10% del conjunto de datos usado para el entrenamiento.



**Figura 72. Accuracy y Validation Accuracy**  
Fuente: Elaboración propia

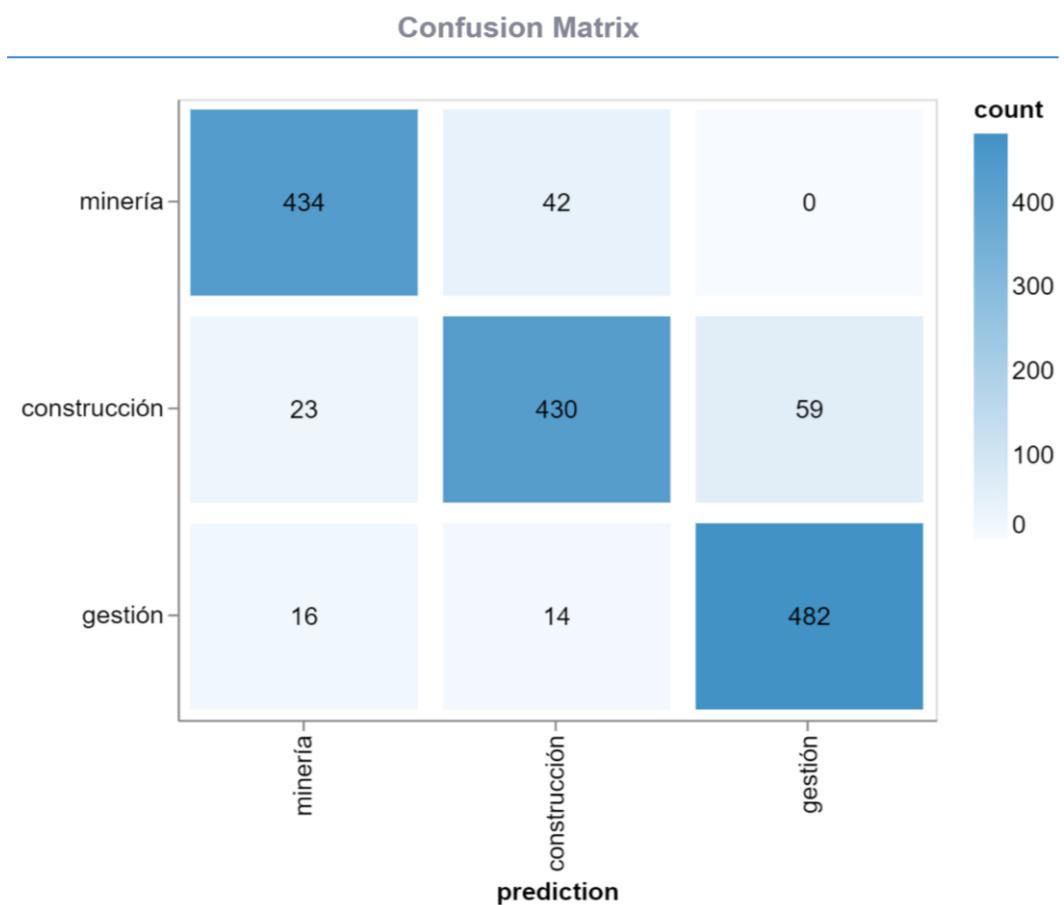
En la figura 73, se muestra el resultado del Loss en cada iteración, se ve una disminución del valor de pérdida alcanzando un resultado de 0.3011, y con Validation Loss un resultado de 0.7856, valores que se van acercando a 0; lo cual significa que han logrado valores bajos de Pérdida, los cuales indican resultados más precisos.



**Figura 73. Lost y Validation Loss**  
Fuente: Elaboración propia

En la figura 74, se muestra la Matriz de Confusión que evalúa el rendimiento del modelo en forma general, representa el resultado de la clasificación usando el modelo ya entrenado; se utilizaron 1500 datos divididos en 500 para cada categoría, donde cada columna de la matriz representa la identificación de las categorías del modelo propuesto, mientras que cada fila representa los valores actuales. Esta matriz nos facilita comprobar de manera rápida si el modelo está confundiendo las clases o no, mostrando a través de la intensidad del color dónde se reúne la mayor cantidad de datos; confirmando así de forma visual, que la mayor cantidad de datos se reúne en la diagonal, lo cual es una prueba rápida de que el modelo es consistente.

Analizando los resultados más detalladamente, la diagonal muestra los resultados predichos correctamente de las clases: minería, construcción y gestión, que obtuvieron los valores de 434, 430 y 482 respectivamente; para la clase minería se obtuvieron los valores de 23 y 16 como falsos positivos y 42 y 0 como falsos negativos; para la clase construcción se obtuvieron los valores de 42 y 14 como falsos positivos y 23 y 59 como falsos negativos; y para la clase de gestión, los valores de 0 y 59 como falsos positivos y 16 y 14 como falsos negativos. Estos resultados obtenidos por la matriz de confusión nos permiten obtener métricas para la validación del modelo de clasificación.



**Figura 74. Matriz de confusión**

**Fuente: Elaboración propia**

Obteniendo los resultados de la matriz de confusión se logra adquirir las métricas para la validación del modelo, resultando valores entre 0 y 1, donde 0 es 0% y 1 representa el 100% con una exactitud para la clasificación de 0,90722. Se obtienen las métricas Precisión, Recall y el puntaje F1 para las clases Minería, Construcción y Gestión; donde el resultado de Precisión confirma que el porcentaje de los resultados que hemos dicho que son la clase positiva, lo son; el resultado de Recall indica que hemos sido capaces de identificar el porcentaje de la clase positiva; y el resultado de F1 Score, combina Precisión y Recall en una sola medida, como se muestra en la tabla 36.

Clases	Precisión	Recall	F1 Score	Exactitud
Minería	0.918	0.912	0.915	0.9072
Construcción	0.885	0.840	0.862	
Gestión	0.891	0.941	0.915	

**Tabla 36. Métricas Precisión, Recall , F1 Score y Exactitud**

Fuente: Elaboración propia

### 5.3 Pruebas y resultados

Para realizar las pruebas de exactitud, eficiencia y fiabilidad, se consideraron 2000 registros nuevos de clientes y se tomaron de igual modo 2000 registros anteriores, del después de implementar la solución y del antes respectivamente. Se obtuvo como tamaño de muestra 158 datos para ambos casos, se realizó el cálculo de la muestra indicado en el Anexo 1.

Así mismo en la figura 75, se muestra la clasificación obtenida en el sistema WEB.

Nombre	Teléfono	Email	Edad	Sexo	Departamento	Estado civil	Cargo	Años exp	Resultado Clasificación	Detalle clasificación
RIVERA, DANIEL	975555554	rivera.daniel@gmail.com	42	M	AREQUIPA	C	OPERARIO	13	CONSTRUCCIÓN	Ver
LOPEZ, CARLOS	945555553	lopez.carlos@gmail.com	27	M	MOQUEGUA	S	CONDUCTOR	3	MINERÍA	Ver
JONES, ANA	965555556	jones.ana@gmail.com	29	M	MOQUEGUA	S	SUPERVISOR	7	MINERÍA	Ver
CHAVEZ, MICHAEL	975555555	chavez.michael@gmail.com	33	M	AREQUIPA	S	JEFE DE OBRA	8	CONSTRUCCIÓN	Ver
ANTONIO, NERY	995555556	antonio.nery@gmail.com	30	M	AREQUIPA	C	ASISTENTE	4	GESTIÓN	Ver
RIVERA, DANIEL	925555553	rivera.daniel@outlook.com	35	M	CUSCO	S	OPERARIO	7	CONSTRUCCIÓN	Ver
CHAVEZ, JUAN	995555553	chavez.juan@gmail.com	36	M	AREQUIPA	S	CONDUCTOR	9	MINERÍA	Ver
YAN, ANA	985555552	yan.ana@gmail.com	29	M	LIMA	S	SUPERVISOR	7	MINERÍA	Ver
CHAVEZ, JUAN	985555552	chavez.juan@gmail.com	35	M	AREQUIPA	S	OFICIAL	9	CONSTRUCCIÓN	Ver
PEDRO, JUAN	905555550	pedro.juan@gmail.com	36	M	CUSCO	S	SUPERVISOR	10	MINERÍA	Ver
FLORES, LUIS	995555559	flores.luis@gmail.com	27	M	AREQUIPA	C	TÉCNICO	7	MINERÍA	Ver
CHAVEZ, MARCO	945555556	chavez.marco@gmail.com	35	M	TACNA	S	OFICIAL	10	CONSTRUCCIÓN	Ver
HERRERA, LUIS	945555551	herrera.luis@gmail.com	36	M	LIMA	S	JEFE DE DEPARTAMENTO	12	GESTIÓN	Ver
MORALES, DANIEL	985555550	moraless.daniel@gmail.com	34	M	AREQUIPA	S	SUPERVISOR	7	CONSTRUCCIÓN	Ver
PEDRO, DANIEL	975555550	pedro.daniel@gmail.com	28	M	AREQUIPA	C	SUPERVISOR	19	MINERÍA	Ver

**Figura 75. Clasificación obtenida en la aplicación web**

Fuente: Elaboración propia

### 5.3.1 Pruebas de aceptación

Para la prueba de aceptación se realizó una encuesta sobre usabilidad y funcionabilidad del sistema, con el fin de conocer el nivel de satisfacción de los usuarios que hacen uso de la aplicación web. Ver Anexo 2.

#### 5.3.1.1 Usabilidad

Enfocado en medir la facilidad de uso del sistema de clasificación, la comodidad, facilidad o complejidad; recogiendo información de los usuarios que hacen uso del sistema mediante las preguntas que se muestran en la tabla 37.

Pruebas de usabilidad						
Seleccione una casilla en una escala del 1 al 5, donde 5 es muy satisfecho y 1 muy insatisfecho		1	2	3	4	5
1	¿El sistema muestra la información de manera clara y comprensiva?					
2	¿La apariencia del sistema es consistente en la pantalla del dispositivo móvil, escritorio, entre otros?					
3	¿El tamaño y tipo de letra utilizados en el sistema permiten su uso sin esfuerzo visual?					
4	¿Es clara la función del menú y los controles proporcionados por el sistema?					
5	¿El sistema carga rápidamente?					

**Tabla 37. Pruebas de usabilidad**

**Fuente: Elaboración propia**

#### 5.3.1.2 Funcionalidad

Enfocado en comprobar las características funcionales del sistema, la rentabilidad y la experiencia de usuario, mediante las preguntas que se muestran en la tabla 38.

Pruebas de funcionalidad						
Seleccione una casilla en una escala del 1 a 5, donde 5 es muy satisfecho y 1 muy insatisfecho		1	2	3	4	5
1	¿La aplicación web permite registrar clientes con facilidad?					
2	¿La aplicación web permite registrar usuarios con facilidad?					

3	¿La aplicación web permite realizar la clasificación de los clientes registrados con facilidad?					
4	¿La aplicación web permite visualizar comprensiblemente la clasificación de clientes por categoría?					
5	¿La exportación de los datos clasificados por la aplicación es comprensible?					

**Tabla 38. Pruebas de funcionalidad**

Fuente: Elaboración propia

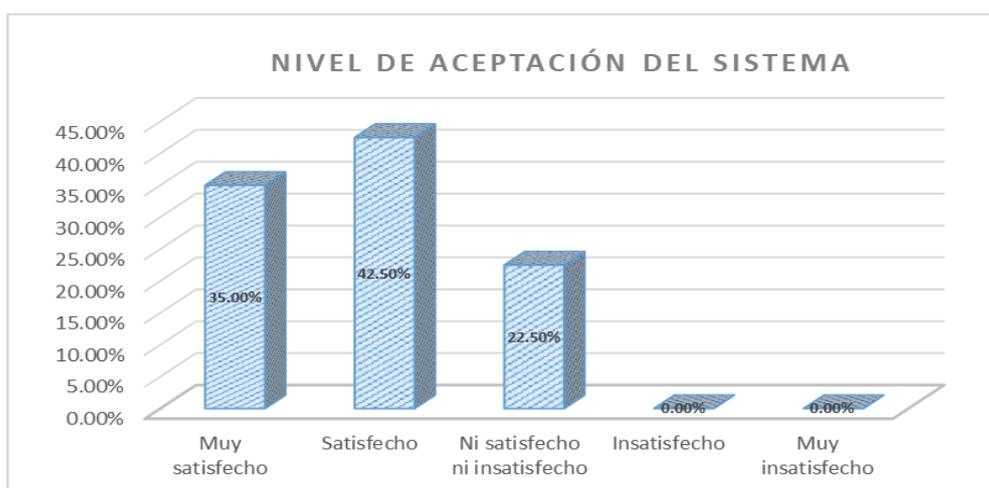
Después de haber terminado, se realizó la tabulación de la información obteniendo los resultados que se muestran en la tabla 39.

Encuestados	Preguntas usabilidad					Preguntas funcionalidad				
	P.1	P.2	P.3	P.4	P.5	P.1	P.2	P.3	P.4	P.5
Jefe de marketing	4	5	4	4	5	5	4	3	3	4
Jefe de ventas	5	5	3	5	3	3	4	4	3	5
Jefe de área de sistemas	4	4	5	4	5	5	4	4	5	3
Supervisor de ventas	4	5	4	3	4	4	5	4	3	5

**Tabla 39. Resultados de encuestas**

Fuente: Elaboración propia

La figura 76, detalla la aceptación de los usuarios del sistema, obteniendo resultados favorables que indican que el 35% de las respuestas fueron “Muy satisfecho”, el 42.50% de las respuestas fueron “Satisfecho” y el 22.50% de las respuestas restantes fueron “Ni satisfecho ni insatisfecho”.



**Figura 76. Resultados de prueba de aceptación**

Fuente: Elaboración propia

### 5.3.2 Pruebas de exactitud

Para analizar la posible mejora de exactitud de clasificación respecto al porcentaje de exactitud obtenido antes de implementar la solución, se tomaron como muestra 158 para el antes y para el después de implementada la solución, los cuales se ubican en las columnas de exactitud indicadas en el Anexo 1.

Con respecto al antes de la implementación se obtuvo un porcentaje de exactitud del 61% como media, y con la implementación de la solución, se obtuvo un 91% de media de exactitud, se puede considerar este valor a priori como una mejora en el porcentaje de exactitud.

### 5.3.3 Resultados de exactitud

Considerando que ambas muestras se tomaron aleatoriamente de datos de antes de implementar la solución y después de implementarla, se consideran independientes; aplicamos la prueba de normalidad a la muestra de exactitud del antes y después, y observamos que el valor de  $P=0.000$  es menor que 0.05, como se muestra en la figura 77, lo cual indica que la muestra no está distribuida siguiendo una normal.

Prueba de Kolmogorov-Smirnov para una muestra

		exactitud_antes	exactitud_desp ues
N		158	158
Parámetros normales <sup>a,b</sup>	Media	,61	,91
	Desv. Desviación	,490	,294
Máximas diferencias extremas	Absoluto	,396	,532
	Positivo	,285	,373
	Negativo	-,396	-,532
Estadístico de prueba		,396	,532
Sig. asintótica(bilateral)		,000 <sup>c</sup>	,000 <sup>c</sup>

Figura 77. Prueba de Kolmogorov-Smirnov para exactitud

Fuente: Resultados obtenidos desde SPSS

Considerando los resultados obtenidos de la prueba de normalidad de las muestras de exactitud, al aplicar la prueba U de Mann Whitney para las 2 muestras independientes se ha obtenido un valor de  $P=0.000$ , como se muestra en la figura 78; considerando que  $P \text{ valor} < 0.005$ , se sigue que estadísticamente existe una diferencia significativa en los valores de ambas muestras; considerando las medias, se sigue que el valor de 90.5 es significativamente mejor que el de 60.75.

## Estadísticos de prueba<sup>a</sup>

	Exactitud
U de Mann-Whitney	8769,000
W de Wilcoxon	21330,000
Z	-6,149
Sig. asintótica (bilateral)	,000

**Figura 78. Prueba U de Mann Whitney para exactitud**  
Fuente: Resultados obtenidos desde SPSS

### 5.3.4 Pruebas de eficiencia

Para analizar la posible mejora en el tiempo de clasificación respecto al tiempo obtenido antes de implementar la solución, se tomaron como muestra 158 datos para el antes y para el después de implementada la solución, los cuales se ubican en las columnas tiempo indicadas en el Anexo 1.

Con respecto al antes de la implementación se obtuvo 7.50 segundos como media, y ya con la implementación se obtuvo 0.004 segundos de media en el tiempo de clasificación; se puede considerar entonces, este valor a priori como una mejora en el tiempo de clasificación.

### 5.3.5 Resultados de eficiencia

Considerando que ambas muestras se tomaron aleatoriamente de datos de antes de implementar la solución y después de implementarla, se consideran independientes; aplicamos la prueba de normalidad a la muestra de exactitud de antes y después, y observamos que el valor de  $P=0.000$  es menor que 0.05, como se muestra en la figura 79, lo cual indica que la muestra no está distribuida siguiendo una normal.

		tiempo_antes	tiempo_después
N		158	158
Parámetros normales <sup>a,b</sup>	Media	7.5051544	.0040139
	Desv. Desviación	.79408595	.00113567
Máximas diferencias extremas	Absoluto	.081	.073
	Positivo	.063	.073
	Negativo	-.081	-.073
Estadístico de prueba		.081	.073
Sig. asintótica(bilateral)		.013 <sup>c</sup>	.037 <sup>c</sup>

**Figura 79. Prueba de Kolmogorov-Smirnov para eficiencia**  
Fuente: Resultados obtenidos desde SPSS

Considerando los resultados obtenidos de la prueba de normalidad de las muestras de exactitud, al aplicar la prueba U de Mann Whitney para las 2 muestras independientes se ha obtenido un valor de  $P=0.000$ , como se muestra en la figura 80; considerando que  $P \text{ valor} < 0.005$ , se sigue que en los valores de ambas muestras estadísticamente existe una diferencia significativa; considerando las medias, se sigue que el valor de 7.5051 es significativamente mejor que el de 0.0040.

### Estadísticos de prueba<sup>a</sup>

	Tiempo
U de Mann-Whitney	,000
W de Wilcoxon	12561,000
Z	-15,371
Sig. asintótica(bilateral)	,000

**Figura 80. Prueba U de Mann Whitney para eficiencia**

**Fuente: Resultados obtenidos desde SPSS**

#### 5.3.6 Pruebas de fiabilidad

Para analizar la posible mejora en la consistencia de datos respecto a la consistencia obtenida antes de implementar la solución, se tomaron como muestra 158 datos para el antes y para el después de implementada la solución, los cuales se ubican en las columnas de consistencia indicadas en el Anexo 1.

Con respecto al antes de la implementación se obtuvo un porcentaje de consistencia del 95% como media y con la implementación de la solución se obtuvo un 100% de media; se puede considerar entonces, este valor a priori como una mejora en el porcentaje de consistencia de datos.

#### 5.3.7 Resultados de fiabilidad

Considerando que ambas muestras se tomaron aleatoriamente de datos de antes de implementar la solución y después de implementarla, se consideran independientes; aplicamos la prueba de normalidad a la muestra de exactitud de antes y después, y observamos que el valor de  $P=0.000$  es menor que 0.05, como se muestra en la figura 81, lo cual indica que la muestra no está distribuida siguiendo una normal.

		antes	despues
N		158	158
Parámetros normales <sup>a,b</sup>	Media	,95	1,00
	Desv. Desviación	,220	,000 <sup>d</sup>
Máximas diferencias extremas	Absoluto	,540	
	Positivo	,409	
	Negativo	-,540	
Estadístico de prueba		,540	
Sig. asintótica(bilateral)		,000 <sup>e</sup>	

**Figura 81. Prueba de Kolmogorov-Smirnov para fiabilidad**

**Fuente: Resultados obtenidos desde SPSS**

Considerando los resultados obtenidos de la prueba de normalidad de las muestras de exactitud, al aplicar la prueba U de Mann Whitney para las 2 muestras independientes se ha obtenido un valor de  $P=0.004$ , como se muestra en la figura 82; considerando que  $P \text{ valor} < 0.005$ , se sigue que en los valores de ambas muestras estadísticamente existe diferencia significativa; considerando las medias, se sigue que el valor de 1 es significativamente mejor que el de 0.95.

### Estadísticos de prueba<sup>a</sup>

	consistencia
U de Mann-Whitney	11850,000
W de Wilcoxon	24411,000
Z	-2,860
Sig. asintótica(bilateral)	,004

**Figura 82. Prueba U de Mann Whitney para fiabilidad**

**Fuente: Resultados obtenidos desde SPSS**

## CONCLUSIONES

1. Con respecto al objetivo específico 1: "Mejorar la exactitud de la clasificación de clientes para la empresa CFCGROUP - Arequipa 2021", se concluye que este se ha logrado, ya que los resultados indican que se ha pasado de 60.75% a 90.50% de exactitud, con un margen de error de menos del 5%.
2. Con respecto al objetivo específico 2: "Reducir el tiempo de clasificación de clientes para la empresa CFCGROUP - Arequipa 2021", se concluye que se logró reducir el tiempo de clasificación, ya que los resultados indican que se ha pasado de 7.5052 a 0.0040 segundos, con un margen de error de menos del 5%.
3. Con respecto al objetivo específico 3: "Reducir las inconsistencias en los datos de clientes para la empresa CFCGROUP - Arequipa 2021", se concluye que este se ha logrado, ya que los resultados indican que se ha pasado de 5.063% a 0% de inconsistencias, con un margen de error de menos del 5%.
4. Considerando los logros de los objetivos específicos se concluye que se ha logrado el objetivo general: "Mejorar la clasificación de clientes para la empresa CFCGROUP - Arequipa – 2021 mediante una aplicación Web", infiriendo que el uso del Aprendizaje Automático en conjunto con una aplicación web es una alternativa aceptable para la clasificación de clientes, al haber obtenido una tasa de acierto del 90.50%.

## TRABAJOS FUTUROS

1. Con respecto a "Mejorar la exactitud de la clasificación de clientes para la empresa CFCGROUP - Arequipa 2021", se recomienda incrementar los datos de entrenamiento utilizados por el modelo de clasificación, para lograr un porcentaje de exactitud mayor.
2. Con respecto a "Reducir el tiempo de clasificación de clientes para la empresa CFCGROUP - Arequipa 2021", se recomienda agrupar los datos de las variables utilizadas cuando se incrementen los datos de entrenamiento, con la finalidad de reducir el tiempo de clasificación del modelo.
3. Con respecto a "Reducir las inconsistencias en los datos de clientes para la empresa CFCGROUP - Arequipa 2021", se recomienda un análisis mayor en los tipos y características de los datos, y de este modo agregar más validaciones con la finalidad de evitar inconsistencias.
4. Con respecto al proyecto en general, se recomienda agregar periódicamente los datos clasificados de manera correcta, con el fin de incrementar los datos de entrenamiento utilizados por modelo de Aprendizaje Automático, y con esto conseguir aumentar el porcentaje de precisión.

## REFERENCIAS BIBLIOGRÁFICAS

1. **ESPINOZA AIRAC, Gianira Xiomara y LEON MUÑOZ, Eduar Fabián.** *Modelo de Machine Learning para la clasificación de estudiantes de acuerdo a su rendimiento académico en el centro de idiomas de la Universidad Nacional del Santa.* s.l. : Universidad Nacional del Santa, 2020.
2. **WBEIMAR OSSA, Giraldo y JARAMILLO MARIN, Veronica.** *Machine Learning para la estimación del riesgo de crédito en una cartera de consumo.* s.l. : Universidad EAFIT, 2021.
3. **BELMAN LÓPEZ, Carlos Eduardo, JIMENEZ GARCIA, Jose Alfredo y VÁZQUES LÓPEZ, José Antonio.** *Clasificación de clientes de la industria bancaria por métodos estadísticos y redes neuronales artificiales.* s.l. : Tecnológico Nacional de Mexico en Celaya, 2019.
4. **ALVAREZ, Fredi.** *Machine Learning en la detección de fraudes de comercio electrónico aplicado a los servicios bancarios.* Mexico : Softwaresocial Consultores, 2020.
5. **AUCAPIÑA PERALTA, Jonnatan Mesias.** *Modelo de elaboración de pronóstico de ventas mediante el uso de redes neuronales artificiales y SVR.* s.l. : Universidad de Cuenca, 2018.
6. **GARCÍA GAZABÓN, Gisela Isabel.** *Modelo Machine Learning para la Clasificación de pacientes en términos del nivel asistencial requerido en una urgencia pediátrica con Área de Cuidados Mínimos.* s.l. : Universidad Tecnológica de Bolívar, 2014.
7. **GONZÁLEZ DIAZ, Humberto.** *Técnicas de Ingeniería informática e inteligencia artificial para clasificación : aplicaciones para el descubrimiento de fármacos y dianas moleculares.* s.l. : Universidad de Coruña, 2013.
8. **VILLEGAS CUBAS, Juan Elías.** *Modelo de Machine Learning en la detección de sitios web Phishing.* s.l. : Universidad Señor de Sipán, 2021.
9. **CHAHUARA FLORES, Lesly y ALIAGA ANCCO, Josselin Jimena.** *Clasificación de las prioridad de atención a Reclamos Presentados por clientes utilizando Machine Learning.* s.l. : Universidad Peruana Union, 2019.
10. **VAZQUEZ LEYVA, Oliver.** *Sistema Predictivo Basado en un modelo Credit Scoring de aprendizaje Automático para la Medición del Riesgo Crediticio en los Créditos PYME de la EDPYME alternativa S.A.* s.l. : Universidad Señor de Sipán, 2019.
11. **LIZARES CASTILLO, Monica.** *Comparación de modelos de clasificación, regresión logística y árboles de clasificación para evaluar el rendimiento académico.* s.l. : Universidad Mayor de San Marcos, 2017.
12. **BOCANGEL CARBAJAL, Jose Luis, y otros.** *Análisis de las características que identifican a un usuario de practisis premium; variables que deciden para convertirse de una cuenta freemium a premium.* s.l. : Universidad Peruana de Ciencias Aplicadas, 2020.
13. **CONCEPCION TIZA, Miguel Angel.** *Sistema de Gestión y clasificación automática de denuncias ambientales mediante aprendizaje máquina.* s.l. : Universidad Peruana de Ciencias Aplicadas, 2021.
14. **NUÑEZ VELARDE, Kenyi Guillermo y PACHECO CHÁVEZ, Philipp Gil.** *Optimización del Proceso de gestión de flota para una empresa de transporte de carga por carretera usando Machine Learning, BI, GPS y SMS Gateway.* s.l. : Universidad Peruana de Ciencias Aplicadas, 2021.
15. **VALDEZ PINO, Edgar Omar.** *Desarrollo de un sistema de proyección de costos y costeo unitario de importaciones con métodos predictivos basados en Machine Learning.* s.l. : Universidad Peruana de Ciencias Aplicadas, 2021.

16. **DÁVILA HURTADO, Luis Alberto.** *Mejora del Proceso de Toma de Decisiones en las Ventas de Abarrotes de una Empresa de Chiclayo, a partir de un Sistema Informático basado en Herramientas OLAP.* s.l. : Universidad Tecnológica del Perú, 2021.
17. **KIONETWORKS.** Kionetworks. [En línea] [Citado el: 08 de 01 de 2022.] <https://www.kionetworks.com/blog/data-center/los-sistemas-de-informacion-de-una-empresa>.
18. **ALTAMETRICS.** AltaMetrics.com. [En línea] [Citado el: 08 de 01 de 2022.] <https://altametrics.com/es/information-systems/information-system-types.html>.
19. **ADOBE.** Adobe Helpx. [En línea] [Citado el: 08 de 01 de 2022.] <https://helpx.adobe.com/es/dreamweaver/using/web-applications.html>.
20. **DEVELOPER MOZILLA.** Mozilla. [En línea] [Citado el: 08 de 01 de 2022.] <https://developer.mozilla.org/es/docs/Web/HTTP/Overview>.
21. **MOZILLA DOCS.** Mozilla. [En línea] [Citado el: 08 de 01 de 2022.] <https://developer.mozilla.org/es/docs/Web/HTML>.
22. **DEVELOPER MOZILLA.** Mozilla. [En línea] [Citado el: 08 de 01 de 2022.] <https://developer.mozilla.org/es/docs/Web/CSS>.
23. **VELASCO, Rubén.** Softzone. [En línea] [Citado el: 08 de 01 de 2022.] <https://www.softzone.es/programas/lenguajes/lenguajes-programacion-web/>.
24. **EGUILUZ PÉREZ, Javier.** *Introducción a JavaScript.* España : s.n., 2009.
25. **ARIAS, Miguel.** *Introducción a php.* 2007.
26. **LARAVEL DOCS.** Laravel.com. [En línea] [Citado el: 08 de 01 de 2022.] <https://laravel.com/docs/6.x>.
27. **CHELSEA ORTIZ, Pamela.** Flexisourceit. [En línea] [Citado el: 08 de 01 de 2022.] <https://flexisourceit.com.au/resources/blog/php-vs-python-vs-java-vs-asp-vs-ruby/>.
28. **ORACLE.** Oracle. [En línea] [Citado el: 08 de 01 de 2022.] <https://www.oracle.com/mx/artificial-intelligence/what-is-ai/>.
29. **HPE.** Hpe. [En línea] [Citado el: 08 de 01 de 2022.] <https://www.hpe.com/lamerica/es/what-is/machine-learning.html#:~:text=El%20aprendizaje%20autom%C3%A1tico%20es%20una,nuevas%20situaciones%20de%20manera%20independiente..>
30. **CAPARRINI, Fernando Sancho.** Dpto. de Ciencias de la Computación e Inteligencia Artificial de la Universidad de Sevilla. [En línea] [Citado el: 08 de 01 de 2022.] <http://www.cs.us.es/~fsancho/?e=231>.
31. **ELI STEVENS, Luca Antiga y VIEHMANN, Thomas.** *Deep Learning with Pytorch.* 2009.
32. **FRICKLAS, Kenneth.** *Machine Learning with TensorFlow.* 2013.
33. **DUBOVIKOV, Kirill.** TowardsDataScience.com. [En línea] [Citado el: 08 de 01 de 2022.] <https://towardsdatascience.com/pytorch-vs-tensorflow-spotting-the-difference-25c75777377b>.
34. **SUCAR, Luis Enrique.** *Redes Bayesianas.* Mexico : s.n., 2009.
35. **P, SOUCY y G, Mineau.** *A simple KNN algorithm for text categorization.* San Jose , USA : s.n., 2001.
36. **RONG, Xiao, JICHENG, Wang y FAYAN, Zhang.** *An approach to incremental SVM learning algorithm.* Canada : s.n., 2000.
37. **IZAURIETA, Fernando y SAAVEDRA, Carlos.** *Redes Neuronales Artificiales.* Chile : Universidad de Concepción.
38. **MATICH, Damián Jorge.** *Redes Neuronales: Conceptos básicos y aplicaciones.* 2001.
39. **HEATON, Jeff.** HeatonResearch. [En línea] [Citado el: 08 de 01 de 2022.] <https://www.heatonresearch.com/2017/06/01/hidden-layers.html>.
40. **GUZMAN ARENAS, Adolfo.** *Clasificación supervisada inducción de árboles de decisión algoritmo.* 2006.

41. **CFCGROUP**. *Informe - Oficina de Recursos Humanos*. s.l. : Cfcgroup, 2021.
42. **SCHWABER, Ken y SUTHERLAND, Jeff**. *La Guía de Scrum*. s.l. : Ken Schwaber and Jeff Sutherland, 2020.
43. **MICROSOFT AZURE**. Microsoft. [En línea] [Citado el: 08 de 01 de 2022.] <https://azure.microsoft.com/es-es/overview/machine-learning-algorithms/#uses>.
44. **MICROSOFT DOCS**. Microsoft. [En línea] [Citado el: 08 de 01 de 2022.] <https://docs.microsoft.com/es-es/dotnet/machine-learning/resources/glossary>.
45. **CHAUHAN, Nagesh Singh**. DataSource. [En línea] [Citado el: 08 de 01 de 2022.] <https://www.datasource.ai/es/data-science-articles/metricas-de-evaluacion-de-modelos-en-el-aprendizaje-automatico>.
46. **ORACLE**. Oracle. [En línea] [Citado el: 08 de 01 de 2022.] <https://www.oracle.com/mx/database/what-is-database/>.
47. **TORRES HERNÁNDEZ, Miguel**. *Symfony Framework: Desarrollo Rápido de Aplicaciones Web*. 2016.
48. **DEVCAMP**. Devcamp. [En línea] [Citado el: 08 de 01 de 2022.] <https://devcamp.es/que-es-libreria-programacion/>.
49. **AMAZON**. Amazon Docs. [En línea] [Citado el: 08 de 01 de 2022.] [https://docs.aws.amazon.com/es\\_es/machine-learning/latest/dg/training-parameters1.html](https://docs.aws.amazon.com/es_es/machine-learning/latest/dg/training-parameters1.html).
50. **IBM CLOUD EDUCATION**. IBM. [En línea] 08 de 01 de 2022. <https://www.ibm.com/cloud/learn/overfitting>.
51. **MICROSOFT DOCS**. Microsoft. [En línea] [Citado el: 08 de 01 de 2022.] <https://docs.microsoft.com/es-mx/sql/machine-learning/r/reference/microsoftml/loss>.

## ANEXOS

### Anexo 1. Cálculo de tamaño de muestra

N=	2000	Población
e=	0.05	Error máximo permisible
p=	0.8	Proporción de aciertos
q=	0.2	Proporción de desaciertos
z=	1.64	Límite de confianza
e <sup>2</sup> =	0.0025	
z <sup>2</sup> =	2.6896	
no=	172.13	$z^2 \cdot p \cdot q / d^2 =$
n=	<b>158</b>	$no / (1 + no/N)$

## Anexo 2. Encuesta de Aceptación del sistema

### Encuesta de satisfacción del sistema

Seleccione una casilla en una escala del 1 a 5, donde 5 es muy satisfecho y 1 muy insatisfecho.

5 - Muy satisfecho

4 - Satisfecho

3 - Ni satisfecho ni insatisfecho

2 - insatisfecho

1 - Muy insatisfecho

- Usabilidad

1. ¿El sistema muestra la información de manera clara y comprensiva?

5	4	3	2	1
---	---	---	---	---

2. ¿La apariencia del sistema es consistente en la pantalla del dispositivo móvil, escritorio, entre otros?

5	4	3	2	1
---	---	---	---	---

3. ¿El tamaño y tipo de letra utilizados en el sistema permiten su uso sin esfuerzo visual?

5	4	3	2	1
---	---	---	---	---

4. ¿Es clara la función del menú y los controles proporcionados por el sistema?

5	4	3	2	1
---	---	---	---	---

5. ¿El sistema carga rápidamente?

5	4	3	2	1
---	---	---	---	---

- Funcionalidad

1. ¿La aplicación web permite registrar clientes con facilidad?

5	4	3	2	1
---	---	---	---	---

2. ¿La aplicación web permite registrar usuarios con facilidad?

5	4	3	2	1
---	---	---	---	---

3. ¿La aplicación web permite realizar la clasificación de los clientes registrados con facilidad?

5	4	3	2	1
---	---	---	---	---

4. ¿La aplicación web permite visualizar comprensiblemente la clasificación de clientes por categoría?

5	4	3	2	1
---	---	---	---	---

5. ¿La exportación de los datos clasificados por la aplicación son comprensibles?

5	4	3	2	1
---	---	---	---	---

**Anexo 3. Toma de exactitud, de tiempo y consistencia de las muestras del antes y después.**

Antes				Después			
Número	Exactitud ( Inexacto = 0, Exacto = 1)	Tiempo (Segundos)	Consistencia ( Inconsistente = 0; Consistente = 1)	Número	Exactitud ( Inexacto = 0, Exacto = 1)	Tiempo (Segundos)	Consistencia ( Inconsistente = 0; Consistente = 1)
1	0	8.3558	1	1	1	0.0032	1
2	1	8.0686	1	2	1	0.0060	1
3	0	6.5468	1	3	1	0.0033	1
4	1	8.1295	1	4	1	0.0029	1
5	1	6.2522	1	5	1	0.0053	1
6	1	7.2156	1	6	1	0.0038	1
7	1	6.7944	1	7	1	0.0034	1
8	1	7.3138	1	8	1	0.0036	1
9	1	7.8152	1	9	1	0.0025	1
10	1	6.1032	1	10	1	0.0053	1
11	1	7.1707	0	11	1	0.0034	1
12	0	7.6576	1	12	1	0.0038	1
13	1	6.0511	1	13	1	0.0053	1
14	1	6.6297	1	14	1	0.0022	1
15	0	6.5289	1	15	1	0.0056	1
16	1	6.9921	1	16	1	0.0051	1
17	1	6.3989	1	17	1	0.0024	1
18	1	8.9789	1	18	1	0.0025	1
19	1	7.9727	1	19	1	0.0043	1
20	0	8.1133	0	20	1	0.0052	1
21	1	7.3202	1	21	1	0.0052	1
22	1	7.8893	1	22	1	0.0055	1
23	1	7.9419	1	23	1	0.0023	1
24	1	7.0108	1	24	1	0.0055	1
25	1	6.0717	1	25	1	0.0039	1
26	1	8.4649	1	26	1	0.0056	1
27	1	8.1345	1	27	1	0.0032	1
28	1	6.4539	1	28	1	0.0048	1
29	1	7.3086	1	29	1	0.0042	1
30	0	6.2205	1	30	1	0.0029	1
31	1	7.9687	1	31	1	0.0059	1
32	1	6.0936	1	32	1	0.0030	1
33	1	6.5988	1	33	1	0.0040	1
34	0	7.6286	1	34	0	0.0030	1
35	0	8.4088	1	35	1	0.0047	1
36	1	7.3562	1	36	1	0.0030	1
37	0	6.6487	1	37	1	0.0037	1
38	1	6.6417	1	38	1	0.0020	1

39	1	8.1944	1	39	1	0.0036	1
40	0	7.5150	1	40	1	0.0028	1
41	0	6.9905	1	41	1	0.0056	1
42	1	8.4401	1	42	1	0.0034	1
43	1	7.5199	1	43	1	0.0029	1
44	0	6.9045	1	44	1	0.0056	1
45	1	8.1330	1	45	1	0.0027	1
46	0	7.6227	1	46	0	0.0023	1
47	0	7.9615	1	47	0	0.0046	1
48	1	6.1979	1	48	1	0.0040	1
49	1	8.3741	1	49	1	0.0045	1
50	0	7.9252	1	50	1	0.0044	1
51	0	6.6634	1	51	1	0.0055	1
52	1	7.7721	1	52	1	0.0049	1
53	1	7.4859	1	53	1	0.0032	1
54	0	6.4397	0	54	1	0.0030	1
55	0	7.4077	1	55	1	0.0040	1
56	1	6.5282	1	56	1	0.0057	1
57	1	8.7860	1	57	1	0.0030	1
58	1	7.6976	1	58	1	0.0052	1
59	1	6.2261	1	59	1	0.0046	1
60	0	7.0590	1	60	1	0.0051	1
61	1	7.3103	1	61	1	0.0050	1
62	1	6.8599	1	62	1	0.0029	1
63	0	7.1828	1	63	0	0.0029	1
64	1	7.9855	1	64	1	0.0029	1
65	0	7.7461	1	65	1	0.0055	1
66	1	6.0547	1	66	1	0.0027	1
67	1	6.0462	1	67	1	0.0041	1
68	1	6.7080	0	68	1	0.0035	1
69	0	7.2789	1	69	1	0.0037	1
70	1	6.1920	1	70	1	0.0040	1
71	0	7.4200	1	71	1	0.0032	1
72	1	8.1361	1	72	1	0.0024	1
73	1	7.7563	1	73	1	0.0038	1
74	0	8.0743	1	74	1	0.0027	1
75	1	7.4633	1	75	1	0.0058	1
76	1	7.7902	1	76	1	0.0034	1
77	1	8.5264	1	77	1	0.0042	1
78	1	7.0351	1	78	1	0.0056	1
79	0	8.9658	1	79	1	0.0045	1
80	1	7.8050	1	80	1	0.0057	1
81	1	7.9546	1	81	1	0.0051	1
82	0	7.8008	1	82	1	0.0045	1
83	1	7.1940	0	83	1	0.0028	1
84	1	7.0966	1	84	1	0.0021	1

85	1	8.9846	1	85	1	0.0045	1
86	0	8.8894	1	86	0	0.0038	1
87	1	7.5393	1	87	1	0.0049	1
88	0	7.9084	1	88	1	0.0058	1
89	1	8.1711	1	89	1	0.0046	1
90	1	7.3955	1	90	1	0.0024	1
91	1	7.3981	1	91	1	0.0046	1
92	1	6.4321	1	92	1	0.0033	1
93	1	8.1513	1	93	1	0.0031	1
94	1	8.6119	1	94	1	0.0032	1
95	0	7.3963	1	95	1	0.0036	1
96	1	7.3575	1	96	1	0.0042	1
97	0	8.2644	0	97	1	0.0036	1
98	0	7.8118	1	98	1	0.0022	1
99	0	8.3174	1	99	1	0.0021	1
100	0	6.4140	1	100	1	0.0043	1
101	1	7.0250	1	101	1	0.0050	1
102	1	7.1480	1	102	1	0.0060	1
103	1	8.5305	1	103	1	0.0035	1
104	0	8.6179	1	104	0	0.0044	1
105	1	6.9071	1	105	1	0.0044	1
106	1	8.4214	1	106	1	0.0041	1
107	0	8.2495	1	107	0	0.0041	1
108	0	7.8190	1	108	1	0.0053	1
109	0	8.6555	1	109	1	0.0030	1
110	0	8.2195	1	110	1	0.0032	1
111	1	8.1068	1	111	1	0.0055	1
112	0	6.7612	1	112	1	0.0044	1
113	1	8.7696	1	113	1	0.0045	1
114	1	8.3347	1	114	1	0.0042	1
115	1	6.0030	1	115	1	0.0046	1
116	1	6.5381	1	116	1	0.0032	1
117	0	7.9880	0	117	1	0.0034	1
118	1	7.3437	1	118	1	0.0021	1
119	0	7.8175	1	119	0	0.0036	1
120	1	6.1498	1	120	1	0.0036	1
121	1	7.8373	1	121	1	0.0032	1
122	0	8.7113	1	122	1	0.0021	1
123	1	7.7692	1	123	1	0.0029	1
124	1	8.4767	1	124	1	0.0035	1
125	0	8.4632	1	125	1	0.0050	1
126	1	6.0327	1	126	1	0.0054	1
127	0	8.2475	1	127	1	0.0021	1
128	0	6.4990	1	128	1	0.0055	1
129	0	8.0127	1	129	0	0.0044	1
130	0	8.0454	1	130	0	0.0020	1

131	0	7.2489	1	131	1	0.0039	1
132	1	7.1593	1	132	1	0.0022	1
133	1	7.9367	1	133	1	0.0025	1
134	0	8.8357	0	134	0	0.0040	1
135	0	6.5151	1	135	0	0.0027	1
136	1	8.6241	1	136	1	0.0044	1
137	1	6.6686	1	137	1	0.0041	1
138	1	8.4920	1	138	1	0.0058	1
139	1	7.8641	1	139	1	0.0027	1
140	0	6.8036	1	140	0	0.0060	1
141	0	8.3845	1	141	1	0.0054	1
142	0	7.4242	1	142	0	0.0038	1
143	1	8.0334	1	143	1	0.0039	1
144	0	7.3906	1	144	1	0.0057	1
145	1	7.6372	1	145	1	0.0058	1
146	1	6.8913	1	146	1	0.0058	1
147	1	7.5662	1	147	1	0.0038	1
148	0	8.3253	1	148	1	0.0033	1
149	0	6.9220	1	149	1	0.0052	1
150	0	7.0757	1	150	1	0.0049	1
151	1	6.2669	1	151	1	0.0048	1
152	0	8.5911	1	152	1	0.0058	1
153	1	6.6066	1	153	1	0.0030	1
154	0	8.6741	1	154	1	0.0042	1
155	0	6.7483	1	155	0	0.0047	1
156	1	7.9589	1	156	1	0.0042	1
157	0	7.2490	1	157	1	0.0023	1
158	0	7.8993	1	158	1	0.0058	1

## Anexo 4. Constancia de conformidad del proyecto de tesis

### CARTA DE CONFORMIDAD

Por medio de la presente me dirijo a ustedes UNIVERSIDAD CONTINENTAL y hago de su conocimiento que la empresa CFCGROUP S.A.C con RUC 20603912838, con domicilio fiscal C.H. Dean Valdivia sector 5 lote 1 MZ R Arequipa – Arequipa. Se está conforme con la ejecución y los resultados del proyecto de tesis titulado “*APLICACIÓN WEB PARA LA CLASIFICACIÓN DE CLIENTES DE LA EMPRESA CFCGROUP*”, realizado por el Bachiller Diego Esteban Santa Cruz Álvarez y el Bachiller Renato Boris Zapana Llanquiche.

De igual manera se informa que el proyecto realizado cumplió con nuestras expectativas.

Atte.

Luis Branco Ticona Manrique

Gerente



Luis Branco Ticona Manrique  
CIP 283990  
DNI.70446773

Arequipa, 27 de febrero del 2022

## Anexo 5. Captura de pantalla de la hoja de cálculo proceso manual

	A	B	C	D	E	F	G	H	I	J	K	L
	Nro	Fecha de registro	Nombres y Apellidos	Dni	Telefono	Email	Edad	Sexo	Departamento	Estado civil	Cargo	Años de experiencia
1												
2	1	21/03/2018	HUMANA DANIEL CÉSAR DALEJO	46004120	936095546	huma_nad@outlook.com	30	M	arequipa	soltero	asistente	3
3	2	11/04/2018	RAMIRO DANIEL DAVID MAMANI	40013580	987744440	ramirodanielda@gmail.com	35	M	arequipa	soltero	operario	2
4	3	11/04/2018	CAROL MONTANA ROSMEL DAVID	46004120	980110070	carolm100@gmail.com	32	M	arequipa	soltero	jefe de departamento	1
5	4	11/04/2018	HUMANA DANIEL CÉSAR DALEJO	46004120	936095546	huma_nad@outlook.com	25	M	arequipa	soltero	asistente	3
6	5	11/04/2018	ÁLVAREZ ROSA JULIO CÉSAR	4640544	957330001	julio.alvarez200@gmail.com	33	M	arequipa	casado	operario	1
7	6	11/04/2018	IBRICO NIÑEZ JUAN KEVIN	70073004	956133000	ibrico.ninez@gmail.com	27	M	lima	casado	supervisor	5
8	7	22/04/2018	ARQUINO CARDENAS	74344360	944340000	arquino77javier@gmail.com	27	M	moquegua	casado	asistente	1
9	8	22/04/2018	RAMOS OLIVERA ERICK YOHANNAN	44085700	980000004	erickramos95@hotmail.com	25	M	moquegua	casado	operario	7
10	9	22/04/2018	FLORES MAÑAC YERIN	70444346	907003036	yerinflores14@gmail.com	26	M	moquegua	casado	jefe de departamento	5
11	10	22/04/2018	POMERO OLIVER JAMIEL	43355436	903065730	lamero_17@hotmail.com	33	M	moquegua	soltero	asistente	1
12	11	22/04/2018	MONTAÑANA JUAN RAMIRO	70044000	904430457	monta_nad@outlook.com	29	M	moquegua	soltero	operador equipo pesado	7
13	12	22/04/2018	ECHEVELE CAJALANA FRANCISCO	43733630	900040444	cajalana10@gmail.com	32	M	moquegua	soltero	asistente	4
14	13	22/04/2018	GAMARRA HUAMARCA MILO COM	41004824	961030410	wgamarra30@hotmail.com	31	M	moquegua	soltero	técnico	2
15	14	22/04/2018	EGGODOBO MANUYAMA JULIO CÉSAR	4304044	947054006	eggo_dob@gmail.com	32	M	moquegua	casado	jefe de obra	3
16	15	22/04/2018	REYES CRATEROL RENEALDO ENRIQUE	40700010	901700746	reyesre10@gmail.com	29	M	moquegua	casado	jefe de almacén	2
17	16	14/05/2018	ALVARADO ALVAREZ MIHAIL ALEVIS	71329687	978346710	mihail24coz@gmail.com	34	M	moquegua	casado	operador de equipo pesado	5
18	17	14/05/2018	NIÑEZA VÉREZ CERGOL CÉSAR EMANUEL	74506000	933330007	Cercol_05_2010@hotmail.com	34	M	moquegua	casado	operador de equipoliviano	4
19	18	14/05/2018	ARAZA GOMEZ CARLOS ALBERTO	35003666	956470046	carazaz1074@hotmail.com	30	M	moquegua	soltero	gerente	7
20	19	14/05/2018	NIÑEZA VÉREZ CERGOL CÉSAR EMANUEL	74506000	933330007	Cercol_05_2010@hotmail.com	26	M	moquegua	soltero	asistente	4
21	20	14/05/2018	RODRIGUEZ CABRERA MAGNAT MARIAM	44000400	900546000	magnat_mariam@gmail.com	31	M	moquegua	soltero	jefe de obra	4
22	21	14/05/2018	GARCIA SANCHEZ JAVIER	44730070	930343040	SANCHEZ_JAVIER_GARCIA@GMAIL.COM	25	M	moquegua	casado	jefe de almacén	3
23	22	14/05/2018	SILVEIRA ROSAMONDA MARIBEL	70600070	947470000	maribelrosamonda@gmail.com	30	M	moquegua	soltero	operador equipo liviano	3
24	23	14/05/2018	ALVARADO ALVAREZ MIHAIL ALEVIS	71329687	978346710	mihail24coz@gmail.com	25	M	moquegua	soltero	jefe de obra	4
25	24	14/05/2018	AGUILAR MIRALCO JUAN JOSÉ MIGUEL	73440604	946033334	juanmiralco30@gmail.com	30	M	moquegua	soltero	asistente	7
26	25	14/05/2018	LIMINGO OLIVER LUIS MIGUEL	47470040	934004070	limingo47luis@gmail.com	32	M	moquegua	soltero	técnico	5
27	26	14/05/2018	RAMOS ROSA JUAN CARLOS	40700010	901700746	ramosrosajuan@gmail.com	26	M	moquegua	casado	operario	7
28	27	14/05/2018	ARAZA GOMEZ CARLOS ALBERTO	35003666	956470046	carazaz1074@hotmail.com	35	M	moquegua	soltero	asistente	1
29	28	14/05/2018	OLIVER RAMOS DONALDO	74000000	943853434	Oliveramosdonaldo@gmail.com	31	M	moquegua	soltero	asistente	4
30	29	14/05/2018	VILLAR VILCA SEBASTIAN	40000004	947004000	sebastianvillar00@gmail.com	32	M	moquegua	soltero	operador equipo pesado	1
31	30	14/05/2018	VILLAR VILCA SEBASTIAN	40000004	947004000	sebastianvillar00@gmail.com	28	M	lima	soltero	asistente	1
32	31	14/05/2018	MOSCOCO GUTIERREZ TONY EDUARDO	70000000	950000001	tony.tong@gmail.com	32	M	arequipa	casado	operador equipo liviano	6
33	32	11/03/2018	FLORES MONTANA JAVIER ALEXANDER	40246700	902460000	jwachs0904@gmail.com	28	M	cusco	casado	operario	4
34	33	14/05/2018	ERNESTO NEMAS RICARDO MARIANO	43000000	902460000	ernesto.nemas2010@gmail.com	26	M	lima	casado	asistente	6
35	34	14/05/2018	JIMENEZ LABAN YUSSEF GALIAN	41242447	966000000	yussef1420@gmail.com	25	M	arequipa	casado	operador equipo pesado	6
36	35	14/05/2018	MINGOLUPI MECHEÑO DORE DEL CARMEN	41044445	900400000	mingolupimecheno30@gmail.com	34	M	cusco	casado	operador equipo liviano	4