

**UNIVERSIDAD CATÓLICA
DE SANTIAGO DE GUAYAQUIL**

**FACULTAD DE ECONOMÍA Y EMPRESA
CARRERA DE NEGOCIOS INTERNACIONALES**

TÍTULO:

**Implementación de machine learning para clasificación de clientes en la empresa
VALENTINA EXPORT S.A.**

AUTORES:

Delgado Moreira, Antonella Nicole

Moreno Dazza, Víctor Moreno

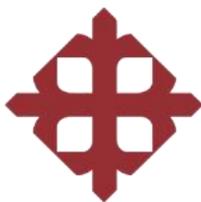
**Trabajo de integración curricular previo a la obtención de título de
LICENCIADO EN NEGOCIOS INTERNACIONALES**

TUTOR:

Ing. Arias Arana, Wendy Vanessa, Mgs.

Guayaquil, Ecuador

16 de febrero del 2024



UNIVERSIDAD CATÓLICA
DE SANTIAGO DE GUAYAQUIL

FACULTAD DE ECONOMÍA Y EMPRESA
CARRERA NEGOCIOS INTERNACIONALES

CERTIFICACIÓN

Certificamos que el presente trabajo de titulación fue realizado en su totalidad por **Delgado Moreira, Antonella Nicole y Moreno Dazza Victor Israel** como requerimiento para la obtención del título de **Licenciado en Negocios Internacionales**.

TUTOR

f. Wendy Arias

Ing. Arias Arana, Wendy Vanessa, Mgs.

DIRECTORA DE LA CARRERA

f. Gabriela Hurtado

Ing. Hurtado Cevallos, Gabriela Elizabeth, Mgs

Guayaquil, a los 16 días del mes de febrero del año 2024



UNIVERSIDAD CATÓLICA
DE SANTIAGO DE GUAYAQUIL

FACULTAD DE ECONOMÍA Y EMPRESA
CARRERA NEGOCIOS INTERNACIONALES

DECLARACIÓN DE RESPONSABILIDAD

Yo, Delgado Moreira, Antonella Nicole y Moreno Dazza, Víctor Israel

DECLARAMOS QUE:

El Trabajo de Titulación, **Implementación de machine learning para clasificación de clientes en la empresa VALENTINA EXPORT S.A.**, previo a la obtención del título de **Licenciados en Negocios Internacionales**, ha sido desarrollado respetando derechos intelectuales de terceros conforme las citas que constan en el documento, cuyas fuentes se incorporan en las referencias o bibliografías. Consecuentemente este trabajo es de mi total autoría.

En virtud de esta declaración, me responsabilizo del contenido, veracidad y alcance del Trabajo de Titulación referido.

Guayaquil, a los 16 días del mes de febrero del año 2024

LOS AUTORES:

f. _____

Delgado Moreira, Antonella Nicole

f. _____

Moreno Dazza, Víctor Israel



UNIVERSIDAD CATÓLICA
DE SANTIAGO DE GUAYAQUIL

FACULTAD DE ECONOMÍA Y EMPRESA
CARRERA NEGOCIOS INTERNACIONALES

AUTORIZACIÓN

Nosotros, Delgado Moreira, Antonella Nicole y Moreno Dazza Víctor Israel.

Autorizamos a la Universidad Católica de Santiago de Guayaquil a la publicación en la biblioteca de la institución del Trabajo de Titulación, **Implementación de machine learning para clasificación de clientes en la empresa VALENTINA EXPORT S.A.**, cuyo contenido, ideas y criterios son de nuestra exclusiva responsabilidad y total autoría.

Guayaquil, a los 16 días del mes de febrero del año 2024

LOS AUTORES:

f. _____

Delgado Moreira, Antonella Nicole

f. _____

Moreno Dazza, Víctor Israel



**UNIVERSIDAD CATÓLICA
DE SANTIAGO DE GUAYAQUIL
FACULTAD DE ECONOMÍA Y EMPRESA
CARRERA DE NEGOCIOS INTERNACIONALES**

REPORTE URKUND



Tesis final Victor Moreno Dazza -
Antonella Delgado Moreira



Nombre del documento: Tesis final Victor Moreno Dazza - Antonella Delgado Moreira.docx
ID del documento: 58f82aa090324349c91548dccb21fa09ece90cbd
Tamaño del documento original: 902,18 kB

Depositante: Wendy Vanessa Arias Arana
Fecha de depósito: 25/1/2024
Tipo de carga: interface
fecha de fin de análisis: 25/1/2024

Número de palabras: 23.847
Número de caracteres: 157.346

f. Wendy Arias

Ing. Arias Arana, Wendy Vanessa, Mgs.

TUTOR

-Antonella Nicole Delgado Moreira

AGRADECIMIENTO

A mis padres por haber forjado una mujer capaz de cumplir sus sueños, muchos de mis logros se los debo a ustedes por su apoyo incondicional, por nunca soltar mi mano para que yo logre cumplir cada una de mis metas, también a mi hermana por ser un pilar fundamental en mi vida y sobre todo a Dios quien me ha guiado y dado fortaleza para seguir adelante.

-V́ctor Israel Moreno Dazza

AGRADECIMIENTO

Quisiera expresar mi ḿs profundo agradecimiento por su inquebrantable apoyo y amor incondicional a lo largo de mi trayectoria acad́mica. Su constante presencia y aliento han sido la fuerza detŕs de cada paso que he dado hacia la consecuci3n de mis metas.

Padre y madre, son mi mayor inspiraci3n y el ejemplo para seguir en cada aspecto de mi vida. Su dedicaci3n y sacrificio han forjado el camino que hoy tengo el privilegio de recorrer.

Gracias por creer en ḿ incluso cuando dudaba de mis propias capacidades. Su confianza infalible me ha dado la fortaleza para enfrentar cada desaf́o con determinaci3n y perseverancia.

Este logro no solo es ḿo, sino tambi3n vuestro. Cada palabra de aliento, cada gesto de apoyo ha sido un pilar fundamental en mi camino hacia el 3xito acad́mico.

- Antonella Nicole Delgado Moreira

DEDICATORIA

A mi abuelo que fue una persona de admirar y que con sus consejos me motivó a luchar por los anhelos de mi corazón.

A mis padres, Antonio delgado y Amarilis Moreira por estar a mi lado en este camino que fue largo y arduo, por cada consejo que fue dado desde el amor, creyendo siempre en mí.

A mi hermana Fiorella que es una persona fundamental en mi vida y su incondicionalidad que es un estímulo constante.

DEDICATORIA

A mis amados padres, Víctor Hugo Moreno y Angela Dazza:

Su incansable dedicación y amor incondicional han sido mi mayor fortaleza a lo largo de este camino. Gracias por ser mis pilares y por guiarme con sabiduría en cada paso que doy.

A mi querida abuela Myriam:

Su cariño infinito y sus palabras de aliento han sido mi refugio en los momentos difíciles. Dedico este logro a ti, con profundo agradecimiento por su amor incondicional y por ser mi fuente de inspiración.

A mi adorada abuela Juanita:

Su valentía y su ejemplo de vida han dejado una huella imborrable en mi corazón. Este logro lo dedico a ti, con profunda admiración y gratitud por enseñarme a enfrentar los desafíos con coraje y determinación.

A mi amada abuela Nancy:

Su alegría y su eterna sonrisa han iluminado mi camino en cada momento. Gracias por enseñarme a valorar las pequeñas cosas que hacen la vida hermosa. Este logro lo dedico a ti, con profundo agradecimiento por tu amor desinteresado y por ser siempre mi fuente de inspiración.

-



**UNIVERSIDAD CATÓLICA
DE SANTIAGO DE GUAYAQUIL
FACULTAD DE ECONOMÍA Y EMPRESA
CARRERA DE NEGOCIOS INTERNACIONALES**

TRIBUNAL DE SUSTENTACIÓN

Gabriela Hurtado

Ing. Hurtado Cevallos, Gabriela Elizabeth, Mgs
DIRECTORA DE CARRERA

f. _____

(NOMBRES Y APELLIDOS)
COORDINADOR DEL ÁREA

f. _____

(NOMBRES Y APELLIDOS)
OPONENTE



**UNIVERSIDAD CATÓLICA
DE SANTIAGO DE GUAYAQUIL
FACULTAD DE ECONOMÍA Y EMPRESA
CARRERA DE NEGOCIOS INTERNACIONALES**

CALIFICACIÓN

f. _____

Delgado Moreira, Antonella Nicole

f. _____

Moreno Dazza, Víctor Israel

Guayaquil, a los 16 días del mes de febrero del año 2024

Implementación del Machine Learning para la clasificación de clientes en la empresa Valentina Export S.A.

INDICE

RESUMEN	XVI
ABSTRACT	XVII
RÉSUMÉ	XVIII
I. INTRODUCCIÓN	2
1. Antecedentes	3
2. Problemática	4
3. Justificación	5
4. Objetivos	6
4.1. Objetivo General	6
4.2. Objetivos Específicos	6
II. MARCO TEÓRICO	7
1. Importancia de los clientes en una empresa	7
2. Clasificación de los clientes	7
3. Exportación	8
4. Inteligencia Artificial	9
5. Machine Learning en las empresas	10
6. Herramienta RStudio	11
III. MARCO CONCEPTUAL	13
1. Machine Learning	13
2. Machine Learning para clasificación de los clientes	13
3. Regresión Logística	14
4. Técnicas de regresión logística	14
5. Importancia de los algoritmos en el Machine Learning	18
6. ¿Por qué es importante tener en cuenta el Machine Learning?	18
7. Aplicación de la regresión logística en la inteligencia artificial	18
IV. MARCO REFERENCIAL	20
V. MARCO METODOLÓGICO	23
1. Diseño de la Investigación	23
1.1. Tipo de investigación	23
1.2. Investigación Descriptiva	23
2. Enfoque de la investigación	24

2.1.	Enfoque Mixto.....	24
3.	Alcance de la investigación.....	24
3.1.	Población seleccionada.....	25
4.	Objetivos del estudio.....	25
4.1.	Optimización de Exportaciones:.....	25
4.2.	Mejora en la Toma de Decisiones:.....	25
4.3.	Maximización de Ingresos:.....	25
4.4.	Reducción del Riesgo Financiero:.....	25
4.5.	Eficiencia en la Gestión de Stock:.....	25
4.6.	Mejora en la Competitividad:.....	26
4.7.	Personalización en la Relación con el Cliente:.....	26
4.8.	Eficiencia Operativa:.....	26
4.9.	Beneficios Esperados:.....	26
5.	Consideraciones Importantes:.....	26
6.	Técnicas de Recopilación de Datos.....	27
6.1.	Entrevistas.....	27
6.2.	Aplicación de la Entrevista.....	27
6.3.	Análisis de la Entrevista.....	29
7.	Observación.....	30
7.1.	Aplicación de la Observación.....	30
VI.	ANÁLISIS DE DATOS.....	32
VII.	CAPITULO 1. ANALIZAR EL PROCESO ACTUAL DE ACERCAMIENTO A CLIENTES.....	34
1.	Proceso actual de acercamiento a clientes.....	34
2.	Descripción actual del proceso de capacitación para el acercamiento al cliente.....	36
2.1.	Problemas Identificados:.....	36
2.2.	¿Cómo ayudaría el modelo a la poca adaptabilidad y falta de integración?.....	36
3.	Abordaje de Riesgo financiero.....	37
4.	Afectaciones negativas de un mal acercamiento a los clientes y falta de estrategias.....	38
5.	Gestión ineficiente del stock debido a la mala organización con los clientes.....	40
6.	Clientes potenciales y crecimiento en el mercado.....	41
VIII.	CAPITULO II: ANÁLISIS Y DETERMINACIÓN DE LOS SEGMENTOS DE LOS CLIENTES.....	42
1.	Determinar los segmentos en los que se quiere clasificar a los clientes.....	42
2.	Importancia de la Clasificación.....	42
2.1.	Apto para Exportación.....	42

2.2.	Ventajas Estratégicas:.....	42
2.3.	No Apto para Exportación:.....	43
2.4.	Desventajas estratégicas:.....	43
3.	Variables por tomar en cuenta para la clasificación de clientes	43
3.1.	Liquidez Financiera.....	43
3.2.	Score Crediticio.....	44
3.3.	¿Por qué tomar en cuenta el score crediticio?	45
3.4.	Historial de Pago (35% del Puntaje)	45
3.5.	Porcentaje de Uso del Crédito (30% del Puntaje)	45
3.6.	Duración del Historial Crediticio (15% del Puntaje).....	45
3.7.	Tipos de Crédito (10% del Puntaje)	46
3.8.	Consultas de Crédito (10% del Puntaje).....	46
3.9.	Tamaño del pedido en dólares.....	46
3.10.	Ubicación Geográfica.....	47
3.11.	Costo de la logística en dólares	49
3.12.	Utilización logística eficiente del proceso.....	50
3.13.	Frecuencia de compra.....	51
IX.	CAPITULO III. IMPLEMENTACIÓN DEL MACHINE LEARNING	53
1.	Implementar el modelo de Machine Learning utilizando las variables con más significancia. 53	
1.1.	Identificación de datos faltantes	53
1.2.	Tratamiento de valores atípicos.....	53
1.3.	Manejo de duplicados.....	54
1.4.	Corrección de errores de entrada.....	54
1.5.	Transformación de datos	54
1.6.	Estandarización y normalización.....	54
X.	CAPÍTULO IV. ANÁLISIS DE LA ESTRATEGIA PARA LOS CLIENTES POR MEDIO DEL SCRIT.....	65
1.	Estrategias para los clientes por medio del Script	65
1.1.	Estrategia de fidelización a los clientes.....	65
1.2.	Estrategia de atención al cliente con una clasificación no apta financiera para exportar .. 68	
1.3.	Estrategia de venta a los clientes con una clasificación buena.....	69
	CAPÍTULO V. CONCLUSIONES Y RECOMENDACIONES	71
1.	Conclusiones	71
2.	Recomendaciones.....	72
XI.	BIBLIOGRAFÍA	73
	ANEXOS	79

Anexo 1. Base de datos	79
Anexo 2. Script en RStudio	81

Índice de Figuras

Figura 1 Modelo de regresión logística.....	15
Figura 2 Modelo de regresión logística binaria.....	15
Figura 3 Modelo de regresión multinomial.....	16
Figura 4 Utilización del sistema regresión logística como herramienta médica.....	18
Figura 5 Clasificación del valor según la fórmula	55
Figura 6 Tabla de coeficientes y sus significados en el Logit 1.....	58
Figura 7 Utilización del gráfico Q-Q del modelo sin optimizar	59
Figura 8 Gráfico de la distancia de Cooks existente en el modelo	59
Figura 9 Gráfico Q-Q del Modelo Optimizado	60
Figura 10 Gráfico de la Distancia de Cook existente en el Modelo Optimizado.....	61
Figura 11 Matriz de Confusión	63

RESUMEN

En esta tesis se detalla la aplicación de la regresión logística como una solución innovadora para segmentar a los clientes de Valentina Export S.A. La metodología utilizada implica la transformación de variables y la limpieza de datos para garantizar la calidad y precisión del modelo. La regresión logística se presenta como una herramienta clave para predecir la probabilidad de que un cliente sea "Apto para Exportar" o "No Apto para Exportar" en función de diversos parámetros, como historial de compras, liquidez financiera y tamaño del pedido.

La implementación práctica de este modelo no solo ha permitido optimizar la asignación de productos a clientes estratégicos, sino que también ha posibilitado la creación de estrategias más efectivas para la toma de decisiones y la atención personalizada al cliente. Este enfoque de machine learning no solo busca mejorar la eficiencia operativa de Valentina Export S.A., sino también brindar una visión informada para la toma de decisiones en un entorno de demanda dinámica y desafiante.

Palabras claves: Machine Learning, Clasificación de clientes, Regresión logística, Exportación.

ABSTRACT

This thesis details the application of logistic regression as an innovative solution to segment the customers of Valentina Export S.A. The methodology used involves the transformation of variables and data cleaning to ensure the quality and accuracy of the model. Logistic regression is presented as a key tool to predict the probability of a customer being “export eligible” or “export ineligible” based on various parameters such as purchase history, financial liquidity, and order size.

The practical implementation of this model not only helped optimize product allocation to strategic customers, but also created more effective strategies for decision-making and personalized customer service. This machine learning approach aims not only to improve the operational efficiency of Valentina Export S.A., but also to provide informed insight for decision-making in a dynamic and challenging demand environment.

Keywords: Machine Learning, Customer Classification, Logistics Regression, Export.

RÉSUMÉ

Cette thèse détaille l'application de la régression logistique comme solution innovante pour segmenter les clients de Valentina Export S.A. La méthodologie utilisée implique la transformation des variables et le nettoyage des données pour garantir la qualité et la précision du modèle. La régression logistique est présentée comme un outil clé pour prédire la probabilité qu'un client soit « éligible à l'exportation » ou « non éligible à l'exportation » en fonction de divers paramètres tels que l'historique des achats, la liquidité financière et la taille de la commande.

La mise en œuvre pratique de ce modèle a non seulement permis d'optimiser l'attribution des produits aux clients stratégiques, mais a également permis de créer des stratégies plus efficaces de prise de décision et de service client personnalisé. Cette approche d'apprentissage automatique vise non seulement à améliorer l'efficacité opérationnelle de Valentina Export S.A., mais également à fournir une vision éclairée pour la prise de décision dans un environnement de demande dynamique et stimulant.

Mots clés : Machine Learning, Classification des Clients, Régression Logistique, Exportation.

I. INTRODUCCIÓN

El presente proyecto aborda la implementación del Machine learning para la clasificación de clientes en la empresa Valentina Export S.A., con el objetivo de optimizar el proceso de acercamiento a los clientes y mejorar la eficiencia en la asignación de recursos. La empresa se enfrenta a desafíos significativos en la toma de decisiones sobre a qué clientes exportar, dada la alta demanda en comparación con la capacidad de producción del producto sembrado. La ausencia de una adecuada segmentación de clientes ha resultado en la subvaloración de ciertos clientes y en la falta de personalización de campañas de marketing, entre otros aspectos.

En este contexto, se propone la implementación de un modelo con Machine Learning para segmentar a los clientes según diferentes parámetros, lo que permitirá identificar mejor los puntos fuertes y clasificar los datos de manera más precisa. Se enfoca en analizar el proceso actual de acercamiento a clientes, determinar los segmentos en los que se quiere clasificar a los clientes, implementar el modelo de Machine Learning utilizando las variables con más significancia, y establecer estrategias para la atención a los clientes segmentados por el modelo. Se espera que la implementación de Machine Learning no solo mejore la eficiencia operativa y la toma de decisiones, sino que también aumente la competitividad de Valentina Export S.A. en el mercado nacional, al permitir la personalización de ofertas y mensajes, la identificación de oportunidades de ventas cruzadas, upselling y la adaptación proactiva a las tendencias cambiantes del mercado.

La estructura de la tesis se divide en cuatro capítulos, cada uno centrado en un aspecto crucial del proyecto. En el capítulo uno se aborda análisis exhaustivo del proceso actual de acercamiento a clientes en Valentina Export S.A. Se examinarán los métodos, prácticas y desafíos existentes en la interacción con los clientes, identificando las áreas de oportunidad y los obstáculos que puedan afectar la eficiencia y efectividad del proceso.

El capítulo dos, se enfoca en la determinación de los segmentos en los que se busca clasificar a los clientes y se establecen criterios claros para la segmentación. El capítulo tres se adentra en la implementación del modelo de Machine Learning, específicamente la regresión logística, utilizando las variables identificadas como más significativas. Se describirán

detalladamente los pasos tomados en RStudio, desde la preparación de datos hasta el entrenamiento y evaluación del modelo.

En el capítulo cuatro, se detalla las estrategias establecidas para la atención a los clientes segmentados. Esto incluye la personalización de servicios, la adaptación de enfoques y cualquier otra iniciativa estratégica que haya surgido como resultado de la clasificación de clientes.

1. Antecedentes

Exportadora de Productos Agrícolas Valentina Export S.A. es una empresa en Ecuador, con sede principal en Guayaquil. La empresa fue fundada el 01 de junio de 2018. Actualmente emplea a 14 personas, su planta principal se encuentra en Santa Ana- Manabí, adicional a esto cabe mencionar que, la empresa surgió inicialmente por el grupo de emprendedores dedicados a la exportación de productos del área de flores y harina de pescado, también al crecimiento socioeconómico y desarrollo sustentable del medio ambiente de su entorno, teniendo en cuenta el control y su misión que es mejorar la capacidad de los agricultores tecnificándolos para tener una producción agrícola sostenible y sustentable. Actualmente solo se desempeñan en el cultivo de la pitahaya.

Valentina Export S.A. pone a disposición a sus clientes una plataforma integral de servicios personalizados referentes a la incursión, crecimiento y consolidación de sus actividades de comercio exterior. Inicialmente la empresa se desempeñaba en la exportación del sector floricultura y a la industria panificadora, durante de la pandemia este proceso se dificultó para la exportadora y decidieron paralizarlo. Para finales del 2020 la empresa empezó con la idea de cosechar pitahaya y para el 2021 ya estaban vendiendo a nivel nacional esta fruta, sobre todo en la sierra ecuatoriana. Para el 2022 la idea de volver a lanzarse al mercado internacional empezó a plantearse esta vez apostando a la fruta del dragón o también conocida la pitahaya, ya que esta estaba ganando popularidad en el extranjero sobre todo en el continente europeo, empezaron a investigar este mercado que es muy difícil posicionar y mover todas sus cartas y contactos, así de esta manera se logró obtener un abanico extenso de clientes potenciales para la empresa.

Cabe recalcar que, VALENTINA EXPORT S.A durante todo el 2023 ha estado perfeccionando sus técnicas de marketing e invirtiendo en máquinas de calidad para el cultivo

de la pitahaya, porque prevén que para inicios del 2024 empiecen con las negociaciones y posteriormente con las exportaciones a diversos países como Francia, Italia, Dinamarca, Alemania y Suiza, ya que, son los mayores consumidores de esta fruta, sobre todo estos países convienen a la empresa y en las relaciones comerciales con Ecuador, aumentando conjuntamente la popularidad y el posicionamiento de VALENTINA EXPORT S.A en este continente. Finalmente, Ecuador al ser pionero en introducir la pitahaya en Europa en el 1999 opta por excelente reputación en mercados internacionales, aumentando la popularidad de la empresa.

2. Problemática

Valentina Export S.A hoy en día no posee un perfil o una segmentación de los clientes, lo cual en muchas ocasiones se ha visto reflejado en múltiples aspectos que resultan en subvalorar a ciertos clientes, poco conocimiento para personalización de campañas de marketing, etc. Mediante la implementación de un modelo con machine Learning podemos segmentar según diferentes parámetros al cliente para así poder identificar mejor muchos puntos fuertes con una correcta clasificación de los datos.

La empresa Valentina Export S.A se encuentra actualmente enfrentando desafíos significativos en la toma de decisiones sobre a qué clientes exportar, dado el panorama de una demanda extraordinariamente alta en comparación con la capacidad de producción del producto sembrado. La empresa ha adoptado un enfoque de exportación por orden de pedido, agotando el stock disponible para satisfacer la creciente demanda. Sin embargo, esta estrategia ha demostrado ser insostenible, ya que la oferta no puede abarcar a todos los clientes interesados.

Con el objetivo de optimizar este proceso y mejorar la eficiencia en la asignación de recursos, la empresa está considerando la implementación de herramientas avanzadas como el Machine Learning. En particular, se busca aplicar técnicas de regresión logística utilizando R Studio para desarrollar un modelo que permita priorizar a qué clientes exportar en función de diversos factores, como historial de compras, ubicación geográfica, tamaño del pedido y otros indicadores relevantes. Este enfoque no solo optimizará la asignación de productos limitados a los clientes más estratégicos, sino que también ofrecerá una visión más precisa de las tendencias del mercado, ayudando así a la empresa a tomar decisiones informadas y mantener su competitividad en el sector de exportación.

La ausencia de una adecuada segmentación de clientes en Valentina Export S.A plantea un desafío importante en cuanto a la asignación de recursos y estrategias. La empresa se encuentra en una posición donde no puede diferenciar eficazmente las necesidades y preferencias de sus clientes, lo que resulta en una falta de personalización en la oferta de productos y servicios. Esta falta de personalización puede llevar a la pérdida de clientes y a una disminución de la retención, lo que impacta negativamente en la rentabilidad y la reputación de la empresa en el mercado nacional. Sin una segmentación adecuada, la empresa puede tener dificultades para adaptarse a cambios en el comportamiento del cliente o en las tendencias del mercado. No podrá identificar oportunidades emergentes o amenazas potenciales de manera oportuna.

Se da a lugar a múltiples problemas en distintos ámbitos y da lugar a estrategias de marketing sin impacto con los diferentes segmentos de clientes. Esto puede resultar en un bajo retorno de la inversión en marketing y una falta de respuesta por parte de los clientes, sin ella, la empresa puede tener dificultades para predecir la demanda de productos y puede incurrir en costos adicionales debido a la falta de planificación. En un mercado competitivo, las empresas que pueden personalizar sus ofertas y servicios según las necesidades específicas de los clientes tienen una ventaja competitiva. La falta de segmentación puede hacer que Valentina Export S.A. sea menos competitiva frente a empresas que comprenden mejor a sus clientes y pueden satisfacer sus demandas de manera más efectiva, por ejemplo, si la empresa no sabe cuáles son los productos complementarios que un cliente podría necesitar, se perderá la oportunidad de aumentar el valor de la venta.

3. Justificación

Se propone la implementación de Machine Learning en la segmentación de clientes en Valentina Export S.A. Se fundamenta en la necesidad de la empresa de mejorar su capacidad para comprender y atender de manera más efectiva a sus clientes en un contexto nacional altamente competitivo. La segmentación de clientes desempeña un papel crucial en el éxito de cualquier organización, ya que permite la personalización de estrategias comerciales, la optimización de recursos y la retención de clientes valiosos.

La elección de Machine Learning como enfoque se justifica debido a su capacidad para manejar grandes volúmenes de datos de manera eficiente y su habilidad para identificar patrones complejos y sutiles en estos datos. En un entorno empresarial donde la información es

un activo valioso, el Machine Learning puede proporcionar una ventaja significativa al permitir la segmentación de clientes de manera precisa y oportuna. Además, su capacidad para aprender y adaptarse con el tiempo asegura que la segmentación se mantenga relevante a medida que evolucionan las preferencias y el comportamiento de los clientes.

El uso de Machine Learning en la segmentación de clientes no solo mejorará la eficiencia operativa y la toma de decisiones, sino que también puede aumentar la competitividad de Valentina Export S.A. en el mercado nacional. Esta tecnología proporcionará a la empresa una ventaja estratégica al permitir la personalización de ofertas y mensajes, la identificación de oportunidades de ventas cruzadas, Upselling (Upselling significa rebasar el límite de precio que se ha marcado un comprador), y la adaptación proactiva a las tendencias cambiantes del mercado

4. Objetivos

4.1. Objetivo General

- Implementar el uso del Machine Learning para la clasificación de clientes en la empresa Valentina Export S.A.

4.2. Objetivos Específicos

- Analizar el proceso actual de acercamiento a clientes.
- Determinar los segmentos en los que se quiere clasificar a los clientes.
- Implementar el modelo de Machine Learning utilizando las variables con más significancia.
- Establecer estrategias para la atención a los clientes segmentados por el modelo.

II. MARCO TEÓRICO

1. Importancia de los clientes en una empresa

Los clientes son personas que se consideran como los usuarios que utilizan los servicios o consumen los productos que comercializa una empresa o de otra persona. Esta se convierte en cliente cuando adquiere los servicios o productos de la empresa, generando en su accionar la rentabilidad del negocio, la promoción de las marcas, la posibilidad de crecimiento de la empresa, el ofrecer el feedback que la empresa necesita para determinar su crecimiento y determinar el proceso de camino como parte del progreso para la empresa (Rodríguez, 2023). La clasificación de los clientes es una estrategia con la cuales las empresas pueden determinar a su público objetivo mediante determinadas características que los definen. Entre las características existentes se encuentra el estatus, el volumen de compra que puede realizar, la antigüedad de compra, la frecuencia de compra, entre otros aspectos que se consideran para determinar el comportamiento del consumidor (UPBE, 2023).

Los clientes consisten en las personas que están dispuestas a pagar por el servicio de una empresa o adquirir los productos que esta misma comercializa. Los clientes también pueden ser entidades que compren los bienes y servicios para desarrollar una determinada actividad, sea esta de consumo o comercio. Cada empresa posee un distinto tipo de clientes según el tipo de mercado en el que se encuentren y con el que la empresa aplique su proceso de competición. Entre los tipos de clientes que existen se encuentran el mercado de consumos, el mercado industrial, el mercado de los revendedores, entre otros. Los clientes son fundamentales para las empresas debido que, por sus acciones de compra de los productos y servicios, las empresas obtienen ingresos, siendo esta su razón de existencia buscando la satisfacción de los clientes o solucionar un determinado problema o carencia (Godás, 2005).

2. Clasificación de los clientes

La clasificación de los clientes es un aspecto fundamental en el desarrollo empresarial debido que, mediante determinados criterios, las empresas pueden establecer su modelo de desarrollo empresarial basadas determinadas características. Los clientes pueden ser clasificados por causa de distintos criterios que son el estatus, el volumen de compra, la antigüedad, la frecuencia de compra, el tipo de producto que compra, entre otros. El proceso de clasificación de los clientes dependerá de actividades como el conocimiento e identificación de

las cualidades de los clientes. Así como el conocimiento de las técnicas de comunicación y habilidades sociales que fomenten la empatía entre la empresa y los consumidores. La importancia de clasificar a los clientes está relacionada con el establecimiento de los procesos comunicativos que sean directos y permitan alcanzar al consumidor determinado según el tratamiento adecuado del cliente en la empresa (Alcázar, 2023)

La clasificación de los clientes también es conocida como el proceso de segmentación del mercado. Y las empresas la utilizan esta estrategia para diferenciar su mercado y potenciar las labores de la organización para conocer de mejor manera los criterios con los cuales la empresa comenzará el desarrollo de nuevas estrategias de comunicación y comercio de sus productos o servicios. La segmentación de los clientes es una oportunidad que garantiza la intervención de estrategias que permitan desarrollar actividades basados los esfuerzos y los recursos utilizados para que la empresa pueda potenciar su imagen de marca, su crecimiento y sostenibilidad en el mercado.

Los tipos de clasificación de las empresas más utilizados se realiza por medio del estatus. Entre los diferentes tipos de estatus se encuentran los clientes actuales, los clientes activos, los clientes inactivos, los clientes potenciales y los clientes probables. Existen varias formas de clasificar a los clientes como utilizar su frecuencia de compra como clientes frecuentes, clientes habituales y los clientes ocasionales. Los criterios que se utilizan para clasificar a los clientes dependerán de la segmentación en función a los objetivos de rentabilidad por el cliente, la segmentación para garantizar una estrategia de formatos adecuada, segmentación para optimizar el surtido, alcanzar nuevos productos o servicios en el mercado. O simplemente como una estrategia de redimensionamiento para beneficio de la segmentación empresarial (Artyco, 2017)

3. Exportación

La exportación es el concepto que se otorga a la acción de exportar un producto elaborado con mano de obra, materia prima y dentro de un territorio nacional, para ser consumido en el mercado internacional. El concepto es utilizado principalmente en economía para definir la compra y venta de los bienes y servicios desde un país hacia el extranjero. Gracias a la exportación se puede predecir el comportamiento del comercio nacional establecido (Harvard, 2015)

Las exportaciones son fundamentales para el crecimiento económico de las empresas, debido que las exportaciones promueven el desarrollo económico de un país desde los años 50s, hasta la actualidad, gracias a la globalización. Este mismo proceso de globalización es el responsable de generar la formación de los bloques comerciales existentes entre los distintos países para reducir las barreras al comercio exterior. La importancia de las ventas en el exterior permite aumentar los niveles de crecimiento económico, tal como ocurrió con Japón, incrementando su capacidad de producción gracias a sus políticas de proteccionismo y el aumento de su capacidad de producción, principalmente de tecnología para fomentar el crecimiento de la productividad. Las naciones suelen medir su crecimiento económico por medio del estudio de la Balanza Comercial, relacionada con las importaciones y las exportaciones, para medir aspectos como la producción, la salida de productos y la salida y entrada de capital (Toledo, 2017).

El proceso de exportación conlleva a una serie de operaciones logísticas relacionadas con las operaciones del departamento de exportación para establecer los tiempos y los costos que conlleva la actividad. Antes que nada, es necesario registrar la empresa como una productora nacional, que posee oferta exportable y una determinación importante para el seguimiento de los acuerdos con las normas de comercio internacional vigente. Se deberá de contar también con los permisos por parte del Registro nacional de Importadores y Exportadores de la Dirección General de las Aduanas para gestionar labores arancelarias como la identificación del tipo de producto, su país de origen y los requisitos que se deben de cumplir para que cada producto pueda ingresar al mercado objetivo. Además, se deberá de contar con el Registro Único Tributario que permita el comercio de los productos en el mercado internacional (Páramo, 2017).

4. Inteligencia Artificial

La inteligencia artificial consiste en un campo de la informática que ha sido desarrollado con el objetivo de simular la inteligencia del ser humano por medio de herramienta tecnológicas. La Inteligencia Artificial se trata de una composición de elementos como el software y el hardware, diseñados por parte de los seres humanos para simular el pensamiento del ser humano por medio de una herramienta tecnológica. La inteligencia artificial es capaz de percibir su entorno y adquirir conocimiento por medio de la interpretación de los datos. Son los sistemas de Inteligencia Artificial los encargados de mediante los datos obtenidos establecer un

razonamiento sobre el conocimiento, el proceso de la información derivada y la toma de decisiones sobre las mejoras acciones que se pueden tomar para lograr los objetivos dados. En pocas palabras, la inteligencia artificial es capaz de aprender por medio de algoritmos y modelos matemáticos que garantizan el procesamiento de grandes cantidades de información sobre los patrones y las reglas establecidas (INCyTU, 2018).

Entre los tipos de Inteligencia Artificial más utilizados se encuentran las herramientas de Software y la inteligencia artificial integrada. La primera es aquella que se compone de un entorno virtual, posee software de análisis de imágenes y posee motores de búsqueda o sistemas de reconocimiento de voz, rostro y huellas digitales. Mientras tanto, la inteligencia artificial integrada, es aquella que su software funciona en piezas tecnológicas como robots, drones o autos. La Inteligencia Artificial puede ser clasificada según su procedencia, la cual se puede determinar cómo débil, fuerte o superinteligente. La inteligencia artificial débil es conocida como la IA estrecha, siendo esta un sistema diseñado para realizar tareas específicas y limitadas. Las IA fuertes son aquellas que pueden razonar, planificar y tomar decisiones complejas de forma autónoma. Y las IA superinteligentes, son aquellas que superan a la inteligencia humana, pero que aún no han sido diseñadas (Gobierno de España, 2023)

5. Machine Learning en las empresas

La aplicación del Machine Learning en las empresas, se fomenta como una herramienta que permite el desarrollo de nexos dinámicos existentes entre los negocios y los clientes, para permitir el desarrollo de los negocios, principalmente las Pymes dentro del mercado virtual. La herramienta es fundamental para promover la venta de los productos y los servicios prediciendo las tendencias comerciales y desarrollar estrategias que fomenten la resolución de los problemas de ciberseguridad. Las empresas que utilizan en la actualidad las herramientas de Machine Learning son las empresas que necesitan aplicar estrategias de soluciones integrales de Machine Learning. Entre los beneficios que existen sobre el sector empresarial se encuentran la anticipación del futuro, el eliminar posibles errores en la estrategia de desarrollo y gestión, a la promoción de la automatización de los sistemas, el ofrecimiento de productos y servicios personalizados, el mejorar la capacidad estratégica empresarial, el ahorrar costos operativos, el mejorar el servicio de los clientes y el desarrollar mayores oportunidades de negociación.

Las herramientas de Machine Learning se han convertido en estrategias fundamentales gracias a que estas herramientas ahorran los costos operativos, evitando posibles errores en la

toma de decisiones y el proceso de comercialización de un producto. La herramienta busca perfeccionar la atención con el cliente debido que en el 89% de los casos se ha demostrado que mientras mejor sea el trato con una determinada empresa, mayor será su interés de invertir su capital en adquirir un determinado producto o servicio. Por ende, los clientes que sean mejor atendidos mediante el uso de herramientas de Machine Learning, será mayor su capacidad de crecimiento gracias a su incremento en las ventas, logrando obtener mayores oportunidades de negociación (Datisión, 2021).

6. Herramienta RStudio

Una herramienta de RStudio consiste en un software informático que utiliza el análisis estadístico para el diseño de gráficos para estudiar determinados factores por parte del investigador. El software ha sido desarrollado por medio de un lenguaje de programación libre, el cual es utilizado para fomentar la curva de aprendizaje de los usuarios debido a lo interactivo que es su utilización. Este lenguaje no necesita que el usuario disponga de amplios conocimientos de programación ni de informático, estando más dirigido al uso de los matemáticos, permitiendo que pueda maquetear o desarrollar prototipos en modelos ya realizados, aumentando la capacidad de desarrollo de sistemas informáticos enfocados a las necesidades de cada usuario.

El sistema utiliza una consola de R, en la cual permite que el usuario se pueda comunicar con el programa y desarrollar mediante comandos sencillos su propio lenguaje de programación, resolviendo operaciones como sumar, restar, dividir, o desarrollar funciones matemáticas complejas como lo serían las funciones logarítmicas, trigonométricas y demás. La interfaz de usuario RStudio funciona como un lenguaje de programación en la cual se fomenta el Studio de un Entorno de Desarrollo Integrado, que fomenta un manejo del sistema más cómodo para los usuarios (González, 2019).

Este sistema es altamente utilizado y popular para la realización de modelos estadísticos como son la regresión logística, hasta la similitud de los sistemas informáticos que se generan mediante las redes neuronales. La herramienta también es capaz de graficar las funciones para que el sector empresarial pueda tomar decisiones más acertadas según los datos utilizados en su modelo de gestión.

Los sistemas de RStudio permiten la implementación de las regresiones logísticas, fundamentales para determinar el comportamiento de los consumidores. Gracias a sus capacidades, son utilizados como una herramienta dentro del estudio cuantitativo, permiten predecir, explicar o determinar el comportamiento de una variable dependiente por medio del estudio de una variable independiente (Ahumada, 2003).

III. MARCO CONCEPTUAL

1. Machine Learning

El Machine Learning consiste en el desarrollo de una disciplina la cual ha sido creada sobre el campo de la Inteligencia Artificial. Para su funcionamiento depende del uso de algoritmos y ordenadores con los cuales es posible la identificación de la información y determinar sus propios resultados por medio de la programación previa. Para el desarrollo de una herramienta de Machine Learning es necesario aplicar de forma estratégica el uso de los algoritmos con los cuales se analizarán los datos. El objetivo es que el sistema mediante una revisión de datos previamente aplicados pueda comenzar a determinar su propio sistema de procesamiento con el modelo de entrada y salida de información, generando así un sistema predictivo en la creación del modelo que causa el proceso de interacción que simula el proceso de aprendizaje interactivo (IBM, 2023).

Existen tres tipos de Machine Learning existentes los cuales se determinan por el tipo de procesamiento de los datos. Se encuentran el aprendizaje supervisado, el aprendizaje no supervisado y el aprendizaje de refuerzo mediante la naturaleza de los datos. En el caso del aprendizaje supervisado, se caracteriza por ser un modelo de datos que es capaz de aplicar la compresión de la información y calificar los datos dentro del proceso de aprendizaje, siguiendo determinados patrones para aplicar sobre el proceso y análisis de la información (Stack, 2021).

2. Machine Learning para clasificación de los clientes

Las herramientas digitales como el Machine Learning permiten desarrollar un proceso de segmentación de los clientes con una eficiencia del 73%, permitiendo reconocer los productos, servicios y las experiencias propias de los consumidores. Gracias a la herramienta de Machine Learning es posible determinar de manera eficiente los análisis de los datos de los clientes y perfilar así de manera más objetiva su capacidad para utilizar sobre las características determinadas de cada cliente (Muller, 2023). La intervención de un sistema de Machine Learning se podría implementar basado dos maneras. La primera intervención sería mediante el sistema de aprendizaje supervisado. Gracias a esta estrategia se pueden determinar las reglas y las configuraciones implementadas dentro del sistema de operaciones y salida de los resultados. Si se aplica la segunda opción de un sistema no supervisado, el Machine Learning

será capaz de crear por sí mismo sus propios criterios para analizar los datos y determinar posibles patrones de comportamiento entre los consumidores (Duckerman, 2021)

3. Regresión Logística

La regresión logística consiste en la implementación de unas técnicas estadísticas con las que se puede medir los datos y determinar las relaciones existentes entre dos factores. Una vez que se determina la relación entre las variables, se podrá predecir el valor en la utilización de los factores por medio de su relación y estudio entre las mismas, utilizando la predicción que existen entre las variables y sus resultados finitos. La importancia en la utilización de las regresiones logísticas está relacionada con su importancia en la utilización en el campo de la inteligencia artificial y Machine Learning, para la realización de actividades complejas relacionadas con el procesamiento de los datos, permitiendo que el sector empresarial se pueda beneficiar de la herramienta para obtener información sobre sus clientes, prediciendo su comportamiento y clasificando a los mismos según determinadas características.

La principal ventaja de la regresión logística es la simplicidad en la utilización, permitiendo la utilización de cálculos matemáticos que no son altamente complejos en comparación con otros métodos. Su velocidad es un factor clave, permitiendo que los modelos de regresión logística tengan la capacidad de procesar grandes volúmenes de datos y por ende los sistemas informáticos no deben ser tan complejos. Además, la flexibilidad de la regresión logística permitirá su utilización para encontrar determinados resultados que sean finitos y ordenarlos según el rango de los valores (Amazon, 2024).

4. Técnicas de regresión logística

La herramienta de utilización logística funcionará como una variable dependiente de datos cualitativos, permitiendo el uso de dos o más categorías de datos pertenecientes a la regresión logística multinomial. Así también se pueden utilizar variables explicativas conocidas como covariables. La regresión logística posee hasta tres finalidades para su utilización. La primera tiene relación con la definición existente entre las covariables y la variable dependiente. La segunda es el análisis entre las covariables y la variable dependiente, determinando la interacción y confusión. Y la tercera está relacionada con la clasificación de los individuos dentro de las respectivas categorías de las variables dependientes.

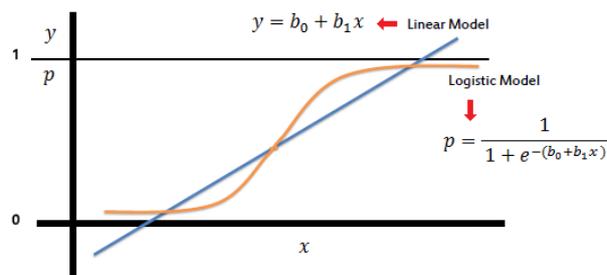
La fórmula de regresión logística es:

$$P(Y) = \frac{1}{1 + e^{-(b_0 + b_1x_1 + b_2x_2 + \dots + b_nx_n)}}$$

Demostrando la cualidad de Y como la probabilidad que ocurra una función exponencial sobre el resto de los coeficientes análogos dependientes de la regresión logística. Los valores existentes sobre esta variable tendrán relación sobre el valor significativo de 0, siendo este altamente improbable de que ocurra. Si el valor es cercano a 1, significará que es altamente probable que se genere la acción (Ferre, 2019).

Figura 1.

Modelo de regresión logística

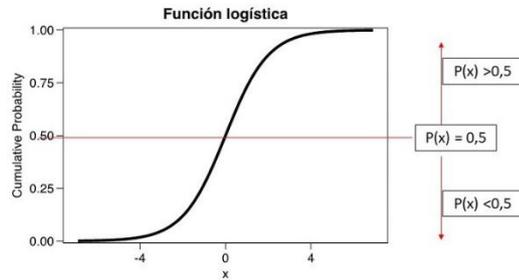


Nota. Representación gráfica de un modelo de regresión logística. Recuperado de (Torres, 2024)

Existen varios tipos de regresión logística que son la regresión logística binaria, la regresión logística multinomial, y la regresión logística ordinal. La regresión logística binaria se determina principalmente para la clasificación binaria de los resultados posibles que se generen sobre la variable dependiente. Para ello, se podrían en la fórmula generarse solo dos resultados posibles que son 0 – 1. Las respuestas deberán redondearse a estos valores y según la cercanía se determinaría el resultado posible. La regresión logística binaria tiene mayor relación entre las variables dependiente cualitativa dicotómica y las variables independientes o explicativas (Elsevier, 2010).

Figura 2.

Modelo de regresión logística binaria

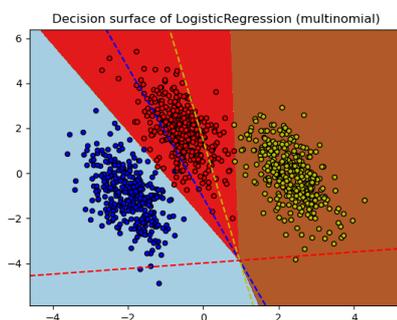


Nota. Representación gráfica de una regresión logística binaria. Recuperado de (Ortega, Ochoa, & Molina, 2022)

El modelo de regresión logística multinomial está relacionado con los problemas que existen entre los resultados posibles, considerando que los resultados sean finitos. Su utilización depende del uso de los valores de los resultados con distintos valores, siendo estos entre 0 a 1, analizando así el rango de los valores continuos. Es la regresión logística la herramienta que permite agrupar los resultados de los valores más cercanos posibles (Amazon, 2024).

Figura 3.

Modelo de regresión multinomial



Nota. Representación gráfica de un modelo de regresión multinomial. Recuperado de (Scikit Learn, 2020)

La regresión logística ordinal está relacionada con el modelo Logit ordenado, el cual es una herramienta de regresión multinomial que permite representar un número determinado de rangos en lugar de los valores reales, permitiendo sobre ello predecir de forma exacta la respuesta de los valores obtenidos en herramientas como encuestas basadas las herramientas como escala de Likert que permiten ser analizadas en función de valores numéricos.

La regresión logística se aplica como una estrategia de análisis matemático en la cual se analizan los datos y se establece una proyección de los valores sobre la variable que sea desconocida o dependiente de la variable conocida. Las regresiones son importantes para medir la relación entre dos o más variables. Al permitir el análisis del comportamiento de una variable se la puede utilizar para determinar el comportamiento de la variable que en este caso sería el comportamiento del consumidor, basada ciertas características que permiten que la empresa pueda mejorar su proceso de toma de decisiones sobre la venta y exportación de los productos agrícolas como la Pitahaya (Vicente, 2021).

El aprendizaje automático se refiere al estudio y desarrollo de algoritmos y técnicas de aprendizaje automático a nivel conceptual, centrándose en fundamentos teóricos, diseño algorítmico y análisis matemático en lugar de detalles de implementación específicos o dominios de aplicación. Su objetivo es proporcionar una comprensión más profunda de los principios fundamentales y las limitaciones del aprendizaje automático, lo que permitirá a los investigadores desarrollar algoritmos novedosos y avanzar en el campo. En el aprendizaje automático abstracto, el énfasis está en formalizar y analizar tareas de aprendizaje, desarrollar modelos matemáticos para procesos de aprendizaje y estudiar las propiedades y el comportamiento de varios algoritmos de aprendizaje. (Pandey, 2023)

Dentro de la IA, el aprendizaje automático (Machine Learning) consiste en la elaboración de programas de ordenador que sean capaces de generar conocimiento a través del análisis de los datos. El conocimiento obtenido será la base para realizar deducciones sobre nuevos datos. Algunos métodos dentro del aprendizaje automático tienen un carácter explicativo (inducción de reglas y árboles de decisión), mientras que otros se caracterizan por un enfoque de caja negra (Black box), como las redes neuronales. (Miranda, Ramos, & Segovia, 2013)

5. Importancia de los algoritmos en el Machine Learning

Los algoritmos de Machine Learning son el alma que mueven los procesos de aprendizaje. Gracias a ellos podemos obtener la información que necesitamos para tomar decisiones o predecir el comportamiento de los datos. Existen diferentes tipos de algoritmos y estos son aplicables dependiendo del tipo de Machine Learning en el cual el algoritmo vaya a funcionar. (Graph Everywhere, 2022)

6. ¿Por qué es importante tener en cuenta el Machine Learning?

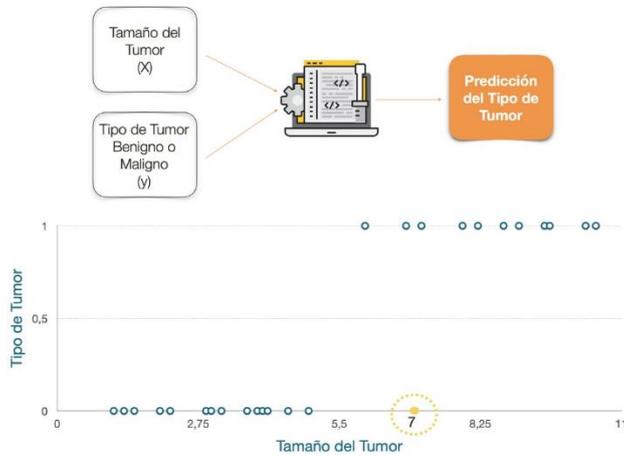
Porque las empresas del futuro utilizarán cada vez más esta tecnología, que está siendo decisiva en la transformación de los procesos de trabajo tal y como los conocemos. Como empresarios, emprendedores y aspirantes a puestos directivos de relevancia, no debemos olvidar que cada vez es mayor el volumen de datos que generamos, que hoy disponemos de herramientas para recogerlos, almacenarlos y procesarlos, y que podemos tomar decisiones más inteligentes, basadas en algoritmos de predicción, y automatizar un mayor número de funciones dentro de la empresa (Cámara Madrid, 2019)

7. Aplicación de la regresión logística en la inteligencia artificial

La aplicación de la regresión logística en la Inteligencia Artificial funcionará como un algoritmo que tendrá relación importante para clasificar las variables, siendo utilizada como una herramienta de estas condiciones hasta en el 70% de los casos. A pesar que ya existen varias herramientas de clasificación de variables, la regresión logística es la herramienta que obtiene los mejores resultados. Su importancia radica en la facilidad de ser implementados para la clasificación de los datos, tomando en cuenta la clase de los datos. Una vez que se obtengan los resultados, será fácil definir su clasificación, siendo la clasificación binaria la forma más utilizada para medir los resultados. El modelo logístico es utilizado como una herramienta que permite estimar la probabilidad de la respuesta binaria que se basa en una de las variables predictoras o variables independientes generadas (Larispardo, 2021)

Figura 4.

Utilización del sistema regresión logística como herramienta médica



Nota. Uso de la regresión logística para medir la probabilidad de formación de un tumor en un paciente en un caso médico. Recuperado de (Gonzalez, 2019)

Gracias al análisis de la regresión logística sobre el análisis predictivo se logra establecer la variable de los datos y explicar la posible relación que existe sobre la variable binaria dependiente y la variable independiente nominal, ordinal y utilizando elementos como los intervalos de razón. Gracias a ello, es posible clasificar en la realización de exámenes como el tamaño del tumor y la probabilidad que este afecte la calidad de vida de un paciente.

IV. MARCO REFERENCIAL

El uso de las herramientas de Machine Learning se ha vuelto una estrategia competitiva que beneficia de forma importante el rendimiento operacional de las empresas. Este cambio en el comportamiento de las empresas es un cambio de la necesidad generado por la revolución de la industria 4.0, la cual incluyen herramientas como IoT, Blockchain, y el uso del resto de herramientas de digitalización que forman parte de la ramificación de la Inteligencia Artificial como lo sería el uso del Machine Learning, el Deep Learning y el Cognitive Intelligence. Entre las empresas que dependen del uso del Machine Learning en la actualidad se encuentran la empresa Nestlé, la cual utiliza los sistemas Machine Learning para fomentar el desarrollo de sus productos con un 60% de la producción de la empresa.

El uso de la herramienta de inteligencia artificial va de la mano con la inversión de la Inteligencia Artificial que depende del aprendizaje automático para mejorar su rendimiento de I+D y generar resultados positivos en su desarrollo empresarial. Nestlé utiliza sus herramientas para desarrollar un sistema de planificación de los procesos para el desarrollo de nuevos productos, el cual actualmente solo genera una inversión de tiempo de seis meses desde que comienza la planificación de la idea. El desarrollo de los productos nuevos está basado en la necesidad de los consumidores. La Inteligencia Artificial ofrece información respecto al fitomejoramiento, la extracción de datos clínicos y el aseguramiento de la calidad de la materia prima (Baking Business, 2022).

Una de las marcas más conocidas que realiza la clasificación de los clientes de manera eficiente es Amazon. Esta empresa desarrolla una estrategia donde se utilizan los Ads en la web, permitiendo descifrar el comportamiento de los consumidores mediante el uso de las plataformas digitales. Para Amazon, existen dos tipos de compradores que son el comprador Normal y el comprador Prime. Los compradores normales son los usuarios que se registran de forma gratuita en la plataforma y compran un producto que les resulta interesante. El cliente paga el envío y recibe su producto. Las compras son poco habituales. No obstante, los compradores Prime son aquellos que la empresa considera que realizan compras anuales que superan los \$800.00 en el año. Para que Amazon implementara una clasificación eficiente de los clientes, fue necesaria la implementación de las herramientas digitales para aumentar su rentabilidad comercial y desarrollar un mejor sistema de información (Ceyca, 2019)

El sector avícola de Europa se ha desarrollado gracias a la implementación de la tecnología, principalmente del Machine Learning para fomentar un óptimo desarrollo comercial mediante la recopilación de los datos de interés que permitan afrontar los desafíos diarios generados en la inteligencia artificial y aumentar sus posibilidades para desarrollo de un sistema de gestión sostenible con mayor bienestar para los animales. El Software de Inteligencia Artificial se utiliza en las granjas avícolas para permitir el control de los animales y lograr que estos crezcan de forma sana tanto para el animal como para el desarrollo de un producto avícola de calidad para el consumidor.

La herramienta de Inteligencia Artificial es capaz de comprender el comportamiento de las aves. La herramienta permite detectar posibles anomalías en las aves. La herramienta digital ha comprobado que las aves clasifican su energía de varias maneras, utilizando mayor energía para su metabolismo y capacidad de producción. Por ejemplo, cuando las gallinas se enferman con bronquitis o influenza aviar, la forma en la que cacarean es diferente. La herramienta permite detectar la enfermedad que padece el ave solo comprobando su tipo de cacaero. La herramienta de inteligencia artificial permite también contar los huevos y determinar el estado de estos para garantizar la calidad del producto (Avinews, 2023)

En el Ecuador, el uso de las herramientas de Machine Learning han permitido que el sector pueda desarrollarse de una forma más eficiente, gracias a que utilizan lenguaje de programación y herramientas digitales para comprender el comportamiento del consumidor. Uno de los sectores que se ha comenzado a desarrollar gracias a la intervención de la tecnología es el sector camaronero, el cual utiliza lenguajes de programación Python, para analizar el comportamiento de los consumidores y establecer la cantidad de producción de camarón que la empresa debe realizar en un periodo de tiempo determinado. Gracias al uso de estas herramientas tecnológicas, es posible que el Ecuador pueda mantenerse e incluso aumentar en su volumen de exportación de camarón del mundo, aumentando la cantidad de compradores del producto que, en la actualidad, el mayor rubro económico es generado por Estados Unidos, que son la cabeza de consumo del mercado del camarón. La herramienta de Machine Learning también es capaz de predecir aspectos relacionados al precio del camarón, el análisis de las exportaciones del camarón para determinar la cantidad que se comercializará posteriormente en el segundo periodo comercial. En el sistema de análisis de la información se utiliza la herramienta de Statmodels, el cual es un modelo de regresión logística y se proyectan los parámetros que el sistema ha ido aprendiendo junto a una serie de datos con los cuales se basa

la teoría de los mínimos cuadrados. En cuanto al paquete de Sklearn es otro modelo de estudio de regresión logística en el cual se analizan los pronósticos existentes en los datos para revelar las estimaciones en los datos conexos (Cordero, 2021).

V. MARCO METODOLÓGICO

1. Diseño de la Investigación

1.1. Tipo de investigación

Para la realización del diseño de un modelo de aprendizaje automático se realizó las respectivas investigaciones de la necesidad existente de un algoritmo, que ayude a través de la identificación de patrones complejos en millones de datos, los cuales permitirán generar los patrones óptimos para la toma de decisiones por parte de los ejecutivos de la empresa Valentina Export S.A.

1.2. Investigación Descriptiva

Este estudio se dirige fundamentalmente, en caracterizar un fenómeno o situación concreta indicando sus rasgos más peculiares o diferenciadores (Morales, 2012). La investigación descriptiva hace referencia cuando se estudia un objeto existente y se busca evolucionarlo, recopila la información obtenida y en base a esta realiza la manifestación de la investigación a realizar. La presente investigación es de carácter descriptivo, ya que, se considerará oportuno realizar un acercamiento a lo que actualmente ha impulsado los procesos internos que lleva la empresa en torno a los clientes y su rentabilidad a la hora de tomar decisiones para una mejor elección de candidatos a exportar, Cabe mencionar que, Valentina Export S.A. ha adoptado un enfoque de exportación el cual puede ser relativamente identificable durante la investigación.

Adicionalmente las diferentes variables existentes de los clientes permitirán evaluar, analizar el entorno y la delimitación del objeto de estudio de la presente investigación el cual es la clasificación de los clientes para una toma de decisiones más rentable. Estos datos se tomaron de las bases de datos gestionados y revisados por la empresa las cuales arrojaban un perfil muy superficial de los clientes, porque los datos a pesar de ser muy útiles no eran tomados en cuenta, puesto que, se atiende a los clientes de manera en que llegaban los pedidos, lo cual crea un riesgo financiero muy alto por no tomar en cuenta datos cruciales a la hora de examinar y evaluar al cliente.

2. Enfoque de la investigación

A través del enfoque o alcance de la investigación se busca definir la estrategia operativa a través de la cual se accederá a los datos fidedignos y su respectiva evaluación o valoración de los mismos. El enfoque permitirá de manera oportuna diseñar eficientemente la estructura procedimental de la investigación y sus respectivas técnicas de recopilación de datos a aplicar.

2.1. Enfoque Mixto.

El enfoque cualitativo también se guía por áreas o temas significativos de investigación. Sin embargo, en lugar de que la claridad sobre las preguntas de investigación e hipótesis preceda a la recolección y el análisis de los datos. (Sampieri, Fernández & Baptista, 2014). El enfoque cualitativo hace referencia a las características de cada proceso que se está estudiando, en el caso particular del presente proyecto son pocos los perfiles a los que se considera realmente necesario evaluar de forma directa con el fin de obtener información relevante que ayude a entender cada proceso, este enfoque ayuda a realizar un análisis profundo. También, se adoptó un enfoque cuantitativo, respaldado por análisis de datos y técnicas de Machine Learning. La solicitud de un registro de ventas, historial de compras y otros datos específicos se alinea con nuestro compromiso de utilizar información cuantitativa para optimizar nuestras estrategias comerciales. La recopilación y análisis de datos cuantitativos nos permiten identificar patrones, tendencias y correlaciones que son fundamentales para tomar decisiones informadas.

3. Alcance de la investigación

El alcance de la tesis se centra en la implementación de técnicas de Machine Learning para la clasificación de clientes en el contexto de una empresa de exportación de pitahaya. La finalidad principal de esta implementación es optimizar y fortalecer el proceso de toma de decisiones comerciales, brindando herramientas efectivas para la identificación y gestión de clientes. Además, la clasificación de clientes a través de Machine Learning contribuirá a una gestión más efectiva del riesgo crediticio. Al evaluar de manera continua la salud financiera de los clientes, el sistema podrá proporcionar alertas tempranas sobre posibles riesgos, permitiendo a la empresa tomar decisiones informadas sobre términos de crédito y límites financieros.

3.1.Población seleccionada

La población considerada para el presente estudio viene dada por la data obtenida de los casi 200 clientes que actualmente maneja la empresa Valentina Export S.A. a partir de la cual se espera con la aplicación del estudio derivar el tipo de cliente más adecuado y óptimo para una exportación más eficiente y rentable.

4. Objetivos del estudio

4.1.Optimización de Exportaciones:

Identificar y priorizar a los clientes más estratégicos y aptos para la exportación, maximizando la eficiencia en la asignación de productos limitados.

4.2.Mejora en la Toma de Decisiones:

Proporcionar a la empresa una herramienta de toma de decisiones basada en análisis predictivos y en el comportamiento pasado de los clientes, permitiendo decisiones más informadas y estratégicas.

4.3.Maximización de Ingresos:

Aumentar los ingresos al enfocarse en exportar a clientes con mayor potencial de crecimiento y rentabilidad, contribuyendo al rendimiento financiero global de la empresa.

4.4.Reducción del Riesgo Financiero:

Mitigar el riesgo de impagos y pérdidas financieras al clasificar a los clientes con una mayor probabilidad de cumplir con los términos de pago establecidos.

4.5.Eficiencia en la Gestión de Stock:

Optimizar la gestión de inventario al exportar productos a clientes con una demanda más predecible y consistente.

4.6.Mejora en la Competitividad:

Fortalecer la posición competitiva de Valentina Export S.A en el mercado al adaptarse de manera ágil y precisa a las demandas cambiantes y comportamientos de los clientes.

4.7.Personalización en la Relación con el Cliente:

Facilitar la personalización en las relaciones comerciales al comprender mejor las preferencias y comportamientos individuales de los clientes.

4.8.Eficiencia Operativa:

Mejorar la eficiencia operativa al automatizar y agilizar el proceso de toma de decisiones relacionado con las exportaciones.

4.9.Beneficios Esperados:

- Identificación más precisa de clientes con mayor potencial de compra.
- Reducción de pérdidas por exportaciones a clientes de riesgo financiero
- Mejora en la planificación de inventario y logística.
- Mayor satisfacción del cliente al recibir productos de manera más oportuna.
- Incremento en la rentabilidad global de las operaciones de exportación.

5. Consideraciones Importantes:

- La implementación de este modelo requiere colaboración interdepartamental y capacitación del personal involucrado en la toma de decisiones.
- La evaluación continua del modelo y su ajuste según sea necesario es esencial para mantener su eficacia a lo largo del tiempo.
- La transparencia en la comunicación con los clientes respecto al proceso de clasificación y toma de decisiones contribuirá a mantener relaciones comerciales sólidas.

6. Técnicas de Recopilación de Datos

Las técnicas de recopilación de datos aplicadas en el presente estudio permitirán ayudar en el progreso de la investigación ya que nos proporciona información relevante de los procesos ya sea a través de diálogos o evaluaciones, con la finalidad que esta información sea útil para el desarrollo de la investigación. Bajo el alcance mixto que se seleccionó y explicó en el bloque anterior se ha considerado la aplicación de las técnicas de Entrevistas y Observación no estructurada.

6.1. Entrevistas.

Desde la perspectiva de Folgueiras (2016) “La entrevista es una técnica de recogida de información que además de ser una de las estrategias utilizadas en procesos de investigación, tiene ya un valor en sí misma” “Tanto si se elabora dentro de una investigación, como si se diseña al margen de un estudio sistematizado, tiene unas mismas características y sigue los pasos propios de esta estrategia de recogida de información” (Folgueiras, 2016). Es decir, la entrevista es un método de recopilación de datos que por medio de un diálogo se logra obtener información, se tomó en consideración dicho método ya que permite realizar un análisis profundo de los procesos que manejan la empresa Valentina Export S.A. para luego procesar la información, detectando cual será de mayor utilidad para la realización del proyecto.

6.2. Aplicación de la Entrevista

Para la aplicación de la entrevista se ha considerado a los diferentes perfiles que podrían ser de mayor aporte al propósito de recabar la mayor información posible del contexto del objeto de estudio para su óptima identificación y análisis bajo un estudio Machine Learning

Perfil de entrevistado #1

- Entrevistado: Sr. Jhon Saavedra
- Cargo: Gerente de Sistemas de Información
- Lugar de Trabajo: Oficinas en Manabí-Ecuador de la empresa Valentina Export S.A
- Aporte Cualitativo:

Al hablar con el encargado del área de tecnología de la información ayudó a entender cómo se utilizan las herramientas tecnológicas para gestionar la información de clientes, stock y procesos de exportación.

Perfil de entrevistado #2

- Entrevistado: Sr. Jhon Moreira
- Cargo: Gerente de Logística
- Lugar de Trabajo: Oficinas en Manabí-Ecuador de la empresa Valentina Export S.A
- Aporte Cualitativo:

El gerente de logística pudo ofrecer información sobre cómo se gestiona el stock y la cadena de suministro. Dio a conocer los desafíos logísticos, cómo se determinan los niveles de stock y cómo se asegura la eficiencia en la entrega.

Perfil de entrevistado #3

- Entrevistado: Sr. Cesar Quiroz
- Cargo: Gerente de Ventas
- Lugar de Trabajo: Oficinas en Manabí-Ecuador de la empresa Valentina Export S.A
- Aporte Cualitativo:

Esta persona fue clave a la hora de proporcionar información sobre cómo se gestionan las relaciones con los clientes durante el proceso de ventas y exportación. Mediante la entrevista realizada se logró reconocer las estrategias de venta, la interacción con los clientes y también los criterios utilizados para decidir a quién exportar.

6.3. Análisis de la Entrevista

Estos roles representan diferentes perspectivas dentro de la empresa y pueden ofrecer una visión integral de cómo Valentina Export S.A aborda la exportación, la gestión de clientes y el manejo de stock. A continuación, se expone a modo de resumen objetivo las diferentes opiniones y aseveraciones emitidas por los entrevistados considerados en el levantamiento de información, donde cada uno de sus aportes son vitales para el desarrollo óptimo de la presente investigación:

- Perspectivas sobre las estrategias de venta y cómo se aborda la exportación desde el punto de vista comercial.
- Criterios utilizados para identificar clientes estratégicos y potencialmente rentables
- Información sobre las relaciones y comunicación con los clientes durante el proceso de exportación
- Detalles sobre cómo se gestiona el stock y se planifica la cadena de suministro.
- Desafíos logísticos específicos y cómo se abordan.
- Información sobre la eficiencia en la entrega de productos y cómo se minimizan los tiempos de espera.
- Identificación de las herramientas y sistemas de información utilizados para gestionar datos de clientes, stock y exportaciones.
- Información sobre si se utiliza o no la analítica avanzada y el análisis de datos para extraer conocimientos significativos y respaldar la toma de decisiones.
- Identificación de desafíos específicos en términos de tecnología de la información y cómo se abordan en el contexto de exportación y gestión de clientes.

7. Observación

7.1. Aplicación de la Observación

Registro de observación de investigación

1. Contexto de la Observación:

- **Fecha y Hora de Inicio:** 16 de noviembre a las 14:00h
- **Ubicación:** Sede Principal de Guayaquil de la empresa Valentina Export S.A.

2. Participantes:

- **Número Total:** 5
- **Descripción de los Participantes:** Se le preguntó y pido información relevante sobre las operaciones de exportación, financieras y logísticas que se llevaban a cabo en la empresa actualmente al gerente de sistemas de información, Gerente de logística y la Gerente de ventas de Valentina Export S.A

3. Instrumentos de Observación:

- **Herramientas Tecnológicas:** Excel, Word, Correo, Reportes Financieros

4. Procedimiento de Observación:

- **Duración:** 2 horas
- **Descripción de la Observación:** Se recolectó documentos de Excel con datos relevantes de los clientes.

5. Registros de Observación:

- **Hora (hh:mm):** 14:00 pm del 14 de noviembre del 2023
- **Descripción detallada:** Excel para organizar la información recopilada en tablas estructuradas. Cada fila representa a un cliente y las columnas contienen datos específicos, como edad, género, historial de compras, etc.

6. Resultados Preliminares:

- **Patrones Identificados:** Durante el proceso de observación, identifiqué patrones como clientes que realizan compras frecuentes, aquellos que muestran preferencia por ciertos productos o servicios, y clientes inactivos.
- **Observaciones Relevantes:** Se observó y recopiló datos relacionados con los clientes de la empresa, incluyendo información demográfica, historial de compras, frecuencia de interacción, etc.

7. Reflexiones del Investigador:

- **Desafíos encontrados:** Uno de los desafíos iniciales fue la calidad y consistencia de los datos recopilados durante la observación. La presencia de valores atípicos, datos faltantes o inconsistencias que requirieron un proceso de limpieza de datos exhaustivo para garantizar la precisión del futuro modelo a implementar
- **Reflexiones Personales:** Explicar los resultados del modelo de manera comprensible para los interesados y garantizar que estos comprendan y confíen en las recomendaciones generadas por el Machine Learning podría haber sido un desafío clave en una implementación exitosa.

8. Conclusión:

- **Resumen de Resultados:** Se logró recopilar y estructurar eficientemente datos relevantes sobre los clientes mediante la técnica de observación.
- **Acciones Futuras:** Realizar integración exitosa del modelo de Machine Learning en los procesos de toma de decisiones empresariales para que la empresa pueda centrar sus esfuerzos en atender a ciertos segmentos de clientes que tienen una mayor probabilidad de realizar la importación de pitahaya de una manera más eficiente y rentable, que evite el riesgo financiero, clientes morosos, etc. Tener una comunicación efectiva de los resultados para asegurar la aceptación y comprensión por parte de los trabajadores y responsables de la empresa.
- **Firma del Investigador:** Víctor Moreno, Antonella Delgado.

VI. ANÁLISIS DE DATOS

El proceso de la implementación del Machine Learning para la clasificación de clientes, empieza por la creación de una base de datos unificada uniendo todas las variables que solicita la empresa a sus clientes, una vez creada se realiza un proceso de "Data Cleaning" o limpieza de datos, es una etapa esencial en la gestión y análisis de cualquier base de datos. Consiste en identificar, corregir y eliminar errores, inconsistencias y datos irrelevantes o duplicados, con el objetivo de garantizar la calidad y confiabilidad de la información contenida en la base de datos.

Luego se procederá con el entrenamiento de los datos con una partición de la data en un 80/20 para evitar un Overfitting (cuando el modelo considere como validado, solo los datos que se han usado para entrenar el modelo, sin reconocer ningún otro dato que sea un poco diferente a la base de datos inicial) y garantizar una clasificación y predicción lo más optima posible, mediante la función "Summary()" aplicado al modelo Logit se podrá determinar las variables con mayor impacto e importancia a tomar en cuenta, aquellas variables las cuales son más determinantes para establecer si un cliente es apto para la exportación.

Se creará un segundo modelo "Logit" llamado "Logit2", en el que indicamos que las variables antes mencionadas de gran porcentaje de significancia son las que se requiere para obtener un modelo más optimizado y preciso para así asegurar una clasificación más precisa. Tras la eliminación de las variables sin significancia considerable, se deberá observar un impacto positivo en los resultados del modelo "Logit2." Al enfocarse en las variables que realmente tienen un impacto significativo en la variable objetivo, el modelo se vuelve más robusto y capaz de proporcionar resultados coherentes y confiables en escenarios fuera de los datos de entrenamiento.

Una vez obtenido un modelo ajustado y optimizado (Logit2) se podrá apreciar mediante el grafico "Q-Q (QuantileQuantile) normal" el cual proporciona una indicación visual de si los datos siguen una distribución normal o no. En este caso mostrará si los puntos se ajustan bien a la línea diagonal o no, lo cual deja ver que si están ajustados a la línea es más probable que los datos sigan una distribución normal.

Mediante el grafico de Distancia de Cook podremos identificar observaciones que pueden tener un impacto considerable en la robustez y validez del modelo. Una vez examinado

detenidamente el gráfico de distancia de Cook generado a partir del modelo (Logit2). Al hacerlo, se observaría si el modelo cumple con lo siguiente:

- **Puntos Dentro de Umbrales Razonables:** Los puntos en el gráfico de distancia de Cook se encuentran mayoritariamente dentro de umbrales razonables, lo que sugiere que las observaciones individuales no están teniendo un impacto excesivo en los resultados del modelo.
- **Ausencia de Valores Atípicos Significativos:** No se han identificado valores atípicos extremos en el gráfico que podrían distorsionar el rendimiento general del modelo.
- **Estabilidad en los Coeficientes:** La influencia de las observaciones en los coeficientes de regresión parece ser relativamente estable, lo que indica que el modelo no está siendo impulsado en gran medida por un conjunto reducido de observaciones.

Luego de este proceso se calculará la probabilidad de éxito de clasificación mediante la función predict () la cual crea una columna en nuestra base de datos que mostrará la probabilidad de exportar o no a cada cliente. Utilizando los valores de probabilidad generados por la función antes mencionada, se establece un umbral de decisión los cuales serán:

$$\geq 0.5 - 1.0 = 1 \quad < 0.5 - 0.0 = 0$$

Los clientes cuyas probabilidades superan este umbral fueron clasificados como “1”, se consideran más aptos para recibir un crédito, ya que tienen una mayor probabilidad de cumplir con sus pagos. Por otro lado, los clientes cuyas probabilidades están por debajo del umbral se clasificaron como “0” ya que podrían no cumplir con los criterios de elegibilidad y, por lo tanto, se considera más riesgoso realizarles la exportación.

VII. CAPITULO 1. ANALIZAR EL PROCESO ACTUAL DE ACERCAMIENTO A CLIENTES.

1. Proceso actual de acercamiento a clientes.

Valentina Export S.A, se lleva a cabo un proceso de acercamiento a los clientes para la exportación de pitahaya, el cual estaba basado en una serie de pasos para establecer y mantener relaciones sólidas con clientes nuevos y clientes fieles. Para captar la atención de nuevos clientes, realizan análisis de mercado periódicos y participan activamente en ferias comerciales y eventos del sector. Utilizan estrategias de marketing digital, redes sociales y campañas de correo electrónico para llegar a posibles clientes, destacando la calidad, la disponibilidad constante y las ventajas competitivas de sus productos.

Cuando identifican a un cliente potencial, su equipo de ventas inicia el contacto mediante llamadas telefónicas o correos electrónicos personalizados. En estas comunicaciones, dependiendo de lo que desee el cliente en ese momento, de esta manera se despejan dudas y se llega a un punto en específico en el proceso de compra. Cabe recalcar que este desarrollo se trata de llevar de una manera más personal, ya que, lo amerita una situación como lo es el de ventas y exportación. Se presentan de manera detallada las ofertas de pitahaya, destacando las certificaciones de calidad, la logística eficiente y la experiencia en exportación. Además, se esfuerzan por comprender las necesidades específicas de cada cliente, adaptando su propuesta para satisfacer sus requisitos particulares.

Está demostrado que llevar el proceso de esta manera crea fidelidad de parte del cliente porque se aprecia el interés de la empresa por responder cualquier tipo de inquietud y solucionar de manera óptima y eficaz, llevándolos así a posicionarse poco a poco con consumidores pequeños y medianos. Para los clientes fieles, mantienen una comunicación regular y proactiva. Siguen un sistema de seguimiento que incluye actualizaciones periódicas sobre nuevos productos, promociones especiales y noticias relevantes del mercado. Esto ayuda a mantener a los clientes informados y comprometidos con la marca. Además, solicitan retroalimentación constante para asegurarse de que están cumpliendo con sus expectativas y para identificar posibles áreas de mejora. El proceso de compra con los clientes que ya son fijos se lleva a cabo por correo con su respectiva cotización, ya que, como son recurrentes se sabe a dónde se envía el producto y de qué manera desean que llegue el mismo. Este proceso es muy ágil porque los

encargados en el área de servicio al cliente son eficaces en cuanto al tiempo de responder las necesidades del mismo.

Durante el proceso de comunicación tanto con clientes fieles y nuevos, se recopila información clave que permite a la empresa comprender mejor las necesidades del cliente y evaluar su idoneidad como socios comerciales. Uno de los elementos esenciales encontrados en el proceso de recopilación de datos, es la solicitud de un registro de ventas, que proporciona una visión detallada de la actividad comercial del cliente. Este registro ayuda a identificar patrones de compra, preferencias de productos y niveles de demanda, lo que es crucial para adaptar la oferta a sus requerimientos específicos.

El historial de compras también desempeña un papel fundamental al proporcionar una perspectiva a largo plazo sobre la relación comercial con el cliente. Esto permite evaluar la consistencia en las transacciones, identificar productos más populares y anticipar posibles necesidades futuras. Este enfoque histórico contribuye a una planificación estratégica más efectiva y a una relación comercial más sólida. Así mismo, se pide ubicación geográfica y el tamaño del pedido que son factores determinantes en la logística y distribución eficientes de los productos. Conocer la ubicación geográfica del cliente permite optimizar la cadena de suministro y garantizar entregas oportunas. Además, comprender el tamaño del pedido les ayuda a ajustar los procesos de producción y gestión de inventario para satisfacer las demandas específicas de cada cliente.

La empresa durante todo el proceso en el cual se comunica con el cliente, también, pide una evaluación financiera del mismo, solicita un informe comercial y crediticio, así como detalles sobre su capacidad de endeudamiento y flujo de efectivo. Estos documentos son esenciales para evaluar la estabilidad financiera del cliente y mitigar posibles riesgos crediticios. La revisión de estados financieros proporciona una comprensión más profunda de la salud financiera de la empresa, permitiendo tomar decisiones informadas sobre términos de pago y límites de crédito.

Además de los aspectos financieros, la reputación del cliente en el mercado es un factor crucial. Buscan referencias y retroalimentación de otras empresas con las que hayan trabajado para evaluar la confiabilidad y la calidad de la relación comercial.

En resumen, el enfoque integral hacia el contacto con clientes nuevos y fieles refleja que hay un compromiso continuo con la excelencia y la satisfacción del cliente en la exportación de pitahaya.

2. Descripción actual del proceso de capacitación para el acercamiento al cliente.

En la actualidad, el proceso de capacitación en Valentina Export S.A. presenta ciertos desafíos que han afectado la eficacia del aprendizaje y la aplicación de conocimientos por parte de los trabajadores. La capacitación se lleva a cabo mediante sesiones teóricas y prácticas, pero se ha identificado una falta de integración con las tecnologías emergentes y una escasa adaptabilidad por parte de los empleados.

2.1. Problemas Identificados:

- **Falta de Integración Tecnológica:**

La capacitación no incorpora de manera efectiva las tecnologías emergentes, incluyendo el modelo de Machine Learning propuesto. Los empleados carecen de conocimientos especializados para comprender y utilizar estas herramientas en su trabajo diario.

- **Enfoque Genérico:**

El enfoque de capacitación tiende a ser genérico y no se adapta completamente a las necesidades específicas de cada departamento. Esto resulta en una aplicación limitada de los conocimientos adquiridos, especialmente en áreas críticas como la gestión del stock y las relaciones con los clientes.

- **Falta de Evaluación Continua:**

La ausencia de una evaluación continua y específica para medir el progreso de los empleados dificulta la identificación de áreas de mejora. La falta de retroalimentación individualizada limita la eficacia de la capacitación.

2.2. ¿Cómo ayudaría el modelo a la poca adaptabilidad y falta de integración?

La implementación de un modelo de Machine Learning para clasificar a los clientes en Valentina Export S.A. no solo beneficiaría la gestión de clientes, sino que también facilitaría la

adaptabilidad de los empleados al proporcionar una herramienta más accesible y comprensible mediante una capacitación adecuada.

- **Claridad en la Identificación de Clientes:**

El modelo de Machine Learning puede simplificar el proceso de clasificación de clientes al automatizar la identificación de patrones y características clave. Esto permite a los empleados comprender de manera más clara y eficiente las distintas categorías de clientes, ya que el modelo realizará gran parte del análisis complejo.

- **Facilitación del Proceso de Toma de Decisiones:**

Con un modelo bien capacitado, los empleados pueden tomar decisiones informadas basadas en la clasificación predicha por el Machine Learning. Al entender los criterios utilizados por el modelo para clasificar a los clientes, los trabajadores pueden adaptar estrategias y enfoques específicos para cada segmento, mejorando así la efectividad en la gestión de relaciones comerciales.

- **Mejora en la Eficiencia Operativa:**

La capacitación centrada en el modelo de Machine Learning permitirá a los empleados comprender cómo utilizar eficientemente las predicciones del modelo en su trabajo diario. Esto se traduce en una mayor eficiencia operativa, ya que los procesos de clasificación y adaptación de estrategias se vuelven más rápidos y precisos.

- **Simplificación de la Interacción con el Modelo:**

La capacitación sobre el uso e interpretación del modelo de Machine Learning simplificará la interacción de los empleados con la herramienta. Al comprender cómo se generan las clasificaciones y cómo interpretar los resultados, los trabajadores se sentirán más seguros y capaces de integrar el modelo en sus actividades cotidianas.

3. Abordaje de Riesgo financiero

Para abordar este riesgo de pérdidas financieras, es esencial implementar un enfoque más estratégico y analítico en la asignación de productos a los clientes. La introducción de un modelo de Machine Learning proporciona una herramienta valiosa para evaluar de manera más

precisa la solvencia de los clientes, identificando aquellos con una mayor probabilidad de cumplir con los términos de pago y reduciendo así el riesgo financiero asociado con las exportaciones. Este enfoque proactivo contribuirá a salvaguardar la estabilidad financiera y la sostenibilidad a largo plazo de Valentina Export S.A en el competitivo mercado de exportación.

La falta de adaptación a cambios en el mercado obstaculiza la capacidad de la empresa para ajustar rápidamente su estrategia de precios, oferta de productos o enfoque geográfico según las condiciones cambiantes. Esto, a su vez, impacta negativamente en la rentabilidad y en la capacidad de mantener una posición competitiva sólida. La asignación por orden de llegada, al ser un enfoque reactivo, deja a la empresa rezagada frente a la competencia que puede estar implementando estrategias más dinámicas y orientadas al mercado. La capacidad de respuesta limitada también hace que la participación de Valentina Export S.A. en nuevos segmentos de mercado o en la introducción de nuevos productos se reduzca, ya que la asignación de productos no se adapta de manera proactiva a las demandas emergentes.

Para abordar este desafío, la implementación de un modelo de Machine Learning proporcionará una herramienta efectiva para analizar patrones históricos, identificar tendencias emergentes y prever cambios en la demanda del mercado. Esto permitiría a Valentina Export S.A. ajustar proactivamente su estrategia de exportación, enfocándose en clientes y mercados que representen las mejores oportunidades en un entorno de negocios dinámico, lo cual puede conducir a una desigualdad en la demanda de productos.

4. Afectaciones negativas de un mal acercamiento a los clientes y falta de estrategias.

La falta de estrategia afecta la planificación de la producción, ya que la empresa puede producir en exceso. Esto conlleva a costos innecesarios y una cadena de suministro menos eficiente. Los problemas logísticos y de almacenamiento también surgen debido a la falta de una visión clara de la demanda esperada, generando desafíos en la coordinación logística y en la optimización del espacio de almacenamiento.

Este enfoque no estratégico también tiene un impacto directo en los costos de almacenamiento. Mantener la producción de pitahaya en stock sin una planificación adecuada aumenta los costos de almacenamiento, reduciendo la rentabilidad general. Las dificultades para atender las demandas específicas de los clientes surgen, ya que la empresa puede no tener

los productos adecuados disponibles, lo que afecta la satisfacción del cliente y las oportunidades de venta.

El problema de la empresa es no saber clasificar a los clientes por su alta demanda y porque no todos pueden ser atendidos de manera eficaz, es por eso que como se mencionó anteriormente existe un riesgo alto por no decidir cuáles son los más aptos. Cabe recalcar que la situación actual de la empresa es inestable en cuanto a la exportación porque no tienen un manejo correcto en esa área del comercio internacional y la comunicación con los clientes se deteriora es por eso que su acercamiento no es el más óptimo, ya que, otro factor que complica la situación comercial de la empresa es el exceso de clientes inestables a los que se les da prioridad en gran o menor parte.

Además, la pitahaya tarda en cosechar entre 110 a 120 días, este tiempo influye mucho ya que los pedidos por los clientes son numerosos y se receptan con anticipación, pero la empresa no hace un análisis adecuado del perfil del cliente, por lo que, a veces presenta retrasos en el pago de las facturas. Por esta razón, se busca eliminar y seleccionar a los mejores candidatos, ya que sin solvencia económica VALENTINA EXPORT S.A presenta atrasos en el pago de mensualidad de sus obligaciones con sus trabajadores y proveedores.

Lo que se busca es establecer estrategias de fidelización a los clientes no solo como un factor de marketing sino también con la calidad del producto, ya que de esta manera al elegir un proveedor de pitahaya recurren directamente a VALENTINA EXPORT S.A y se posicionan en el mercado internacional. Por consiguiente, también se busca determinar los segmentos en los que se quiere clasificar a los clientes, adicionalmente cabe mencionar que el mayor consumo de esta fruta se encuentra en Europa y Estados Unidos. Por otro lado, en el ámbito demográfico se centran en todo tipo de clase social, edad, género, cultura o religión, ya que esta posee un consumo elevado en el extranjero.

Es importante determinar la importancia de la planificación y estrategias de las ventas y lo que esto influye positivamente a la exportadora, por lo que, en el entorno actual cada día surgen nuevos competidores y cambia la forma en que los consumidores satisfacen sus necesidades, sobre todo es imperativo que cada unidad de fabricación gestione la planificación de ventas y estudie continuamente los factores que influyen en el mercado. Es por esto que VALENTINA EXPORT S.A como factor importante debe examinar las características del plan de ventas y el análisis de mercado al que se quieren dirigir, además de emplear un plan

estratégico de ventas de una división de fabricación que establece un pronóstico de ventas aproximado de pitahaya y describir las metas y estrategias de ventas clave para llevarlas a cabo de la mejor manera posible, como se mencionó anteriormente es esencial lograr una buena relación con el cliente para que este proceso siempre sea un ganar-ganar.

Esto también incluye estudiar el volumen de ventas y el presupuesto de ventas, objetivos fijos y previsiones de ventas, este plan de ventas debe ser táctico, progresivo e incremental para cubrir múltiples períodos futuros, ya que la demanda puede cambiar con el tiempo, por lo que es necesario establecer una planificación de la misma, porque la pitahaya o también llamada fruta del dragón tiene como mayor época de consumo en el extranjero en el verano por su ola de calor, pero en el invierno baja considerablemente las ventas.

5. Gestión ineficiente del stock debido a la mala organización con los clientes.

Una gestión ineficiente de stock puede resultar en desperdicio de recursos, ya que se producen excesos, lo que puede ocasionar la pérdida de la calidad o que se eche a perder. Esto no solo tiene implicaciones operativas, sino que también puede conducir a pérdidas financieras significativas. La falta de una estrategia sólida dificulta la adaptación a cambios en la demanda del mercado, dejando a la empresa vulnerable a fluctuaciones imprevistas y limitando su capacidad para responder ágilmente a nuevas oportunidades comerciales. Como resultado, el stock producido de pitahaya puede agotarse rápidamente, dificultando una gestión equitativa de inventario.

La falta de una estrategia clara en la recepción de pedidos contribuye a la desorganización. Al ordenar los pedidos por orden de llegada, sin tener en cuenta criterios como el comportamiento de pago del cliente o el volumen de la orden, se generan dificultades en la gestión de inventario y en la preparación para la demanda real. Una previsión de la demanda altamente precisa es la única forma en que pueden predecir qué país necesita los productos, en este caso cuál de los clientes sería más factible proveer, lo que a su vez es la única manera de garantizar una alta disponibilidad por parte de VALENTINA EXPORT S.A para los clientes, así al mismo tiempo el nivel mínimo de la demanda de acuerdo al stock que presente la empresa.

Es importante mencionar que el llevar un orden en la recepción de pedidos es importante, en VALENTINA EXPORT S.A estos se ordenan por orden de llegada, ya que, de ellos depende la urgencia de este. Así, se evita malentendidos con los clientes y en el peor de

los casos, retrasos. También hay que recalcar que como la empresa quiere empezar a exportar nuevamente la pitahaya en este caso se llevará a cabo la recepción de pedidos por el comportamiento del cliente en los pagos y en la cantidad porque obviamente se prevé tener un abanico de clientes internacionales extenso, pero con un gran volumen de pedido.

6. Clientes potenciales y crecimiento en el mercado.

Las ventas de la empresa y su reconocimiento en la sierra ecuatoriana van en aumento, aunque, como se mencionó anteriormente los compradores mayoristas actuales no poseen el mejor perfil, pues la fruta es de gran interés para las personas minoristas por lo que esta se vende de manera eficaz.

Es de gran relevancia mencionar que el comportamiento del consumidor recae en esta fruta exótica por su calidad e imagen que proyecta en el mercado, VALENTINA EXPORT S.A se quiere enfocar en estos grandes consumidores porque primeramente poseen una facilidad de pago sin retrasos, esto significa mucho para la empresa porque así de esta manera no se tiene complicaciones con los ingenieros a cargo del seguimiento de la cosecha de la fruta.

También es de conocimiento que llegar al mercado internacional es un gran paso para la empresa, ya que, como se mencionó en los antecedentes inicialmente la empresa exportaba flores y harina de pescado, pero, hace dos años dejó de hacerlo por problemas con el stock del producto y por la pandemia. Con la implementación del Machine Learning en sus procesos se prevé mejorar el posicionamiento con los clientes y volver al ruedo en el mercado extranjero con la respectiva innovación y agilidad en el proceso de cosecha de la pitahaya. Además de esta manera al elegir los mejores perfiles se podrá realizar con anticipación los respectivos pedidos, ya que, como son en gran volumen así se agiliza y evita cualquier tipo de retrasos.

En este caso se tiene previsto que al momento de la selección de los futuros clientes el proceso de la recepción de pedidos sea más automatizado, mediante una extensión telefónica única para cada cliente para que al momento de la llamada sea contestada con rapidez, adicionalmente por el cambio de horario en muchos países chocan por lo que de esta manera se trabajará 24 horas. También hay que mencionar que, si no es por llamada telefónica también habrá un encargado por país para la venta de pitahaya de esta manera se llega a abarcar el mercado.

VIII. CAPITULO II: ANÁLISIS Y DETERMINACIÓN DE LOS SEGMENTOS DE LOS CLIENTES

1. Determinar los segmentos en los que se quiere clasificar a los clientes

La clasificación tiene como objetivo facilitar la toma de decisiones estratégicas en relación con las exportaciones. Los clientes clasificados como "Apto para Exportación" podrían beneficiarse de oportunidades internacionales, mientras que los clasificados como "No Apto para Exportación" podrían requerir una evaluación adicional o estrategias comerciales alternativas. En resumen, la división de clientes en los segmentos "Apto para Exportación" y "No Apto para Exportación" se basa en la capacidad del modelo de regresión logística para analizar y aprender patrones a partir de los datos de entrenamiento, proporcionando así una herramienta útil para la toma de decisiones a la hora de atender clientes que podría representar tanto, un riesgo financiero, como un negocio rentable.

2. Importancia de la Clasificación

La distinción entre "Apto para Exportación" y "No Apto para Exportación" es crucial para tomar decisiones estratégicas. Permite focalizar la prioridad en aquellos clientes que tienen una base financiera sólida y están mejor preparados para aprovechar las oportunidades internacionales. El modelo de regresión logístico en RStudio será entrenado para clasificar la idoneidad de los clientes para exportación. Esta clasificación se logra dividiendo a los clientes en dos segmentos distintos:

2.1. Apto para Exportación

- Los clientes que son asignados a este segmento contarán con características y comportamientos financieros que sugieren una solidez financiera y una capacidad para cumplir con los requisitos y desafíos asociados con la exportación.

2.2. Ventajas Estratégicas:

- **Negociación de Términos Favorables:** Estos clientes pueden beneficiarse de condiciones comerciales más favorables, como plazos de pago extendidos o límites de crédito más altos, gracias a su posición financiera robusta.

- **Contribución al Éxito de Exportación:** Se espera que la participación de estos clientes en actividades de exportación contribuya al éxito general de las operaciones de exportación de la empresa.

2.3.No Apto para Exportación:

- Los clientes clasificados en este segmento serán los que posean características que podrían indicar mayores riesgos o limitaciones financieras para participar en actividades de exportación.

2.4.Desventajas estratégicas:

- **Negociación de Condiciones más Restringidas:** Estos clientes pueden requerir términos comerciales más restrictivos debido a los riesgos asociados. Esto podría incluir plazos de pago más cortos o requisitos de garantía adicionales.
- **Presentaran dificultades operativas:** Clientes en este segmento pueden enfrentar dificultades operativas, como capacidad de pago insuficiente o problemas de fluidez financiera, que los hacen menos aptos para participar en mercados internacionales.

La clasificación como "No Aptos para Exportar" no significa necesariamente que estos clientes deban descartarse por completo. Puede ser necesario un enfoque estratégico alternativo, como la mejora de la capacidad operativa o la mitigación de riesgos financieros antes de considerar la exportación.

3. Variables por tomar en cuenta para la clasificación de clientes

3.1.Liquidez Financiera

La liquidez es esencial para la estabilidad financiera y puede influir en la capacidad de un cliente para cumplir con los compromisos a corto plazo. Medir el flujo de efectivo de un cliente de manera efectiva requiere un enfoque integral que tenga en cuenta tanto los aspectos financieros como los operativos del negocio, debido a que se lo calcula restando los ingresos menos los gastos del cliente. La comunicación abierta y la colaboración con el cliente son esenciales para comprender completamente su situación financiera y ofrecer una correcta clasificación. Dentro del ámbito financiero, particularmente en el contexto empresarial, un indicador crucial que aborda la solidez financiera es el índice de solvencia. Este indicador

ofrece una perspectiva sobre la capacidad de una empresa para cumplir con sus obligaciones financieras a corto plazo, evaluando su robustez.

El cálculo se realiza al dividir el activo corriente (AC), que comprende los recursos que pueden convertirse en efectivo en menos de un año, entre el pasivo corriente (PC), que abarca las deudas con vencimiento en el mismo período. La fórmula para el índice de solvencia es la siguiente: Índice de Solvencia (IS) = AC/PC.

La interpretación de este índice se presenta de la siguiente manera:

- Si el índice de solvencia es superior a 1, esto indica que la empresa cuenta con suficiente liquidez para cumplir con sus deudas a corto plazo.
- Si el índice de solvencia es igual a 1, significa que la empresa posee activos líquidos equivalentes al valor de sus deudas a corto plazo, lo cual podría acarrear problemas financieros en el futuro.
- Si el índice de solvencia es inferior a 1, la compañía se encuentra en una situación crítica, ya que no dispone de los activos más líquidos necesarios para cubrir sus deudas a corto plazo. En estos casos, las empresas suelen recurrir a procesos de reestructuración para refinanciar sus deudas o generar más liquidez mediante la venta de activos.

3.2.Score Crediticio

En el actual entorno de incertidumbre económica y desaceleración global, el puntaje crediticio se vuelve excepcionalmente crucial. Este puntaje es una evaluación numérica de la solvencia crediticia de un individuo o empresa, y en momentos como los actuales, cobra una importancia sin precedentes. La capacidad para acceder a préstamos, tarjetas de crédito o incluso para obtener una hipoteca es una preocupación creciente que debe ser tomada en cuenta a la hora de analizar factores claves de un cliente.

En este contexto, contar con un buen puntaje crediticio marca una gran diferencia cuando se trata de disfrutar de estabilidad financiera. Un puntaje crediticio positivo abre las puertas a opciones crediticias favorables y condiciones más ventajosas. Permite a las personas o compañías acceder a préstamos con tasas de interés más bajas, obtener límites de crédito más altos y, en general, facilita la gestión de asuntos financieros.

En un periodo de incertidumbre económica, donde la obtención de crédito puede ser más difícil, mantener un puntaje crediticio sólido se convierte en un activo valioso. Ayuda a mitigar riesgos financieros y proporciona una sensación de seguridad en medio de la incertidumbre económica actual. En resumen, tener un buen puntaje crediticio se vuelve esencial para asegurar la estabilidad financiera y la capacidad de acceder a oportunidades crediticias en un momento en que estas preocupaciones son cada vez más frecuentes. Esta variable desglosa y califica su puntaje crediticio en función de los siguientes criterios: historial de pagos, uso del crédito, tipos de Crédito, duración del crédito y consultas.

3.3.¿Por qué tomar en cuenta el score crediticio?

El score crediticio toma en cuenta diferentes variables, por lo cual, en una sola variable podemos abarcar varios aspectos y plasmarlos como un valor que refleje la situación de esa persona o empresa.

3.4.Historial de Pago (35% del Puntaje)

Un historial de pago sólido es un indicador clave de la fiabilidad financiera de un cliente. Si un cliente tiene antecedentes de realizar pagos a tiempo, esto sugiere una gestión financiera responsable y confiable. En el contexto de la exportación, donde los plazos y compromisos financieros son críticos, contar con clientes que han demostrado consistencia en el pago es esencial para minimizar el riesgo crediticio.

3.5.Porcentaje de Uso del Crédito (30% del Puntaje)

Evaluar cuánto crédito utiliza un cliente en relación con su límite total proporciona información sobre su capacidad para manejar responsabilidades financieras. Un cliente que utiliza una gran parte de su crédito disponible podría indicar una mayor presión financiera. Este factor es crucial al considerar la capacidad del cliente para asumir compromisos de exportación y cumplir con los pagos en los plazos acordados.

3.6.Duración del Historial Crediticio (15% del Puntaje)

Un historial crediticio prolongado es un indicativo positivo para los acreedores. En el contexto de exportación, trabajar con clientes que tengan un historial crediticio estable y de larga duración proporciona una mayor seguridad en términos de la capacidad del cliente para mantener prácticas financieras sólidas a lo largo del tiempo.

3.7. Tipos de Crédito (10% del Puntaje)

La diversidad en la gestión de diferentes tipos de deuda demuestra la capacidad del cliente para manejar diversas obligaciones financieras. La habilidad para gestionar eficientemente tarjetas de crédito, préstamos a plazos, cuentas minoristas y otros tipos de crédito sugiere una versatilidad financiera. Esta versatilidad es valiosa al considerar la capacidad del cliente para cumplir con compromisos de exportación de diversas naturalezas.

3.8. Consultas de Crédito (10% del Puntaje)

El número y la frecuencia de consultas de crédito pueden ser indicativos del comportamiento crediticio del cliente. Un historial extenso de solicitudes puede ser una señal de búsqueda activa de crédito, lo cual puede ser considerado un riesgo crediticio. En el contexto de la exportación, donde se establecen relaciones a largo plazo, es esencial evaluar si el cliente tiene un historial de consultas que pueda indicar un riesgo potencial.

3.9. Tamaño del pedido en dólares

El factor conocido como el tamaño del pedido en dólares se caracteriza por ser un factor fundamental en la caracterización del cliente. Se le suele referir como el EOQ, y representa a una técnica de gestión que tiene relación con la cadena de suministros para determinar el tamaño del pedido y que se lo pueda clasificar por lotes. La estrategia en la intervención de esta técnica tiene relación con el evitar el desabastecimiento y el exceso de las existencias de los pedidos, logrando equilibrar los costos de inventario con los costos de oportunidad de la empresa. Esta variable funciona como un punto en el cual la empresa podrá aplicar una gestión más eficiente respecto al control de los inventarios, a fin que estos puedan obtener un desempeño óptimo sobre los bienes que se deben de comprar para mantener a flote la producción. Además, el control de los inventarios está relacionado con los costos en el tiempo y las garantías necesarias respecto al producto que ha sido mantenido respecto los niveles de inventarios equilibrados (Kuuse, 2023).

El uso de este modelo fundamenta información básica sobre el desarrollo del sector empresarial, permitiendo determinar los procesos de producción para regular los factores del almacenamiento, el mantenimiento de la demanda anual del producto, el mantenimiento de los costos del pedido y el almacenamiento constante del sistema de información y desarrollo del proceso de almacenamiento de la materia prima. Es decir, se logra que el producto esté siendo

elaborado de forma constante, considerando elementos como la demanda anual y los costos de fabricación, para que la empresa pueda establecer sus necesidades de producción según el tipo del cliente y la temporada de venta. Para desarrollarlos de forma más eficiente, se encuentran los supuestos básicos en los que se establece a la demanda constante, el tiempo de entrega, el costo del pedido y el costo del mantenimiento.

La herramienta ofrece información a la empresa respecto al tamaño ideal del pedido, las unidades de producción o materia prima que se deben de considerar. Así, como la cantidad de dinero que la empresa debe de esperar por parte de los clientes, siendo este la cantidad de dinero que el cliente está dispuesto a pagar por el acceso al pedido. La herramienta también ha sido desarrollada con la finalidad de establecer otros factores como la reducción de los costos de transporte, la reducción de los costos de los pedidos, la adaptación de las necesidades presentes y reales del mercado, la mejora del flujo de caja, las mejoras en el índice de rotación de existencias y la mayor rentabilidad en la empresa (Kuuse, 2023).

Las variables a determinar será la cantidad de dinero que el cliente estará dispuesto a pagar por el acceso al producto agrícola. Basada este factor, la empresa podrá definir la cantidad de mano de obra que podrá necesitar para el proceso de cultivo y cosecha del producto agrícola. Además, de la materia prima necesaria para el proceso de producción que van desde las semillas, el agua, hasta los pesticidas y fertilizantes necesarios para desarrollar el proceso. Todos estos valores deberán estar inventariados para evitar el desabastecimiento y el exceso de la existencia.

3.10. Ubicación Geográfica

Una de las variables más importantes respecto a los clientes es la ubicación geográfica. Este factor representa la capacidad dentro de la empresa de encontrar un punto específico sobre la superficie terrestre para desarrollar un plan de negocios que considere todos los costos incurridos para que el producto llegue hasta el cliente. Siempre será fundamental priorizar la buena ubicación de los clientes, puesto que cuando estos tengan un fácil acceso al producto, será más rentable tanto para el cliente como para la empresa la realización de los negocios. Cabe destacar que la ubicación geográfica está relacionada entre el segmento del mercado al cual se dirige la empresa. En este caso al ser una empresa orientada a la producción agrícola, deberá estar cercano a los lugares de cosecha y cultivo, para tener un mejor acceso a los

productos y aplicar un buen seguimiento a fin de determinar un buen proceso de calidad en el cultivo.

Por otro lado, la empresa debería contar con otros factores como son las sucursales, a fin que se pueda tener un mejor acceso a los clientes. Mientras más acercamiento exista entre los clientes y la empresa, será más eficiente que la empresa logre el aumento de su potencial en la elección de la ubicación del negocio, de forma que se pueda llegar con facilidad y el acceso a los mismos. En cuanto el espacio adaptado a las necesidades particulares es un factor que determina la localización de las necesidades del negocio. Este elemento está enfocado en que el espacio elegido por la empresa pueda cumplir con las necesidades del negocio, para que la rentabilidad o la compra del espacio sea prometedora. En cuanto a la localización de la competencia tiene relación con los problemas de los competidores, puesto que serán la razón que la empresa pueda reducir su capacidad de ganancias. Una buena ubicación garantizará que los competidores no afecten de forma considerable al negocio (Centro de Negocios Zaragoza, 2023).

Otro de los factores que se deben de considerar respecto a la importancia de la ubicación de la empresa, son los aspectos de proximidad y la facilidad para los proveedores. Una empresa que esté cercano a los proveedores será capaz de reducir los costos de movilización, reducción del costo de la materia prima y mejorar las condiciones necesarias, siendo una alta incidencia en el entorno agrícola. Es por ello que en este sector las empresas se encuentran ubicados cercano a las grandes ciudades o a sus plantaciones. En cuanto a la ubicación, se encuentran otros factores importantes relacionados con la comunicación, fomentando que el acceso a las grandes infraestructuras pueda permitir un proceso de comunicación efectiva para el funcionamiento adecuado de los negocios, valorando la facilidad de acceso para los clientes (Melior Centros de Negocios, 2014).

Las variables que se consideran para la ubicación de la empresa serán:

- Proximidad al sector de las plantaciones agrícolas
- Presencia de la competencia
- La cercanía con los proveedores
- Características del espacio elegido
- Mejoras en el proceso de comunicación

3.11. Costo de la logística en dólares

Los costos de la logística se caracterizan por ser todos los gastos que se incurren para el desarrollo del transporte de un producto desde el lugar de origen donde fue producido hasta el consumidor final, incluyendo todo su trayecto por intermediarios. Todos los costos logísticos se consideran un factor que puede aumentar el costo del producto final en el mercado, aumentando sus costos de forma rutinaria. Existen varios tipos de costos logísticos que se deben de considerar como es el costo del inventario, el cual representa que todos los gastos estén asociados con los pedidos, el mantenimiento y la gestión del inventario, fundamentales para la operación de un negocio. Además, es en este tipo de costos que se incurren respecto los costos de los pedidos, problemas con el mantenimiento y la posible escasez dentro del sistema de producción.

La administración de la cantidad de unidades disponibles forma parte de la producción respecto a la demanda de un bien y su impacto en el resultado final. Se pueden presentar algunos problemas relacionados con los costos de la logística como es algún factor negativo en la gestión de la cartera de suministros y el resultado de no lograr rastrear de manera adecuada las grandes cantidades de datos en tiempo real. Los costos de la logística consideran algunos aspectos necesarios que podrían causar problemas en los costos relacionados con los pronósticos deficientes, los retrasos de los envíos, el incremento del costo de producción, el inventario que se encuentra desperdiciado, la baja tasa de rotación, los costos generales, los errores de entradas para los datos y la desaparición de los clientes.

Una variable fundamental es el costo de distribución, el cual determina el aumento de los costos de la empresa relacionados con el transporte de la mercadería. Para lograr una óptima eficiencia, es necesario que el proveedor disponga de factores como la programación, los tiempos de producción y la contribución con los costos de la logística. Esta variable es la más importante puesto que garantiza que la optimización de la cadena de valor sea la adecuada para realizar un seguimiento prudente sobre los puntos de contacto de los proveedores según la variabilidad de sus costos (UCSP, 2023).

Otro de las variables que se debe considerar respecto al costo de almacenamiento, el cual es un factor responsable para medir el rendimiento de la empresa y la administración de los costos de almacenamiento. Dentro de la empresa, se debería de comenzar por medio del análisis de la frecuencia del desempeño, la capacidad de precisión de la selección de los clientes,

la tasa de contracción del inventario y los porcentajes con los cuales los clientes pueden ir aumentando de manera consecutiva. Esta normativa garantiza el rastreo eficiente de la evaluación de los almacenes y la planificación sobre la demanda que fomente el exceso del inventario que esté ligado sobre sus cadenas de suministros (UCSP, 2023).

En el dinámico mundo del comercio internacional, la eficiencia y la precisión en la gestión logística son elementos esenciales para el éxito empresarial, por lo tanto, la empresa se maneja en este ámbito con dos packs que buscan satisfacer las distintas necesidades de las empresas en sus operaciones globales.

3.12. Utilización logística eficiente del proceso

3.12.1. Pack 1: Solución Logística Eficiente

Costo Logístico: \$1,000 - \$15,000

Descripción: El Pack de Solución Logística Eficiente está diseñado para empresas que buscan una solución logística eficiente y rentable. Incluye servicios integrales, desde la gestión de inventario hasta la entrega final. El equipo especializado se encargará de coordinar todas las actividades logísticas para garantizar una cadena de suministro fluida y sin contratiempos.

Servicios Incluidos:

- **Gestión de Inventario:** Supervisión y control eficiente del inventario.
- **Transporte Internacional:** Coordinación de transporte marítimo o aéreo, según sea necesario.
- **Despacho Aduanero:** Manejo de trámites aduaneros para una importación/exportación sin problemas.
- **Seguimiento en Tiempo Real:** Herramientas tecnológicas para el seguimiento en tiempo real de la carga.

Incoterms Recomendados

- **FOB (Free On Board):** Para controlar los costos y garantizar una gestión logística eficiente hasta la entrega al transportista principal.

3.12.2. Pack 2: Solución Logística Premium

Costo Logístico: \$15,000 - \$25,000

Descripción: El Pack de Solución Logística Premium está diseñado para empresas que buscan una solución logística premium con servicios adicionales y mayor atención personalizada. Incluye una gestión logística integral con un enfoque proactivo para adaptarse a las necesidades específicas de la empresa.

Servicios Incluidos

- **Gestión de Inventario Avanzada:** Estrategias avanzadas de gestión de inventario para optimizar los niveles y reducir costos.
- **Transporte Internacional Prioritario:** Priorización en el transporte marítimo o aéreo para entregas más rápidas.
- **Despacho Aduanero Prioritario:** Proceso aduanero más rápido y eficiente.
- **Asesoría Logística Personalizada:** Consultoría personalizada para adaptar la solución a las necesidades específicas de la empresa.
- **Seguro de Carga:** Cobertura de seguro ampliada para mayor protección.

Incoterms Recomendados

- **CIF (Cost, Insurance, and Freight):** Incluye seguro de transporte, proporcionando una capa adicional de seguridad para la carga.

3.13. Frecuencia de compra

Uno de los factores más importantes a tomar en cuenta para el proyecto es la frecuencia de la compra. Este se encuentra relacionado con el comportamiento del cliente y la cantidad de veces que adquiere los productos en la empresa. para que exista una alta rentabilidad del

negocio, es necesario que la empresa considere principalmente clasificar a los clientes según su promedio de frecuencia de compra. De esta manera, se garantiza que los clientes puedan obtener alguna ventaja representativa respecto a su proceso de compra. Para lograrlo, la empresa se basará en el uso de la Big Data para determinar el proceso de compra y su frecuencia en un tiempo determinado de un año (Ortega, 2023).

La importancia de la frecuencia de compra para la empresa representa una oportunidad única para que se pueda aplicar una estrategia que permita la retención de los clientes indispensables y ofrecer beneficios a los compradores habituales. Estos dos factores tienen relación con la importancia de la retención de los clientes, siendo una cualidad indispensable para todos los negocios, puesto que resulta un factor de mayor rentabilidad que garantiza a la empresa mayores ganancias. Es más económico y rentable para una empresa mantener a sus clientes habituales que ganar un nuevo cliente. Por ende, la herramienta digital estará orientada a desarrollar una estrategia donde la frecuencia de compra será medida respecto a la tasa de repeticiones con los clientes, siendo este uno de los indicadores más utilizados para el seguimiento del caso. Además, los compradores habituales siempre serán más económicos de adquirir que los nuevos clientes debido que no es necesario invertir cantidades exorbitantes en publicidad.

Para aumentar la frecuencia de compra en la empresa, es necesario que se otorguen ciertas ventajas respecto al comportamiento del cliente y su relación con la empresa. Los clientes entienden las variables fundamentales como los productos de alta calidad, el precio del mercado y de recibir más por un menor valor. La empresa mediante la herramienta tecnológica debería ser capaz de iniciar con su programa de fidelización con el cliente, buscando otorgar recompensas a los clientes más fieles para aumentar su retención y la frecuencia de compra. El poco marketing que la empresa utilice debería ser adecuado para que se fomente la fidelización y la retención de los clientes. Las campañas se desarrollarán de manera específica para cada cliente, buscando liderar con sus necesidades de consumo en el mercado y utilizarla para la promoción de los productos en el mercado nacional e internacional (Ortega, 2023).

IX. CAPITULO III. IMPLEMENTACIÓN DEL MACHINE LEARNING

1. Implementar el modelo de Machine Learning utilizando las variables con más significancia.

La regresión logística se basa en el concepto de convertir una variable lineal, combinación de características (predictores) del cliente, en una probabilidad de pertenecer a una de las dos categorías: "Apto para Exportar" (1) o "No Apto para Exportar" (0). La variable que queremos predecir se llama "variable dependiente" o "variable de respuesta", mientras que las características que se utilizan para la predicción son las "variables independientes" o "predictores". En el proceso de modelado de regresión logística, se seleccionan los coeficientes del modelo de manera que maximicen la "probabilidad" de los resultados. Debido a la no linealidad del modelo, se requiere un algoritmo iterativo para estimar los parámetros con precisión. Además, dado que la regresión logística utiliza valores discretos y no sigue una distribución normal, no es adecuado utilizar el modelo de regresión lineal simple cuando se trata de una variable de respuesta dicotómica (con solo dos valores posibles).

El modelo de regresión logística primero pasa por un proceso llamado Data Cleaning o limpieza de datos, en la herramienta RStudio se refiere al conjunto de actividades que se realizan para identificar y corregir problemas en los conjuntos de datos con el objetivo de mejorar su calidad y utilidad. Estos problemas pueden incluir datos faltantes, valores atípicos, errores de entrada, valores duplicados y otros tipos de inconsistencias.

La limpieza puede ayudar a mejorar los siguientes aspectos del modelo:

1.1. Identificación de datos faltantes

- Utilizar funciones como **is.na ()** o **complete. Cases ()** para identificar las observaciones o variables con valores faltantes.
- Decidir cómo manejar los valores faltantes, ya sea eliminando las observaciones, imputando valores o utilizando otras estrategias.

1.2. Tratamiento de valores atípicos

- Identificar valores atípicos utilizando visualizaciones, como gráficos de caja (**boxplot**) o estadísticas descriptivas.

- Decidir si los valores atípicos deben eliminarse o ajustarse.

1.3. Manejo de duplicados

- Utilizar funciones como **duplicated()** para identificar observaciones duplicadas.
- Decidir si los duplicados deben eliminarse o si representan datos válidos.

1.4. Corrección de errores de entrada

- Revisar y corregir errores tipográficos o errores de codificación utilizando funciones de manipulación de cadenas de texto o expresiones regulares.

1.5. Transformación de datos

- Convertir variables a los formatos correctos.
- Crear nuevas variables o transformar existentes para satisfacer los requisitos del análisis.

1.6. Estandarización y normalización

- Asegurarse de que las variables estén en la misma escala o formato cuando sea necesario.

Algunas funciones y paquetes útiles en RStudio para realizar estas tareas incluyen **dplyr**, **tidyr**, **ggplot2** para manipulación y visualización de datos, así como otras funciones base de R. En este caso se utilizó la transformación de datos, debido a que las variables como, el costo de la logística, la liquidez financiera y la variable dicotómica de “Aprobado” estaban en formato binario por lo tanto mediante el comando “Factor()” se convirtió una variable a un factor, que es una variable que toma un conjunto finito de valores distintos, generalmente representando categorías o niveles, R asigna un número entero a cada nivel de la variable, lo que facilita su manipulación en análisis estadísticos y modelado.

Los factores son útiles en estadística y análisis de datos porque permiten realizar análisis de varianza (ANOVA), regresiones y otras operaciones estadísticas de manera más eficiente para variables categóricas. También son importantes en la creación de gráficos y visualizaciones, donde se pueden utilizar para especificar el orden y la apariencia de las categorías.

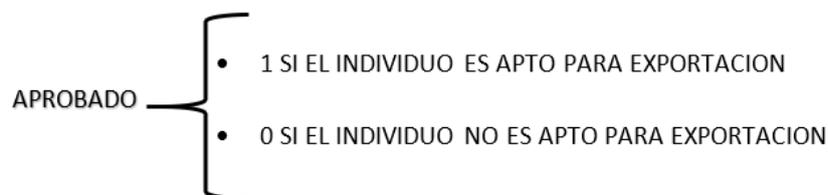
En el caso del costo logístico los niveles fueron dos, el primero representado por el 0 que fue convertido a la etiqueta “Pack 1” y la segunda representada por el 1, se convirtió en la etiqueta “Pack 2”. La variable de liquidez financiera al momento de ser convertida en factor se le asignó el nivel 0 a la etiqueta “No” y el nivel 1 a la etiqueta “Si” y por último la variable dicotómica llamada aprobada se le asignó al nivel 0 la etiqueta “No” y al nivel 1 la etiqueta “Si” al igual que con la variable de liquidez financiera.

La variable dicotómica es necesaria e importante ya que la regresión logística se utiliza comúnmente para predecir la probabilidad de un evento binario, donde la variable de respuesta toma uno de dos valores posibles (por ejemplo, 0 o 1, Sí o No, Éxito o Fracaso). Las variables dicotómicas se utilizan para representar variables categóricas con dos categorías. En el script, se tiene una variable convertida en factor llamada "Aprobado" con dos categorías posibles ("Si" y "No").

Por lo tanto, la variable “Aprobado” representa

Figura 5.

Clasificación del valor según la fórmula de aprobado



Nota. Clasificación del valor según la fórmula donde el aprobado significa que la variable es dicotómica. Recuperado de Propio Autor

Esto permite incorporar la información de la variable categórica en el modelo de regresión logística. Si tienes más de dos categorías, generalmente se crean $k-1$ variables

dicotómicas, donde k es el número total de categorías. Esto se conoce como "codificación de dummy" o "codificación uno contra todos".

Luego de este paso, se procedió a utilizar la función `set.seed(123)`, la cual se utiliza en RStudio para establecer una semilla inicial en el generador de números pseudoaleatorios de R. Esto significa que si utilizas funciones que generan números aleatorios en tu script o modelo, al establecer una semilla con `set.seed()`, los resultados serán reproducibles. Es decir, si ejecutas el mismo script varias veces con la misma semilla, obtendrás los mismos resultados cada vez.

Esto es beneficioso en varias situaciones. Primero, proporciona consistencia en los resultados, lo que es esencial para la reproducibilidad de la investigación o el análisis. Cuando se comparte el código con otros o se vuelve a ejecutar el propio código en el futuro, tener resultados consistentes facilita la validación y la comprensión de los resultados obtenidos.

En segundo lugar, establecer una semilla es crucial cuando trabajas con modelos estadísticos o algoritmos que implican elementos aleatorios, como el muestreo aleatorio o la inicialización de pesos en algoritmos de aprendizaje automático. La consistencia en la semilla garantiza que los mismos conjuntos de datos de entrenamiento y las mismas configuraciones iniciales se utilicen en cada ejecución del modelo. Esto es esencial para evaluar de manera justa y comparable el rendimiento del modelo en diferentes ejecuciones.

El tercer paso es el entrenamiento de los datos, esto mediante el comando `createdatapartition()` el cual es una función del paquete `caret` en R que se utiliza para crear particiones estratificadas de un conjunto de datos. Esta función es bastante útil para el modelo ya que, se está trabajando en problemas de aprendizaje supervisado, esto genera que se necesite dividir el conjunto de datos en subconjuntos de entrenamiento y prueba.

El objetivo este entrenamiento es encontrar los valores de los coeficientes que minimizan la función de verosimilitud, lo que significa que el modelo es capaz de predecir de manera efectiva la probabilidad de pertenecer a la categoría positiva dada la información en las variables predictoras.

Es importante señalar que el modelo se entrenó en el conjunto de entrenamiento con una proporción de 80/20 lo cual significa que, el 80% de los datos serán los que se entrenen, mientras que, el 20%. El proceso de entrenamiento busca encontrar un equilibrio entre ajustar el modelo a los datos de entrenamiento y evitar el sobreajuste (overfitting).

Esta partición se utiliza comúnmente en problemas de aprendizaje supervisado, donde se está entrenando un modelo para hacer predicciones y luego se evalúa en un conjunto de datos no visto para medir su rendimiento en datos no vistos. La elección específica de la proporción puede variar según la naturaleza del problema y la cantidad de datos disponibles

Por esto, podemos observar que luego del comando “`createdatapartition`”, dentro de los paréntesis de la función, se pone `p=0.8`, que es lo que indica que se quiere tomar una partición de 80/20, seguido del comando “`list= FALSE`” que controla el formato de salida de la función. Cuando “`list`” se establece en `FALSE`, la función devuelve un vector de índices. Este vector se puede utilizar directamente para seleccionar las observaciones del conjunto de datos original que formarán el conjunto de entrenamiento, mientras que el conjunto de prueba sería simplemente el complemento de ese conjunto. En este caso, establecer “`list = FALSE`” es conveniente porque proporciona directamente los índices en un formato simple de vector.

El cuarto paso se trató de crear el modelo mediante la función “`glm()`”, para así poder obtener los primeros resultados de nuestro modelo de regresión logística, al cual le llamaremos “`Logit`”.

La función “`glm()`” en este script se utilizó para ajustar el modelo de regresión generalizada. “`GLM`” es la abreviatura de “`Generalized Linear Model`”. A diferencia de los modelos lineales tradicionales, este modelo lineal generalizado permitió modelar respuestas que no siguen una distribución normal y no necesariamente tienen una relación lineal con las variables predictoras.

La sintaxis básica de la función `glm()` es la siguiente:

- **Formula:** Especifica la relación entre la variable de respuesta y las variables predictoras. Por ejemplo, `y ~ x1 + x2` indica que estás modelando la variable de respuesta `y` en función de las variables predictoras `x1` y `x2`. En el caso del modelo la variable “`Approved~`” señaló que esa era la variable de respuesta y debido al “`~`” se indicó que las demás serían las variables predictoras.
- **Data:** Especifica el conjunto de datos que contiene las variables utilizadas en la fórmula. Se tomo como data la base de datos que al inicio llamamos como `DataBaseValExp`, porque de esta se sacó las variables que se usaron antes para la formula.

- **Familia:** Esta especifica la familia de distribuciones que modela la variable de respuesta. Para regresión logística binaria, se utiliza **family = binomial**. Para regresión Poisson, se utiliza **family = poisson**, y así sucesivamente. Se puso binomial ya que nuestro tipo de regresión a utilizar sería la logística.

Con el Logit ya creado ahora mediante la función `summary()` se pudo obtener información sobre el ajuste del modelo y evaluar la significancia estadística de los coeficientes.

Figura 6.

Tabla de coeficientes y sus significados en el Logit 1

```

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)   -3.0651437  0.3085654  -9.934 < 2e-16 ***
FinancialLiquiditySi  3.6884741  0.3205519  11.507 < 2e-16 ***
CreditScore     0.2083049  0.0510838   4.078 4.55e-05 ***
LogisticCostPack 2  -0.1767909  0.2721761  -0.650  0.5160
OrderAmount     0.0005292  0.0001743   3.036  0.0024 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

Nota. Desarrollo de la intervención de los coeficientes utilizando la herramienta digital. Recuperado de Propio Autor (2024).

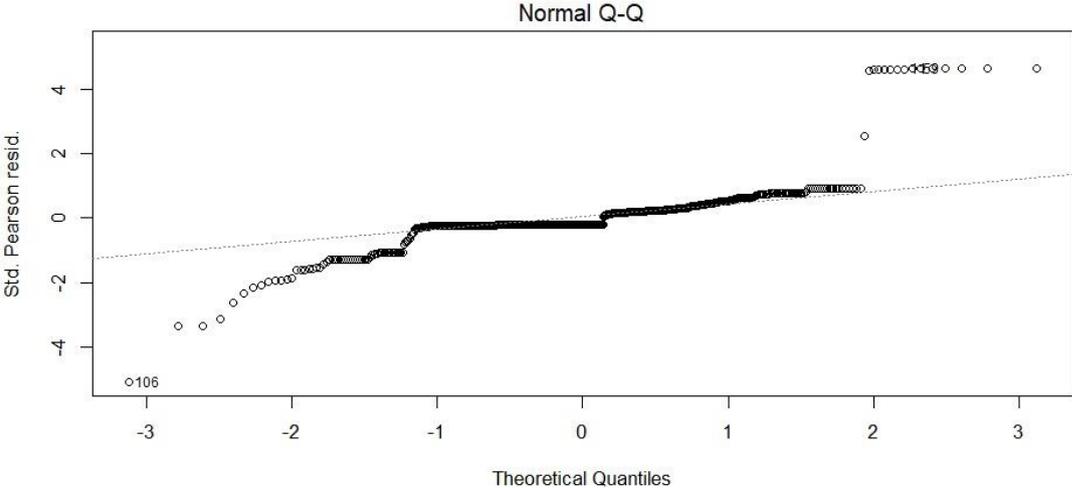
En RStudio, se ha establecido un sistema de clasificación que asigna niveles de estrellas a ciertos criterios. Estos niveles de estrellas corresponden a niveles específicos de confianza estadística. En concreto, se asigna el valor de 3 estrellas a un nivel de confianza del 99%, 2 estrellas a un nivel de 95% y 1 estrella a un nivel de 10%. Por lo tanto, con esta tabla se observó que las variables con más significancia fueron, el tamaño de pedido en dólares, el score crediticio y la liquidez financiera.

Así se vio cuáles son las variables que poseen mayor importancia a la hora de elegir a que cliente exportar o no, ya que, influyen significativamente en la determinación de la elegibilidad y la viabilidad de elegir a un cliente potencialmente exitoso. Luego mediante la

función “plot(Logit)” arrojó como resultado el gráfico Q-Q normal con el conjunto de datos. La disposición de los puntos en el gráfico permitió evaluar visualmente si los cuantiles de los datos se ajustaban a una distribución normal.

Figura 7.

Utilización del gráfico Q-Q del modelo sin optimizar

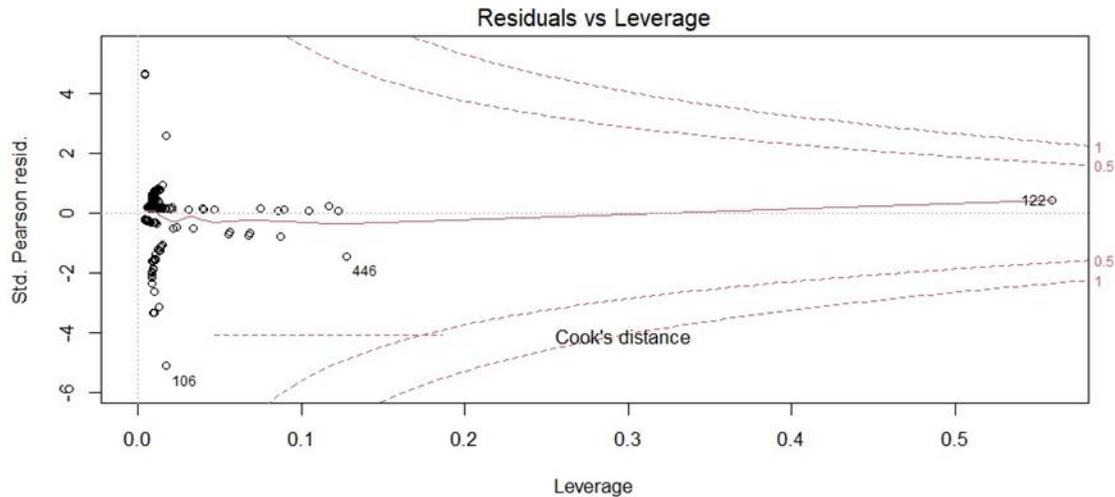


Nota. Ejemplo gráfico de la utilización del gráfico Q-Q existente en el modelo sin optimizar. Recuperado de Propio Autor (2024).

En este caso, la observación de que todos los valores estaban dispuestos de manera correcta en el gráfico Q-Q normal sugiere que los datos seguían aproximadamente una distribución normal. La línea diagonal en el gráfico sirvió como referencia, y la proximidad de los puntos a esta línea indicó la validez de la asunción de normalidad para los datos. La interpretación detallada de los resultados del gráfico Q-Q incluyó la evaluación de cualquier desviación notable de la linealidad en la disposición de los puntos. También, arrojó el gráfico de la distancia de Cooks, es valioso para entender cómo cada observación afecta el ajuste del modelo y para identificar posibles puntos atípicos o influyentes.

Figura 8.

Gráfico de la distancia de Cooks existente en el modelo



*Nota. Ejemplo gráfico de la utilización de la distancia de Cooks en el modelo.
Recuperado de Propio Autor (2024).*

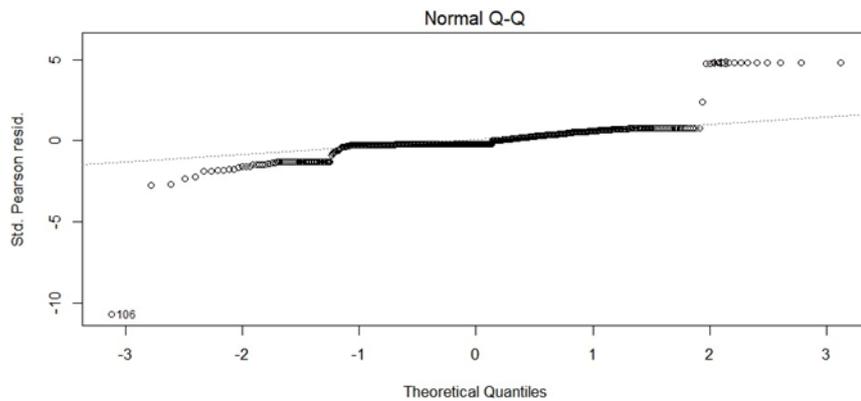
El análisis del gráfico de la Distancia de Cook proporcionó una evaluación en profundidad de la influencia de cada observación en el modelo. El gráfico sugiere que el modelo no fue sensiblemente afectado por observaciones atípicas o influyentes. La atención a la presencia de puntos atípicos o influyentes ayudó a garantizar la validez y la precisión del modelo de regresión, contribuyendo así a la toma de decisiones fundamentada en el análisis de datos.

Una vez hecho el primer Logit se procedió a crear uno más ya que de esta manera generaría una optimización del modelo y así que este enfoque simplificara el modelo al eliminar variables que podrían tener una contribución limitada o no significativa al modelo. Al eliminar variables menos relevantes, la interpretación de los efectos de las variables restantes en la variable de respuesta se volvió más clara y precisa.

Este segundo logit llamado “Logit2” dio dos gráficos más:

Figura 9.

Gráfico Q-Q del Modelo Optimizado

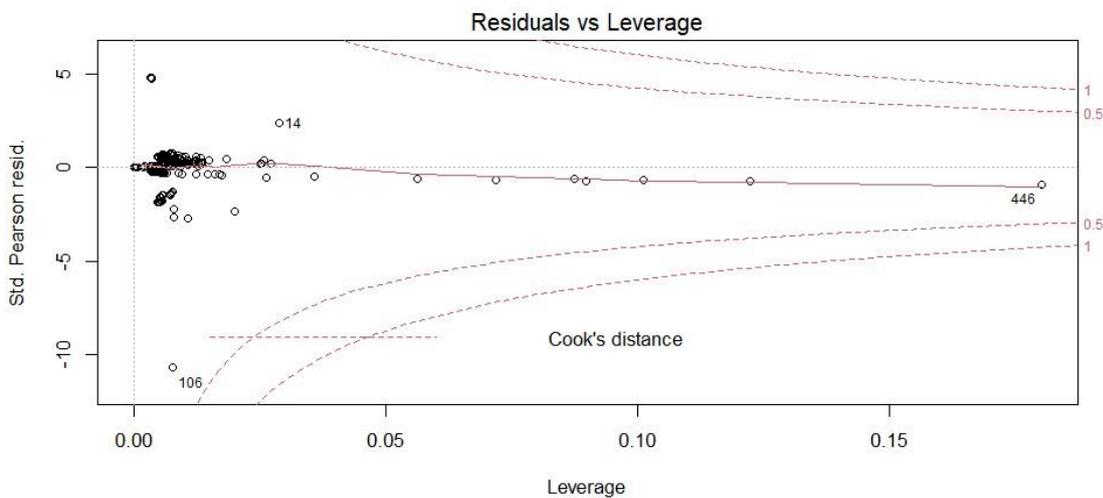


Nota. Ejemplo gráfico de la utilización del modelo Q-Q existente en el modelo Optimizado. Recuperado de Propio Autor (2024).

Cuando se observa que el gráfico de cuantiles-cuantiles (Q-Q plot) en el segundo Logit está más ajustado y los datos están más pegados a la línea diagonal, esto sugiere una mejora en la normalidad de los residuos en comparación con el primer Logit. Al crear el segundo Logit con solo las variables más significativas, se eliminó la influencia de variables que contribuían a la falta de normalidad en los residuos del primer modelo, esto resultó en residuos más cercanos a una distribución normal.

Figura 10.

Gráfico de la Distancia de Cook existente en el Modelo Optimizado



Nota. Gráfico de la Distancia de Cook en el Modelo Optimizado. Recuperado de Propio Autor (2024).

El gráfico de la Distancia de Cook en el segundo logit está más ajustado y los datos están más pegados entre sí y a la línea central, esto sugiere una disminución en la influencia de observaciones atípicas o influyentes en comparación con el primer logit. Al crear este modelo para optimizarlo se eliminó la variable con menos significancia lo cual hizo que haya reducido la presencia de variables menos relevantes que podrían haber contribuido a la identificación de observaciones como influyentes, reduciendo así la influencia de observaciones atípicas o influyentes en el ajuste general del modelo.

Lo siguiente fue calcular la probabilidad de éxito de la clasificación con la función “predict ()”, esta función se usa para obtener las probabilidades predichas de pertenencia a una clase específica según las siguientes variables:

- **Modelo logístico:** Es el modelo de regresión logística ajustado previamente utilizando el Logit2.
- **Newdata:** Representa los datos para los cuales deseas hacer predicciones. En este caso fue nuestra base de datos (DataBaseValExp), en la cual crearemos una nueva variable llamada “Prob”.
- **Type = "response":** Este argumento especifica el tipo de predicción que deseas. En el contexto de regresión logística, "response" se utiliza para obtener las probabilidades predichas. Las probabilidades están en la escala de 0 a 1 y representan la probabilidad de pertenecer a la categoría positiva.

Después de ejecutar este código, “Prob” contendrá las probabilidades predichas para cada observación en nuevos datos. Se usarán estas probabilidades para tomar decisiones de clasificación según un umbral específico. Es por esto que, una vez creada la variable “Prob” se la utilizó para clasificar a los clientes, mediante la función “ifelse”, haciendo que tomara de la base de datos la variable anteriormente creada (“Prob”)

- **DataBaseValExp["Prob"]:** Se refiere a la columna llamada "Prob" en el DataFrame (llamado **DataBaseValExp**). Esto representa las probabilidades predichas por el modelo de regresión logística.

- **DataBaseValExp["Prob"] >= 0.5**: Esta es la condición evaluada. Se está verificando si los valores en la columna "Prob" son mayores o iguales a 0.5.
- **ifelse(DataBaseValExp["Prob"] >= 0.5, 1, 0)**: Para cada fila en la columna "Prob", si el valor es mayor o igual a 0.5, se asigna el valor 1; de lo contrario, se asigna el valor 0. El resultado de esta operación se asigna a la columna "Class", reemplazando los valores originales.

Por último, se utilizó la función `table()` para construir una matriz de confusión en la cual se explica lo siguiente:

- **DataBaseValExp[, "Approved"]**: Representa la variable real ("Approved") en tu conjunto de datos. Estos son los valores reales o verdaderos de las observaciones.
- **DataBaseValExp[, "Class"]**: Representa la variable predicha ("Class") en el conjunto de datos. Estos son los valores predichos por el modelo.
- **dnn = c("Actual", "Predicho")**: Especifica los nombres que se asignarán a las dimensiones de la tabla de contingencia. En este caso, "Actual" se refiere a los valores reales y "Predicho" se refiere a los valores predichos.

Figura 11.

Matriz de Confusión

	Predicho	
Actual	0	1
No	305	78
Si	22	285

Nota. Utilización de las variables para la matriz de confusión. Recuperado de Propio Autor (2024).

Entonces al lanzar la matriz de confusión esta indicó que hay un total de 383 clientes a los que no se les exporto, de los cuales, 305 clientes estaban bien clasificados y 78 clientes si pudieran haber sido factibles, lo que quiere decir que en términos de porcentajes se perdió un 20% de clientes que sí podrían haber sido aptos para una exportación sin riesgo financiero. También, se indicó que hay un total de 307 clientes a los cuales, si se les atendió, pero 285 estaban bien clasificados y 22 estaban dentro la clasificación no cumplir con los criterios de elegibilidad por lo tanto a un 93% de los clientes en el “Si” se los habría clasificado de manera correcta y un 7% no debió haber sido clasificado con como apto para exportar.

X. CAPÍTULO IV. ANÁLISIS DE LA ESTRATEGIA PARA LOS CLIENTES POR MEDIO DEL SCRIT

1. Estrategias para los clientes por medio del Script

1.1. Estrategia de fidelización a los clientes

La fidelización de la clientela consiste en una estrategia de mercaderero en la que se busca la retención de los clientes que ya han sido atendidos en la empresa, para que se establezcan como clientes constantes, siendo una ventaja para la empresa debido que se ahorran en costos de obtención de nuevos clientes. La forma en la que se mantienen a los clientes es por medio de las experiencias positivas que los clientes han logrado obtener por medio del servicio de la empresa. Por lo que, para este caso, utilizando la herramienta script, se generará una serie de interacciones positivas entre el consumidor y la empresa para mantener siempre una confianza al largo plazo con el cliente. Los costos para mantenimiento de los clientes constantes pueden llegar a costar hasta cinco veces menos, permitiendo que la empresa exportadora pueda obtener una mejor estrategia orientada a la retención de los clientes gracias a un buen servicio (Zendesk, 2023).

Entre las ventajas que se encuentran en la estrategia de fidelización de los clientes, es la venta del servicio de exportación para los mismos clientes, lo que representaría una importante ventaja orientada a la experiencia positiva del servicio. Las empresas que aplican una estrategia de fidelización del cliente obtienen beneficios como la lealtad, el incremento de la frecuencia de compra del servicio, aumento de la rentabilidad, la consecución de recomendaciones en el mercado y la obtención de la respectiva retroalimentación que mejorará el estilo y el servicio ofrecido por la empresa. Para una empresa es más rentable ofrecer una experiencia positiva con un cliente ya habitual en lugar de generar riesgos con clientes desconocidos, generando un mejor costo de adquisición del cliente, siendo la base sobre el consumidor que compra por primera vez el servicio de exportación (Aguero, 2014). Para la empresa, la herramienta permitiría que este sea el cliente con mayor importancia e interés, permitiéndole que una vez que haya atendido a todos los clientes y se termine el stock, los clientes más fieles que han adquirido la mayor cantidad de veces el servicio, recibirían beneficios como:

- Un trato preferencial y personalizado
- Deberán de realizar un pago anual de \$60.000 para mantener su privilegio

- Tiempo de espera corto (3 a 5 días)
- Garantías de protección del producto hasta la llegada del mismo al cliente
- Un descuento del 5% al 10% en los nuevos servicios de exportación
- Información en tiempo real sobre el estado del producto exportado en tiempo real
- Ingreso en una plataforma para el sorteo trimestral de premios (descuento del 50% en envío de productos)

Los clientes que recomienden otras empresas interesadas en el servicio de exportación de los galpones de pitahaya, obtendrán un descuento adicional del 5% en el servicio. Lo que evidenciaría para la empresa una estrategia de gestión orientada al servicio personalizado y a las ventas por volumen, estableciendo a la empresa como una marca altamente recomendable por los beneficios y principalmente por los precios que posee en el mercado. Cabe destacar que los beneficios otorgados por la empresa de exportación se lo han realizado en función a los productos que formen parte del Incoterm Free on Board, por las siglas de FOB. Lo cual, es necesario tomar en consideración para que la empresa pueda desarrollar una estrategia comercial que considere aspectos como el valor comercial de la mercadería, el costo del empaquetado, los costos de etiquetado, los costos generados en la aduana por salida del producto, los trámites a tomar en consideración, los gastos del puerto, entre otros. La herramienta tomará siempre en consideración estos beneficios a los clientes que reportaron una mayor incidencia en el acceso al servicio (Ospina, 2020).

La estrategia de personalización del servicio estará orientada con la inversión por parte de la empresa en atender de forma personalizada al cliente con calidad. Para ello, es necesario considerar la manera en la que la empresa está atendiendo a la clientela, generándose posibles problemas como la falta de atención sobre esta estrategia. Una empresa productora de pitahaya lo primero que va a necesitar es obtener información relacionada sobre el servicio, las ventajas que se ofrecen y la diferencia del servicio de la empresa respecto a la competencia. Todo el personal que trate con la persona deberá estar totalmente capacitada, a fin que se pueda lograr un respectivo enfoque de acuerdo a las necesidades del consumidor. Lo más recomendable es ofrecer una atención personalizada, sea por chat de las aplicaciones digitales como What's App, o por medio del uso de redes sociales. Los trabajadores deberán de obtener una cita directa con los clientes para ofrecer de mejor forma el servicio y la orientación respectiva sobre las ventajas de seguir siendo clientes de la empresa exportadora (Murillo, 2018).

Cabe destacar que el proceso de comunicación implementado deberá ser por medio de una comunicación personalizada, considerando que no todos los clientes serán iguales. Lo más recomendable es invertir en la personalización, utilizando los datos que la base de datos ha ido almacenando a fin que se pueda conocer de antemano las necesidades del cliente respecto a la exportación de su producto. Cabe destacar que la empresa deberá lidiar con distintas formas de tratar al cliente, dejando en evidencia su estrategia que le permita destacar en la competencia. La empresa exportadora también se comprometerá en el asesoramiento sobre las mejores técnicas de producción y comercio de sus productos de forma gratuita, ofreciendo el beneficio cada tres meses.

Todos los clientes que el sistema coloque en este punto serán considerados como clientes VIP, obteniendo así su club o membresía VIP para todos los consumidores, alcanzando el acceso a las recompensas y los beneficios que se apliquen sobre los programas de fidelización que el cliente por lo general espera obtener con el desarrollo de su accionar comercial. El club VIP es una oportunidad para que la marca pueda afianzarse en el mercado. La estrategia del club VIP se basa en que los clientes recibirán un trato diferencial directo gracias a su importancia comercial que representan para la empresa. Con ello, la empresa aplica una estrategia que beneficia su relación con sus clientes objetivos y el propio mercado, siendo estos los clientes que la empresa no deseará perder, haciéndolos en todo momento sentirse especiales por tener una atención superior al resto de clientes. Lo que otorga a los clientes que desean exportar sus productos de pitahaya una ventaja competitiva en el mercado internacional (UNFPA Ecuador, 2020).

El grupo de clientes que conforman la clientela VIP, obtendrá ciertas características importantes en el mercado, como la compra habitual que realizan del servicio de exportación, invirtiendo importantes sumas de dinero que la empresa podrá aprovechar como un flujo importante de capital que la empresa aprovechará para mejorar su servicio, generando así un ciclo repetitivo con tendencia a ir aumentando en su efectividad para la empresa. Los clientes VIP por lo general son los que invierten una mayor cantidad de dinero, representando mayores ganancias para la empresa exportadora. Una de las mayores importancias que la empresa posee es el convertirse en una marca portavoz de la empresa, generando recomendaciones en el mercado, representando así a la publicidad gratuita que se necesita para aumentar su potencial en el mercado. Los clientes VIP por lo general tendrán en alta estima el valor de la marca, siendo clientes satisfechos con el servicio recibido. Por lo cual, la empresa exportadora deberá

de ofrecer un mejor servicio post venta para darle el seguimiento respectivo al producto exportado al mercado internacional (Zendesk, 2023).

1.2.Estrategia de atención al cliente con una clasificación no apta financiera para exportar

Cuando el sistema determina que el cliente posee una calificación financiera no apta para el contrato del servicio de exportación, lo que las empresas tradicionales por lo general realizan es dejar de lado al cliente, no ofreciéndole una solución que se adapte a su necesidad, perdiendo así al cliente que irá por un servicio que se adapte a su propia capacidad económica. Las grandes empresas exportadoras por lo general dejan escapar a los clientes con poca capacidad de producción y por ende de poco impacto en el mercado, solo por priorizar a los clientes de amplia escala. Por ende, en el caso de la empresa exportadora ha considerado el no dejar de lado a los clientes potenciales que se encuentra en el mercado con poca capacidad económica. En el mercado del Ecuador, el mercado comercial lo conforman más del 99% como empresas Pymes, lo que supondría un importante flujo económico para la empresa exportadora el no dejar de lado a los clientes y adaptarse a sus necesidades de consumo (UASB, 2021).

Para ello, la empresa exportadora ha desarrollado una propuesta conocida como “El cliente primero”, en el cual permitirá que los clientes con una calificación baja puedan recibir un servicio digno sin afectar la calidad del servicio de la empresa exportadora. La empresa realiza una estrategia de priorizar a sus clientes, pero en este caso se lo realiza por orden de llegada. A todos los clientes se les ofrecerá la capacidad de acceder al servicio en un grupo de clientes donde deberán llenar el stock del embarque de sus productos para que puedan ser exportados al mercado internacional. Este proceso de tiempo cabe destacar no será inmediato como el de los clientes VIP, generando esperas de hasta dos semanas para que los clientes puedan acceder al servicio de envío de su producto. Todos los costos del servicio serán menores que el VIP y la empresa exportadora obtendrá sus ganancias por volumen.

De igual forma que los clientes VIP y los clientes con financiamiento positivo, estos recibirán ciertos beneficios como son:

- Garantías de protección del producto hasta la llegada del mismo al puerto
- Tiempo de espera limitado (10 a 20 días)
- Descuento del 2% al 5% para los nuevos envíos de los productos

Incluso, considerando las estadísticas de las empresas Pymes que existen en el Ecuador, este debería ser el grueso o las mayores ganancias de la empresa. Por lo que, es fundamental que la empresa considere que los clientes que forman parte de este segmento de mercado siempre estén vigentes, ofreciéndole ciertos beneficios que les permitan adaptarse a sus necesidades de consumo. Además, se les deberá de ofrecer la posibilidad si aumentan su capacidad financiera de convertirse en clientes VIP, para que exista un proceso de cambio en los clientes, con una tasa de crecimiento de los clientes VIP de al menos el 5%. Lo que representaría en el mercado para la empresa, una mejora de su imagen de marca al no dejar abandonar en ningún momento a ningún cliente, ofreciéndole beneficios importantes que les permitan crecer en el mercado.

La estrategia también representa el evitar las posibles pérdidas de los clientes para la empresa. Todas las pérdidas de los clientes representan una reducción de su impacto e imagen en el mercado. Por lo cual, es recomendable la intervención de esta estrategia como una herramienta para evitar la pérdida de los clientes. La incompatibilidad con los clientes con la empresa puede considerarse como un problema grave que las empresas que no mejoran de forma inmediata representarían un impacto negativo en el mercado. La empresa ha sido capaz de escuchar todos los reclamos, necesidades y quejas de los clientes, adaptándose a todas las necesidades para cumplir con la mayor demanda posible (Gómez, 2023).

1.3.Estrategia de venta a los clientes con una clasificación buena

Para los clientes con una calificación positiva, estos serán los clientes que sean los de segunda importancia para la empresa, solo superados por los clientes VIP. Los clientes con calificación positiva incluirán algunas empresas de pitahaya que gracias a un mayor impacto en el mercado por medio de la exportación del producto. Estos clientes deberán ser priorizados ofreciéndoles no solo el servicio, siendo también el asesoramiento respectivo sobre los mercados internacionales para que estas empresas productoras puedan aumentar su competitividad y rentabilidad en el mercado. Los clientes positivos en el mercado serán los clientes que el sistema al momento de terminar el stock del servicio, les ofrecerán el servicio VIP, a fin de obtener los mejores beneficios por parte de la empresa.

Para que la empresa exportadora pueda contar con un 50% de sus clientes con financiamiento positivo, debería de disminuir sus criterios basados en su capacidad de producción y ganancias generadas en el mercado. Por lo cual, la estrategia se debería enfocar en que este público pueda gozar de ciertos beneficios como:

- Control de la exportación de su producto en tiempo real
- Garantías de protección del producto hasta la llegada al puerto y entrega al cliente
- Tiempo de espera limitado (5 a 12 días)
- Descuento del 5% al 10% para los nuevos envíos de los productos

Los clientes una vez que han quedado satisfechos con el servicio, deberán de ser recomendados con el servicio VIP justo después de su tercer servicio adquirido, de esta manera la estrategia se pondrá en marcha para su debida implementación, ya que, estos clientes disponen de una buena presencia financiera en el mercado, adicional a esto se debe recomendar las ventajas que representa para el cliente tener un trato preferencial y personalizado en el mercado de las exportaciones. Además, a este público se le ofrecerá asesoramiento mensual, constante y gratuito sobre las políticas de exportación que existen en el mercado extranjero, los principales mercados donde la pitahaya es de alto valor en el mercado. Así también sobre las variaciones de sus precios en el mercado extranjero. Para la empresa exportadora, ningún cliente deberá ser considerado como de poca importante, puesto que todos los consumidores son importantes y representan una oportunidad para que la empresa pueda mejorar su imagen de marca en el mercado, y sea reconocida como una de las pocas organizaciones de exportación que poseen beneficios reales que mejorarían al sector agrícola en el Ecuador, considerándose como una marca responsable con el sector Pymes del país.

CAPÍTULO V. CONCLUSIONES Y RECOMENDACIONES

1. Conclusiones

La implementación del Machine Learning para la clasificación de clientes en Valentina Export S.A. ha representado un cambio significativo en la optimización de los procesos de acercamiento a los clientes y en la toma de decisiones estratégicas. A través de la aplicación del modelo de regresión logística en RStudio, se logró clasificar a los clientes en segmentos "Apto para Exportación" y "No Apto para Exportación", lo que ha permitido establecer estrategias personalizadas para la atención a cada segmento.

Se identificó que, dentro del proceso de acercamiento al cliente, la empresa enfrenta un problema, destacando la falta de segmentación de clientes y la dificultad en la toma de decisiones sobre a qué clientes exportar, dada la alta demanda en comparación con la capacidad de producción del producto. Por lo tanto, la identificación de variables significativas para la clasificación de clientes ha brindado a la empresa una mayor comprensión de su base de clientes, variables significativas para la clasificación de clientes, como historial de compras, ubicación geográfica, tamaño del pedido y otros indicadores relevantes.

Esta implementación no solo mejoró la eficiencia operativa y la toma de decisiones, sino que también aumentó la competitividad de Valentina Export S.A. en el mercado nacional e internacional. Este proyecto representa una contribución significativa para la empresa, al proporcionar un enfoque más estratégico y rentable en la selección de clientes para exportación. Con esta implementación a la base de clientes, se creó una base de datos que permitió al modelo de regresión logística predecir la probabilidad de cada cliente de ser apto o no para la exportación.

La adopción de estrategias para cada segmento de cliente creado por el modelo refleja el compromiso continuo de Valentina Export S.A. con la excelencia y la satisfacción del cliente en la exportación de pitahaya. Esta iniciativa sentó las bases para un crecimiento sostenible y una mayor eficacia en la gestión de las relaciones con los clientes, posicionando a la empresa como un referente en el sector de exportación de productos agrícolas.

2. Recomendaciones

Considerar la implementación de un plan de inversión y ganancias en nuestras estrategias. Esto proporcionará un mayor control y una clara comprensión del proceso de implementación, así como de los posibles resultados a obtener. La formulación de un plan detallado no solo facilitará la toma de decisiones financieras, sino que también permitirá evaluar de manera más efectiva el rendimiento de nuestras estrategias.

Para garantizar un rendimiento eficiente en el manejo de la base de datos y lograr resultados óptimos en nuestro proyecto, es crucial implementar correctamente los métodos en el programa R Studio. La eficaz aplicación de estos métodos en el análisis y manipulación de datos no solo mejora la velocidad y efectividad de las operaciones, sino que también contribuye de manera significativa al éxito general del proyecto. Por lo tanto, se sugiere dedicar tiempo y atención a la optimización y adecuada utilización de las funciones dentro de R Studio, asegurando así un manejo eficiente de la base de datos y un rendimiento destacado en todas las fases del proyecto.

Se debe considerar la implementación de un plan sólido para el mantenimiento continuo del modelo de regresión logística utilizado en la segmentación de clientes. Este plan debería abordar la necesidad de actualizaciones periódicas y mejoras a medida que se acumulan nuevos datos y se profundiza en la comprensión del comportamiento del cliente.

XI. BIBLIOGRAFÍA

- Amazon. (23 de enero de 2024). ¿Qué es la regresión logística? Obtenido de <https://aws.amazon.com/es/what-is/logistic-regression/>
- Aguero, L. (1 de marzo de 2014). Estrategia de fidelización de clientes. <https://repositorio.unican.es/xmlui/bitstream/handle/10902/4474/%5B2%5D%20Ag%C3%BCero%20Cobo%20L.pdf>
- Ahumada, J. (3 de marzo de 2003). *R para principiantes*. Obtenido de https://cran.r-project.org/doc/contrib/rdebuts_es.pdf
- Alcázar, P. (31 de octubre de 2023). *El Cliente*. Obtenido de https://www.edebe.com/ciclosformativos/zona-publica/UT09_830030_LA_CEyAC_CAS.pdf
- Aprende Machine Learning. (13 de mayo de 2018). *Regresión Lineal en español con Python*. Obtenido de <https://www.aprendemachinelearning.com/regresion-lineal-en-espanol-con-python/>
- Artyco. (18 de mayo de 2017). *Cómo clasificar y segmentar a tus clientes*. Obtenido de <https://artyco.com/como-clasificar-segmentar-clientes/>
- Astorga, J. (1 de diciembre de 2014). *Aplicación de modelos de regresión lineal para determinar las armónicas de tensión y corriente*. Obtenido de http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S1815-59012014000300008
- Avinews. (22 de octubre de 2023). *La inteligencia artificial permitirá una avicultura más sostenible y con mayor bienestar animal*. Obtenido de <https://avinews.com/la-inteligencia-artificial-permitira-una-avicultura-mas-sostenible-y-con-mayor-bienestar-animal/>
- Baking Business. (12 de enero de 2022). *Nestlé emplea inteligencia artificial y aprendizaje automático para mejorar la innovación*. Obtenido de

<https://www.bakingbusiness.com/articles/57913-nestle-employing-ai-machine-learning-to-improve-innovation>

Cámara Madrid. (30 de diciembre de 2019). Obtenido de <https://www.mba-madrid.com/empresas/impacto-del-machine-learning-ambito-empresarial/>

Centro de Negocios Zaragoza. (17 de Diciembre de 2023). La localización de la empresa y su importancia. Obtenido de <https://centrodenegocioszaragoza.es/localizacion-empresa-importancia/>

Ceyca, S. (7 de agosto de 2019). *Tipos de compradores en Amazon*. Obtenido de <https://fideliza.com.mx/2019/08/07/tipos-de-compradores-en-amazon/>

Cordero, B. (1 de diciembre de 2021). *Algoritmos de Aprendizaje Supervisado para Proyección de Ventas de Camarón Ecuatoriano con Lenguaje de Programación Python*. Obtenido de <https://revistas.ute.edu.ec/index.php/economia-y-negocios/article/view/996/883>

Datision. (5 de agosto de 2021). *10 aplicaciones de machine Learning en los negocios*. Obtenido de <https://datision.com/blog/aplicaciones-machine-learning-negocios/>

Duckerman, W. (2 de marzo de 2021). *¿Cómo utilizar el Machine Learning para la segmentación de clientes?* Obtenido de <https://es.linkedin.com/pulse/c%C3%B3mo-utilizar-el-machine-learning-para-la-de-clientes-n%C3%A9stor-wario>

Elsevier. (1 de agosto de 2010). Modelo de regresión logística para estimar la dependencia según la escala de Lawton y Brody. Obtenido de [https://www.elsevier.es/es-revista-medicina-familia-semergen-40-articulo-modelo-regresion-logistica-estimar-dependencia-S1138359310001322#:~:text=La%20Regresi%C3%B3n%20Log%C3%ADstica%20Binaria%20\(RLB,objetivo%20de%20obtener%20una%20estimaci%C3%B3n](https://www.elsevier.es/es-revista-medicina-familia-semergen-40-articulo-modelo-regresion-logistica-estimar-dependencia-S1138359310001322#:~:text=La%20Regresi%C3%B3n%20Log%C3%ADstica%20Binaria%20(RLB,objetivo%20de%20obtener%20una%20estimaci%C3%B3n)

Ferre, M. (4 de abril de 2019). Regresión logística. Obtenido de <https://gauss.inf.um.es/feir/45/>

Gobierno de España. (19 de abril de 2023). *Qué es la Inteligencia Artificial*. Obtenido de <https://planderecuperacion.gob.es/noticias/que-es-inteligencia-artificial-ia-prtr>

- Godás, L. (1 de octubre de 2005). *La importancia del cliente*. Obtenido de <https://www.elsevier.es/es-revista-offarm-4-articulo-la-importancia-del-cliente-13079599>
- Gómez, D. (4 de octubre de 2023). 6 razones de pérdida de clientes (y cómo evitarla). <https://blog.hubspot.es/service/motivos-perdida-de-clientes>
- González, L. (28 de junio de 2019). Regresión Logística – Teoría. Obtenido de <https://aprendeia.com/algorithm-regresion-logistica-machine-learning-teoria/>
- González, R. (22 de marzo de 2019). *¿Qué es R y RStudio?* Obtenido de https://gonzalezgouveia.com/que-es-r-y-rstudio/#Que_es_R_y_RStudio
- Graph Everywhere*. (2022). Obtenido de <https://www.grapheverywhere.com/algoritmos-de-machine-learning/#:~:text=Los%20algoritmos%20de%20machine%20learning,el%20comportamiento%20de%20los%20datos.>
- Harvard. (1 de Julio de 2015). *Exportaciones*. Obtenido de <https://scholar.harvard.edu/files/vrios/files/201507-mexicoexports.pdf>
- IBM. (14 de octubre de 2023). *¿Qué es Machine Learning?* Obtenido de <https://www.ibm.com/mx-es/analytics/machine-learning>
- Kuuse, M. (13 de febrero de 2023). *¿Qué es la cantidad económica de pedido (EOQ) y la fórmula EOQ?* Obtenido de <https://manufacturing-software-blog.mrpeasy.com/es/la-cantidad-economica-de-pedido-eoq/>
- INCyTU. (1 de marzo de 2018). *Inteligencia Artificial*. Obtenido de https://www.foroconsultivo.org.mx/INCyTU/documentos/Completa/INCYTU_18-012.pdf
- Larisparto. (1 de octubre de 2021). *¿Qué es y para qué sirve una regresión lineal en machine Learning?* Obtenido de <https://platzi.com/blog/que-es-regresion-lineal/>
- Melior Centros de Negocios. (16 de Julio de 2014). La importancia de la ubicación de tu empresa. Obtenido de <https://www.melior.es/noticias/importancia-ubicacion-empresa>

- Miranda, M., Ramos, M., & Segovia, M. (enero de 2013). Técnicas de inteligencia artificial aplicadas a la resolución de problemas económico-financieros: análisis de los factores determinantes del éxito exportador. México: ResearchGate.
- Muller, P. (7 de marzo de 2023). *Cómo mejorar la experiencia para el cliente con el Machine Learning*. Obtenido de <https://www.zendesk.com.mx/blog/de-que-manera-las-empresas-pueden-mejorar-la-experiencia-para-el-cliente-con-el-aprendizaje-de-maquinas/>
- Murillo, L. (1 de enero de 2018). Fundamentos de servicio al cliente. <https://digitk.areandina.edu.co/bitstream/handle/areandina/3523/93%20FUNDAMENTOS%20DE%20SERVICIO%20AL%20CLIENTE.pdf?sequence=1>
- Ortega, C. (12 de diciembre de 2023). Frecuencia de compra: Qué es, importancia y cómo calcularla. Obtenido de <https://www.questionpro.com/blog/es/frecuencia-de-compra/#:~:text=%C2%BFQu%C3%A9%20es%20la%20frecuencia%20de,la%20pertinencia%20de%20los%20resultados.>
- Ortega, Ochoa, & Molina. (14 de febrero de 2022). Regresión logística binaria simple. Obtenido de <https://evidenciasenpediatria.es/articulo/7943/regresion-logistica-binaria-simple>
- Ospina, A. (24 de Julio de 2020). Fidelización del cliente: concepto, importancia, consejos y métricas. <https://www.rdstation.com/blog/es/fidelizacion-del-cliente/>
- Pandey, A. (Agosto de 2023). Machine Learning. International Journal for Research in Applied Science and Engineering Technology. *Machine Learning*, 864.
- Páramo, L. (1 de diciembre de 2017). *Proceso de Exportaciones*. Obtenido de <https://core.ac.uk/download/pdf/326424075.pdf>
- UASB. (1 de abril de 2021). Observatorio de la PyME. https://www.uasb.edu.ec/observatorio-pyme/wp-content/uploads/sites/6/2021/04/faq_53.pdf
- UNFPA Ecuador. (29 de octubre de 2020). Adaptar la Estrategia VIP a un formato accesible para distintos tipos de discapacidad. <https://untalent.org/jobs/adaptar-la-estrategia-vip-a-un-formato-accesible-para-distintos-tipos-de-discapacidad>

- Rodríguez, J. (20 de enero de 2023). *¿Qué es un cliente? Definición e importancia en las empresas*. Obtenido de <https://blog.hubspot.es/sales/que-es-un-cliente>
- Scikit Learn. (1 de enero de 2020). Trazar la regresión logística multinomial y de One-vs-Rest. Obtenido de https://qu4nt.github.io/sklearn-docs/auto_examples/linear_model/plot_logistic_multinomial.html
- Stack, P. (13 de octubre de 2021). *Métodos de aprendizaje supervisado y no supervisado para la estimación de Microestructura Cerebral de datos en DWMR*. Obtenido de <https://cimat.repositorioinstitucional.mx/jspui/bitstream/1008/1129/1/TE%20835.pdf>
- Toledo, W. (28 de marzo de 2017). *El rol de las Exportaciones en el crecimiento económico: Evidencia de una muestra de países de América Latina y el Caribe*. Obtenido de <https://www.scielo.org.mx/pdf/remy/v34n89/2395-8715-remy-34-89-78.pdf>
- Torres, L. (23 de enero de 2024). *¿En qué consiste la regresión logística? ¿Qué es la regularización?* Obtenido de <https://www.themachinelearners.com/regresion-logistica-regularizacion/>
- UCSP. (1 de diciembre de 2023). Costos logísticos: ¿qué son y qué tipos existen? Obtenido de <https://postgrado.ucsp.edu.pe/articulos/que-es-costo-logistico/#:~:text=Los%20costos%20log%C3%ADsticos%20son%20todos,y%20todos%20los%20pasos%20intermedios.>
- UPBE. (14 de octubre de 2023). *¿Cuántos tipos de clientes existen según su comportamiento y satisfacción?* Obtenido de <https://www.upbe.ai/blog/tipos-clientes/>
- Vicente, K. (1 de enero de 2021). *Efecto de las exportaciones, el consumo y la inversión en el crecimiento económico del Perú durante el Periodo 1985 a 2018*. Obtenido de <https://www.urp.edu.pe/pdf/id/34322/n/4-vicente-karin-modelo-vec-pbi.pdf>
- Folgueiras Bertomeu, P. (2016). La entrevista.
- Morales, F. (2012). Conozca 3 tipos de investigación: Descriptiva, Exploratoria y Explicativa. Recuperado el, 11(3).

Sampieri, R., Fernández, C., & Baptista, L. (2014). Definiciones de los enfoques cuantitativo y cualitativo, sus similitudes y diferencias. *RH Sampieri, Metodología de la Investigación*, 22.

Peláez, A., Rodríguez, J., Ramírez, S., Pérez, L., Vázquez, A., & González, L. (2013). La entrevista. Universidad autónoma de México. [En línea]. [Online]. [cited 2012 septiembre 30]. Disponible en: http://www.uam.es/personal_pdi/stmaria/jmurillo/InvestigacionEE/Presentaciones/Curso_10/E

Zendesk. (8 de marzo de 2023). ¿Cómo atender a un cliente VIP? Guía práctica para darle una atención excepcional. <https://www.zendesk.com.mx/blog/como-atender-a-un-cliente-vip-guia-practica-para-darle-una-atencion-excepcional/#:~:text=El%20acr%C3%B3nimo%20VIP%20se%20refiere,especiales%20brind%C3%A1ndoles%20una%20atenci%C3%B3n%20superior.>

Zendesk. (19 de diciembre de 2023). 7 estrategias de fidelización de clientes que dan resultado. <https://www.zendesk.com.mx/blog/estrategias-fidelizar-clientes/>

ANEXOS

Anexo 1. Base de datos

```
DataBaseValExp<-read.csv("../Desktop/Aprobacion_De_Expo.csv")
View(DataBaseValExp)
#regresion logistica o regresion logaritmica

#SE MANEJA DE FORMA BINOMIAL

#Edicion de variables de numericas a categoricas
#DATA CLEAN

DataBaseValExp$LogisticCost<-factor(DataBaseValExp$LogisticCost,
                                     levels = c(0,1),
                                     labels = c("Pack 1","Pack 2"))

DataBaseValExp$Financiamiento<-factor(DataBaseValExp$Financiamiento,
                                       levels = c(0,1),
                                       labels = c("No","Si"))

DataBaseValExp$Aprobacion<-factor(DataBaseValExp$Aprobacion,
                                   levels = c(0,1),
                                   labels = c("No","Si"))

#Sembrando semilla
set.seed(123)
|
#ENTRENAMOS LOS DATOS

library(caret)

Entrenamiento<-createDataPartition(DataBaseValExp$Aprobacion,
                                    p=0.8, list = FALSE)
```

```

#CREAMOS EL MODELO

#se pone esa linea con alt 126
Logit<-glm(Approved~.,
           data = DataBaseValExp[Entrenamiento,],
           family = binomial)

summary(Logit)

par(mfrow=c(2,2))
plot(Logit)

###MODELO OPTIMIZADO###

Logit2<-glm(Approved~FinancialLiquidity+Creditscore+OrderAmount,
            data = DataBaseValExp[Entrenamiento,],
            family = binomial)

summary(Logit2)

par(mfrow=c(2,2))
plot(Logit2)

#CALCULAMOS LA PROBABILIDAD DE EXITO DE CLASIFICACION

DataBaseValExp$Prob<-predict(Logit2, newdata = DataBaseValExp,
                             type = "response")

#Clasificamos
DataBaseValExp$Class<-ifelse(DataBaseValExp["Prob"]>=0.5,1,0)

#MATRIZ DE CONFUSION

table(DataBaseValExp[, "Approved"], DataBaseValExp[, "Class"],
      dnn = c("Actual", "Predicho"))

```

Anexo 2. Script en RStudio

	FinancialLiquidity	CreditScore	LogisticCost	OrderAmount	Approved	Prob	Class[, "Prob"]
318	No	0	Pack 2	100000	Si	1.00000000	1
133	Si	6	Pack 1	51100	Si	1.00000000	1
66	Si	1	Pack 2	50000	Si	1.00000000	1
7	Si	0	Pack 2	31285	Si	0.99999996	1
244	Si	5	Pack 1	26726	Si	0.99999985	1
246	Si	1	Pack 2	18027	Si	0.99996561	1
68	Si	3	Pack 1	15108	Si	0.99989340	1
181	Si	20	Pack 1	15000	Si	0.99999663	1
585	Si	0	Pack 1	13212	Si	0.99946009	1
118	Si	15	Pack 1	11202	Si	0.99992938	1
168	Si	0	Pack 1	11177	Si	0.99841711	1
165	Si	0	Pack 2	10561	Si	0.99780858	1
30	Si	6	Pack 2	10000	Si	0.99914489	1
238	Si	1	Pack 1	9800	Si	0.99733512	1
218	Si	5	Pack 1	8851	Si	0.99807132	1
576	Si	1	Pack 2	8000	Si	0.99312252	1
161	Si	4	Pack 1	7544	Si	0.99527879	1
231	Si	7	Pack 1	7059	Si	0.99671130	1
235	Si	13	Pack 1	6700	Si	0.99884587	1
206	Si	8	Pack 1	6590	Si	0.99657209	1
497	Si	0	Pack 2	5860	Si	0.97426567	1
53	Si	0	Pack 2	5800	Si	0.97345768	1
589	Si	5	Pack 2	5777	Si	0.99026870	1
285	No	2	Pack 1	5552	No	0.55288120	1
562	Si	0	Pack 1	5298	Si	0.96566067	1
446	No	0	Pack 1	5200	No	0.40444286	0

590	Si	7	Pack 2	5124	Si	0.99089893	1
33	Si	15	Pack 2	5000	Si	0.99812432	1
140	Si	2	Pack 1	5000	Si	0.97319333	1
506	Si	16	Pack 1	5000	Si	0.99847380	1
579	Si	10	Pack 1	4700	Si	0.99385227	1
493	Si	14	Pack 1	4607	Si	0.99716380	1
379	No	0	Pack 2	4500	No	0.31922511	0
362	No	0	Pack 2	4208	No	0.28662687	0
597	Si	8	Pack 2	4159	Si	0.98770517	1
549	Si	7	Pack 2	4071	Si	0.98422007	1
34	Si	0	Pack 2	4000	Si	0.93399803	1
343	No	0	Pack 2	4000	No	0.26466337	0
353	No	0	Pack 2	4000	No	0.26466337	0
456	No	0	Pack 1	3552	No	0.22116459	0
158	Si	14	Pack 1	3376	Si	0.99457428	1
581	Si	8	Pack 2	3290	Si	0.98066747	1
185	Si	7	Pack 1	3257	Si	0.97593056	1
149	Si	7	Pack 2	3065	Si	0.97342545	1
188	Si	7	Pack 1	3065	Si	0.97342545	1
184	Si	7	Pack 1	3000	Si	0.97252121	1
253	Si	11	Pack 1	3000	Si	0.98778288	1
559	Si	0	Pack 1	3000	Si	0.89290087	1
600	Si	11	Pack 2	3000	Si	0.98778288	1
69	Si	7	Pack 1	2954	Si	0.97186330	1
288	No	1	Pack 1	2803	No	0.19021272	0

127	Si		12	Pack 2		2510	Si	0.98713019		1
236	Si		5	Pack 1		2503	Si	0.94737267		1
560	Si		7	Pack 2		2384	Si	0.96233282		1
148	Si		14	Pack 1		2300	Si	0.99045253		1
42	Si		11	Pack 1		2283	Si	0.98224689		1
501	Si		3	Pack 2		2279	Si	0.91363685		1
496	Si		0	Pack 1		2206	Si	0.84561829		1
150	Si		14	Pack 1		2200	Si	0.98993903		1
528	Si		0	Pack 2		2197	No	0.84499565		1
154	Si		11	Pack 2		2184	Si	0.98131008		1
487	No		0	Pack 1		2100	No	0.11638675		0
558	Si		16	Pack 1		2079	Si	0.99288236		1
146	Si		6	Pack 1		2072	Si	0.94629076		1
173	Si		0	Pack 2		2028	Si	0.83292011		1
255	No		0	Pack 1		2010	No	0.11157859		0
160	Si		1	Pack 2		2000	Si	0.85792827		1
247	Si		1	Pack 2		2000	Si	0.85792827		1
492	Si		2	Pack 2		2000	Si	0.88129229		1
462	No		4	Pack 1		1950	No	0.21749672		0
232	Si		2	Pack 2		1704	Si	0.86390630		1
186	Si		9	Pack 1		1655	Si	0.96331290		1
152	Si		14	Pack 1		1602	Si	0.98624627		1
570	Si		4	Pack 1		1583	Si	0.89999569		1
575	Si		4	Pack 1		1465	Si	0.89423478		1
10	Si		0	Pack 2		1442	Si	0.78523375		1

189	Si	5	Pack 2	1430	Si	0.91074503	1
574	Si	11	Pack 2	1400	Si	0.97197176	1
500	Si	3	Pack 1	1391	Si	0.86864818	1
8	Si	0	Pack 1	1349	Si	0.77681961	1
554	Si	14	Pack 1	1344	Si	0.98426637	1
119	Si	6	Pack 2	1332	Si	0.92254606	1
365	No	2	Pack 1	1300	No	0.11534373	0
209	Si	7	Pack 1	1270	Si	0.93408419	1
23	Si	6	Pack 2	1260	Si	0.91977994	1
552	Si	11	Pack 1	1249	Si	0.96971133	1
47	Si	23	Pack 2	1236	Si	0.99736908	1
385	No	0	Pack 2	1212	No	0.07607480	0
210	Si	16	Pack 1	1210	Si	0.98877446	1
348	No	0	Pack 1	1210	No	0.07600046	0
21	Si	17	Pack 1	1208	Si	0.99084042	1
251	Si	11	Pack 1	1200	Si	0.96894055	1
591	Si	8	Pack 1	1200	Si	0.94378369	1
205	Si	1	Pack 1	1187	Si	0.79706578	1
372	No	0	Pack 2	1110	No	0.07236745	0
580	Si	5	Pack 2	1097	Si	0.89534931	1
176	Si	0	Pack 2	1065	Si	0.74969235	1
334	No	0	Pack 2	1062	No	0.07068106	0
432	No	0	Pack 1	1058	No	0.07054217	0
375	No	0	Pack 1	1004	No	0.06869180	0
126	Si	6	Pack 2	1000	Si	0.90902625	1

135	Si	12	Pack 2	1000	Si	0.97183273	1
137	Si	3	Pack 1	1000	Si	0.84319468	1
213	Si	15	Pack 2	1000	Si	0.98464187	1
259	No	0	Pack 1	1000	No	0.06855654	0
367	No	0	Pack 1	1000	No	0.06855654	0
474	No	0	Pack 2	1000	No	0.06855654	0
511	Si	2	Pack 2	1000	Si	0.81391556	1
587	Si	9	Pack 2	1000	Si	0.94889486	1
628	No	0	Pack 1	1000	No	0.06855654	0
524	Si	10	Pack 1	990	No	0.95781779	1
512	Si	0	Pack 1	960	Si	0.73912389	1
556	Si	6	Pack 1	948	Si	0.90672536	1
598	Si	3	Pack 1	918	Si	0.83737283	1
491	Si	1	Pack 1	892	Si	0.77064909	1
233	Si	4	Pack 2	857	Si	0.85973282	1
166	Si	0	Pack 1	837	Si	0.72638227	1
145	Si	9	Pack 1	827	Si	0.94426950	1
3	Si	0	Pack 1	824	Si	0.72501316	1
128	Si	7	Pack 2	809	Si	0.91738132	1
631	No	0	Pack 2	809	No	0.06237831	0
523	Si	1	Pack 1	800	Si	0.76193284	1
198	Si	12	Pack 1	790	Si	0.96862660	1
437	No	3	Pack 1	769	No	0.10796661	0
599	No	1	Pack 1	768	Si	0.07410458	0
689	No	0	Pack 1	750	No	0.06057738	0
214	Si	9	Pack 1	742	Si	0.94185510	1

62	Si		9	Pack 1		730	Si	0.94150643	1
266	No		0	Pack 1		722	No	0.05973982	0
37	Si		12	Pack 2		713	Si	0.96736468	1
518	Si		5	Pack 2		690	Si	0.87338824	1
459	No		0	Pack 1		687	No	0.05870812	0
201	Si		1	Pack 1		678	Si	0.75002754	1
169	Si		0	Pack 2		639	Si	0.70507618	1
129	Si		2	Pack 1		610	Si	0.78062618	1
139	Si		6	Pack 2		600	Si	0.88994460	1
193	Si		0	Pack 1		600	Si	0.70076746	1
203	Si		6	Pack 1		600	Si	0.88994460	1
507	Si		2	Pack 1		591	Si	0.77889988	1
636	No		2	Pack 2		587	No	0.08207300	0
147	Si		3	Pack 2		582	Si	0.81169199	1
123	Si		12	Pack 2		567	Si	0.96483603	1
2	Si		6	Pack 1		560	Si	0.88785468	1
35	Si		5	Pack 2		560	Si	0.86558527	1
199	Si		3	Pack 2		560	Si	0.80990645	1
38	Si		2	Pack 2		551	Si	0.77523384	1
179	Si		4	Pack 1		540	Si	0.83826265	1
577	Si		11	Pack 1		540	Si	0.95652474	1
682	No		0	Pack 1		537	No	0.05447309	0
646	No		0	Pack 2		501	No	0.05350036	0
39	Si		2	Pack 1		500	Si	0.77049734	1
70	Si		1	Pack 2		500	Si	0.73195907	1

100	Si		2	Pack 2		500	No	0.77049734	1
187	Si		1	Pack 1		500	Si	0.73195907	1
220	Si		4	Pack 2		500	Si	0.83537286	1
234	Si		0	Pack 2		500	Si	0.68955682	1
419	No		2	Pack 1		500	No	0.07867135	0
508	Si		19	Pack 1		500	Si	0.99118325	1
522	Si		5	Pack 2		500	Si	0.86184887	1
601	Si		4	Pack 1		500	Si	0.83537286	1
277	No		0	Pack 1		484	No	0.05304674	0
488	No		0	Pack 2		475	No	0.05280806	0
561	Si		6	Pack 1		458	Si	0.88236794	1
67	Si		11	Pack 1		456	Si	0.95463867	1
383	No		0	Pack 1		456	No	0.05230751	0
249	Si		2	Pack 1		444	Si	0.76521628	1
293	No		1	Pack 1		444	No	0.06316822	0
63	Si		2	Pack 2		400	Si	0.76100817	1
101	Si		0	Pack 2		400	No	0.67811909	1
167	Si		0	Pack 2		400	Si	0.67811909	1
200	Si		9	Pack 1		396	Si	0.93098089	1
687	No		2	Pack 2		394	No	0.07470130	0
658	No		0	Pack 1		390	No	0.05060336	0
143	Si		8	Pack 1		375	Si	0.91561515	1
620	Si		0	Pack 1		375	No	0.67522527	1
134	Si		6	Pack 1		367	Si	0.87727799	1
681	No		0	Pack 1		364	No	0.04994656	0

486	No	2	Pack 1	351	No	0.07314394	0
647	No	0	Pack 1	351	No	0.04962120	0
455	No	0	Pack 2	350	No	0.04959626	0
638	No	0	Pack 1	350	No	0.04959626	0
547	Si	6	Pack 2	347	No	0.87613423	1
540	Si	0	Pack 1	340	No	0.67115137	1
548	Si	6	Pack 1	327	Si	0.87498133	1
555	Si	4	Pack 1	321	Si	0.82193192	1
452	No	0	Pack 1	316	No	0.04875519	0
9	Si	0	Pack 1	314	Si	0.66810825	1
40	Si	11	Pack 2	300	Si	0.95092746	1
48	Si	3	Pack 1	300	Si	0.78782164	1
56	Si	0	Pack 2	300	Si	0.66646379	1
121	Si	1	Pack 1	300	Si	0.71069690	1
202	Si	0	Pack 2	300	Si	0.66646379	1
245	Si	2	Pack 1	300	Si	0.75125331	1
439	No	1	Pack 1	300	No	0.05880695	0
510	Si	1	Pack 2	300	Si	0.71069690	1
377	No	0	Pack 1	286	No	0.04802432	0
46	Si	11	Pack 2	284	Si	0.95053093	1
567	Si	11	Pack 2	284	Si	0.95053093	1
613	No	1	Pack 1	283	No	0.05831110	0
144	Si	1	Pack 1	278	Si	0.70829788	1
122	Si	67	Pack 2	258	Si	0.99999950	1
336	No	0	Pack 2	251	No	0.04718479	0
142	Si	15	Pack 1	247	Si	0.97729591	1

675	No	0	Pack 1	246	No	0.04706600	0
18	Si	10	Pack 1	245	Si	0.93868772	1
676	No	0	Pack 1	237	No	0.04685290	0
521	Si	1	Pack 1	234	Si	0.70346493	1
337	No	0	Pack 2	228	No	0.04664072	0
95	Si	0	Pack 2	225	No	0.65758582	1
41	Si	12	Pack 2	221	Si	0.95806836	1
115	Si	3	Pack 2	210	No	0.77975313	1
281	No	0	Pack 1	204	No	0.04607935	0
317	No	0	Pack 1	204	No	0.04607935	0
12	Si	0	Pack 1	200	Si	0.65460142	1
54	Si	0	Pack 2	200	Si	0.65460142	1
542	Si	0	Pack 1	200	No	0.65460142	1
563	Si	0	Pack 1	200	Si	0.65460142	1
641	No	0	Pack 2	200	No	0.04598642	0
650	No	6	Pack 1	200	No	0.14269165	0
240	Si	6	Pack 1	196	Si	0.86720070	1
395	No	0	Pack 1	195	No	0.04587050	0
401	No	0	Pack 1	184	No	0.04561646	0
327	No	0	Pack 1	179	No	0.04550143	0
300	No	2	Pack 2	173	No	0.06701058	0
208	Si	6	Pack 1	168	Si	0.86548537	1
569	Si	8	Pack 2	162	Si	0.90649041	1
652	No	1	Pack 2	160	No	0.05483878	0
180	Si	14	Pack 1	158	Si	0.97093205	1
660	No	1	Pack 2	154	No	0.05467448	0
178	Si	0	Pack 1	150	Si	0.64859613	1
592	Si	3	Pack 2	150	Si	0.77425306	1
671	No	0	Pack 1	150	No	0.04483970	0
112	Si	3	Pack 1	147	No	0.77397552	1
411	No	5	Pack 1	146	No	0.11627189	0
642	No	0	Pack 2	141	No	0.04463620	0
402	No	0	Pack 1	140	No	0.04461365	0
553	Si	9	Pack 2	134	Si	0.92152424	1
106	Si	20	Pack 1	130	No	0.99127696	1
295	No	1	Pack 1	126	No	0.05391385	0
347	No	0	Pack 2	122	No	0.04420949	0
263	No	0	Pack 1	120	No	0.04416480	0
673	No	0	Pack 2	117	No	0.04409785	0
320	No	0	Pack 2	113	Si	0.04400872	0
366	No	1	Pack 2	112	No	0.05353729	0
551	Si	1	Pack 2	109	Si	0.68948632	1
616	No	2	Pack 2	108	No	0.06489229	0
632	No	1	Pack 2	108	No	0.05343015	0
287	No	2	Pack 2	105	No	0.06479605	0
43	Si	1	Pack 1	100	Si	0.68846597	1
94	Si	0	Pack 2	100	No	0.64254344	1
342	No	0	Pack 1	100	No	0.04372026	0
472	No	0	Pack 1	100	No	0.04372026	0



DECLARACIÓN Y AUTORIZACIÓN

Nosotros, **Delgado Moreira, Antonella Nicole** con C.C: # 0932411788 y **Moreno Dazza, Víctor Israel**, con C.C # 0931605505 autores del trabajo de titulación curricular: **Implementación de machine learning para clasificación de clientes en la empresa VALENTINA EXPORT S.A.**, previo a la obtención del título de **Licenciado en Negocios Internacionales** en la Universidad Católica de Santiago de Guayaquil.

1.- Declaro tener pleno conocimiento de la obligación que tienen las instituciones de educación superior, de conformidad con el Artículo 144 de la Ley Orgánica de Educación Superior, de entregar a la SENESCYT en formato digital una copia del referido trabajo de titulación para que sea integrado al Sistema Nacional de Información de la Educación Superior del Ecuador para su difusión pública respetando los derechos de autor.

2.- Autorizo a la SENESCYT a tener una copia del referido trabajo de titulación, con el propósito de generar un repositorio que democratice la información, respetando las políticas de propiedad intelectual vigentes.

Guayaquil, 16 de febrero del 2024

f. _____

Delgado Moreira, Antonella Nicole

C.C: 0932411788

f. _____

Moreno Dazza, Víctor Israel

CC: 0931605505



REPOSITORIO NACIONAL EN CIENCIA Y TECNOLOGÍA

FICHA DE REGISTRO DE TESIS/TRABAJO DE TITULACIÓN

TEMA Y SUBTEMA:	Implementación de machine learning para clasificación de clientes en la empresa VALENTINA EXPORT S.A.		
AUTOR(ES)	Delgado Moreira, Antonella Nicole Moreno Dazza, Víctor Moreno		
REVISOR(ES)/TUTOR(ES)	Ing. Arias Arana, Wendy Vanessa, Mgs.		
INSTITUCIÓN:	Universidad Católica de Santiago de Guayaquil		
FACULTAD:	Facultad de Economía y Empresa		
CARRERA:	Negocios Internacionales		
TÍTULO OBTENIDO:	Licenciado en Negocios Internacionales		
FECHA DE PUBLICACIÓN:	16 de febrero del 2024	No. DE PÁGINAS:	88
ÁREAS TEMÁTICAS:	Economía, Innovaciones, Negocios		
PALABRAS CLAVES/ KEYWORDS:	Machine Learning, Clasificación de clientes, Regresión logística, Exportación.		
RESUMEN	<p>En esta tesis se detalla la aplicación de la regresión logística como una solución innovadora para segmentar a los clientes de Valentina Export S.A. La metodología utilizada implica la transformación de variables y la limpieza de datos para garantizar la calidad y precisión del modelo. La regresión logística se presenta como una herramienta clave para predecir la probabilidad de que un cliente sea "Apto para Exportar" o "No Apto para Exportar" en función de diversos parámetros, como historial de compras, liquidez financiera y tamaño del pedido. La implementación práctica de este modelo no solo ha permitido optimizar la asignación de productos a clientes estratégicos, sino que también ha posibilitado la creación de estrategias más efectivas para la toma de decisiones y la atención personalizada al cliente. Este enfoque de machine learning no solo busca mejorar la eficiencia operativa de Valentina Export S.A., sino también brindar una visión informada para la toma de decisiones en un entorno de demanda dinámica y desafiante.</p>		
ADJUNTO PDF:	<input checked="" type="checkbox"/> SI	<input type="checkbox"/> NO	
CONTACTO CON AUTOR/ES:	Teléfono: +593 95 989 5722 +593 96 981 7418	E-mail: Víctor Israel Moreno Dazza Vimd1493@gmail.com Antonella Nicole Delgado Moreira. Antonellanicoledegado03@gmail.com	
CONTACTO CON LA INSTITUCIÓN (COORDINADOR DEL PROCESO UTE)::	Nombre: Freire Quintero, Cesar Enrique		
	Teléfono: +593 99 009 0702		
	E-mail: cesar.freire@cu.ucsg.edu.ec		
SECCIÓN PARA USO DE BIBLIOTECA			
Nº. DE REGISTRO (en base a datos):			
Nº. DE CLASIFICACIÓN:			
DIRECCIÓN URL (tesis en la web):			