



UNIVERSIDAD CÉSAR VALLEJO

**FACULTAD DE INGENIERÍA Y ARQUITECTURA
ESCUELA PROFESIONAL DE INGENIERÍA DE SISTEMAS**

**Machine Learning Para La Gestión De Ventas En La
Empresa Vértice Empresarial S.A.C**

TESIS PARA OBTENER EL TÍTULO PROFESIONAL DE:

Ingeniero de Sistemas

AUTOR:

Castañeda Rojas, Miguel Angel David (ORCID: 0000-0003-3083-1182)

ASESOR:

MG. Perez Farfan Ivan Martin (ORCID: 0000-0001-5833-9400)

LÍNEA DE INVESTIGACIÓN:

Sistemas de Información y Comunicaciones.

LIMA – PERÚ

2020

DEDICATORIA

Dedicado a las personas que nunca perdieron las esperanzas en mí a pesar de las circunstancias de los problemas.

AGRADECIMIENTO

Primero a Dios, Mis Padres y todas las personas que me brindaron su apoyo en todo momento.

PRESENTACIÓN

La siguiente investigación abarca el tema de Machine Learning, y sobre su uso para mejorar la productividad y eficiencia mediante el uso de algoritmos lógicos. El problema consistía en que se tenía el proceso de gestión de ventas no sistematizado por parte del área de sistemas además tenía retrasos al momento de realizar dichas gestiones o reportes. Por estos motivos se decidió utilizar la metodología CRISP DM, ya que se manejaban datos en gran cantidad y como sabemos esta metodología trabaja minuciosamente con ellos.

La implementación de Machine Learning fue de gran ayuda dando resultados positivos, además sistematizando los procesos para lograr una buena gestión de ventas.

En conclusión, la implementación de Machine Learning aportó un gran margen de mejora en los indicadores productividad y la eficiencia en los asesores para la gestión de ventas, así como en el área de sistemas en la sistematización de los reportes en un menor tiempo posible.

Índice de contenidos

Dedicatoria	ii
Agradecimiento	iii
Índice de contenidos.....	iv
Índice de tablas	v
Índice de figuras.....	vii
Resumen.....	vi
Abstract.....	vii
I.INTRODUCCIÓN	1
II. MARCO TEÓRICO	5
III. METODOLOGÍA	19
3.1 Tipo y diseño de investigación.....	20
3.2 Variables y Operacionalización	21
3.3 Población muestra y muestreo.....	23
3.4 Técnicas e Instrumentos de Recolección de Datos	24
3.5 Procedimientos.....	28
3.6 Método de análisis de datos	28
3.5 Aspectos éticos	30
IV.RESULTADOS.....	31
V.DISCUSION	46
VI. CONCLUSIONES	48
VII. RECOMENDACIONES.....	50
REFERENCIAS	52
ANEXOS.....	56

Índice Tablas

Tabla 1 Selección de Metodología	18
Tabla 2 Validación del indicador Productividad	25
Tabla 3 Validación del indicador Eficiencia	26
Tabla 4 Confiabilidad del Indicador Eficiencia	27
Tabla 5 Confiabilidad del Indicador Productividad	26
Tabla 6 Medidas Descriptivas del nivel de Eficiencia antes y después de la implementación de Machine Learning	32
Tabla 7 Medidas Descriptivas del nivel de Productividad antes y después de la implementación de Machine Learning	33
Tabla 8 Prueba Normalidad Eficiencia	35
Tabla 9 Prueba Normalidad Productividad	37
Tabla 10 T Student Eficiencia	40
Tabla 11 T Student Productividad	43

Índice de figuras

Figura 1: Productividad Diaria	2
Figura 2: Eficiencia Diaria	3
Figura 3: Metodología CRISP DM	18
Figura 4: Confiabilidad del Indicador Eficiencia	27
Figura 5: Confiabilidad del Indicador Productividad	27
Figura 6: Formula Para Hallar el T-Student	28
Figura 7: Distribución según T-Student	29
Figura 8: Índice de Nivel de Eficiencia Antes y Después de la implementación de Machine Learning	33
Figura 9: Índice de Nivel de Productividad Antes y Después de la implementación de Machine Learning	34
Figura 10: Eficiencia PreTest	36
Figura 11: Eficiencia PostTest	36
Figura 12: Productividad PreTest	38
Figura 13: Productividad PostTest	38
Figura 14: Nivel De Eficiencia	39
Figura 15: T Student Eficiencia	41
Figura 16: Distribución T Student	41
Figura 17: Nivel De Productividad	42
Figura 18: T Student Productividad	44
Figura 19: Distribución T Student	44

RESUMEN

El proyecto de investigación llamado Machine Learning Para La Gestión De Ventas En la Empresa Vértice Empresarial S.A.C. consta de una implementación del sistema de aprendizaje automático Machine Learning, donde agiliza el proceso de ventas, el cual tenía deficiencias en sus procesos de servicio como en la búsqueda de clientes. En esta investigación se utilizó la metodología CRISP DM ya que se trabajó con gran cantidad de datos y se acomodaba a las etapas y necesidades del proyecto. Con un gestor de Datos Sql server 2017 y el sistema de programación Python.

La investigación que se realizara es pre-experimental y el enfoque a realizar es cuantitativo. La implementación de recolección de datos fue realizada mediante el fichaje y encuestas, además de una previa encuesta para identificar los problemas frecuentes.

Se concluye que el Machine Learning influyo positivamente en la gestión de ventas en la Empresa Vértice Empresarial S.A.C.

PALABRAS CLAVES:

Machine Learning, Gestión de Ventas, CRIPS DM

ABSTRACT

The research project called Machine Learning for Sales Management at the company Vértice Empresarial S.A.C. constant in the implementation of a machine learning machine learning system, for the improvement of the sales process, the quality deficiencies in its service processes as in the search for customers. In this investigation, verify the CRISP DM methodology, since it will be seen with a large amount of data and will adapte the stages and needs of the project. With a SQL Server 2017 Data manager and the Python programming system.

The research carried out is pre-experimental and the approach to be carried out is quantitative. The data collection technique was carried out through the signing and surveys, in addition to a previous survey to identify frequent problems.

It is concluded that Machine Learning positively influences sales management in the company Vértice Empresarial S.A.C

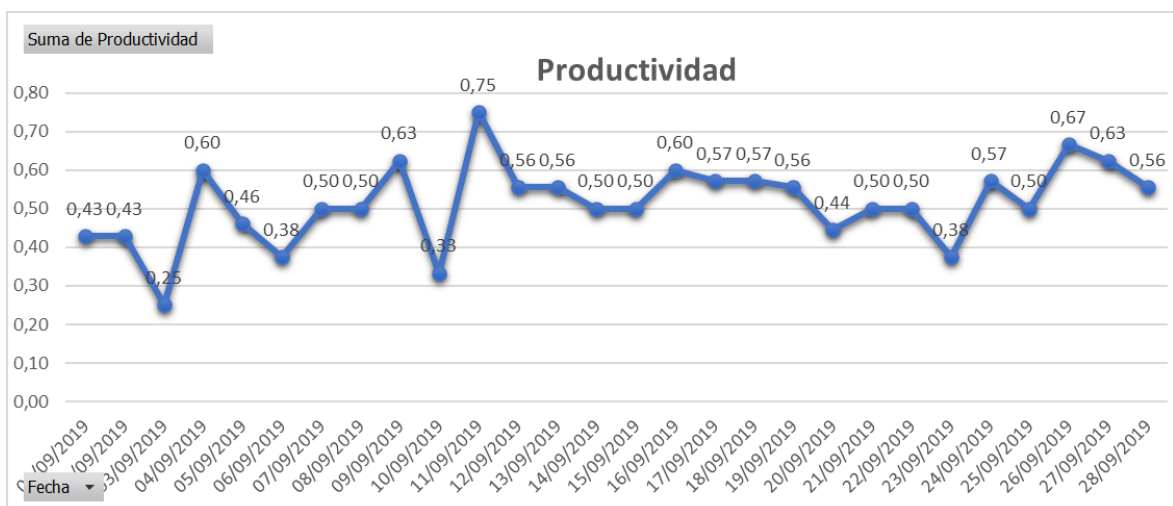
KEYWORDS:

Machine Learning, Sales Management, CRIPS DM

I.INTRODUCCIÓN

La empresa Vértice Empresarial se encuentra establecida en el sector de ventas por telecomunicaciones. El área de operaciones se encarga de velar por la productividad en la empresa. Así mismo el área de gestión se encarga de contactar la mayor cantidad de clientes por llamadas telefónicas mediante una plataforma score, en esta plataforma son ingresados los números telefónicos manualmente por los asistentes de sistemas. El asesor esperaba que ingrese una llamada eficiente la cual generaba futuros retrasos en la operación al momento de su gestión de venta. Mientras tanto los asistentes de sistemas buscaban los números eficientes en gestiones anteriores tanto en las bases o archivos de Excel. Todo este proceso provocaba que la eficiencia de la gestiones sea baja y no se logre realizar muchas ventas o tipos de servicios, ya que al llamar a todos los números del cliente para tener un contacto eficiente podía tomar días. Así mismo la realización de reportes, solicitados por el área de operaciones demoraba debido a que no se tenía recursos suficientes para su elaboración. Estos informes eran generados manualmente en hojas de cálculo Excel por cada asistente de sistemas, los cuales tienen a cargo para su análisis. Una vez generados dichos reportes eran enviados al área de operaciones para que evalúen y determinen dónde enfocar la productividad de sus ventas.

Figura 1: Productividad Diaria

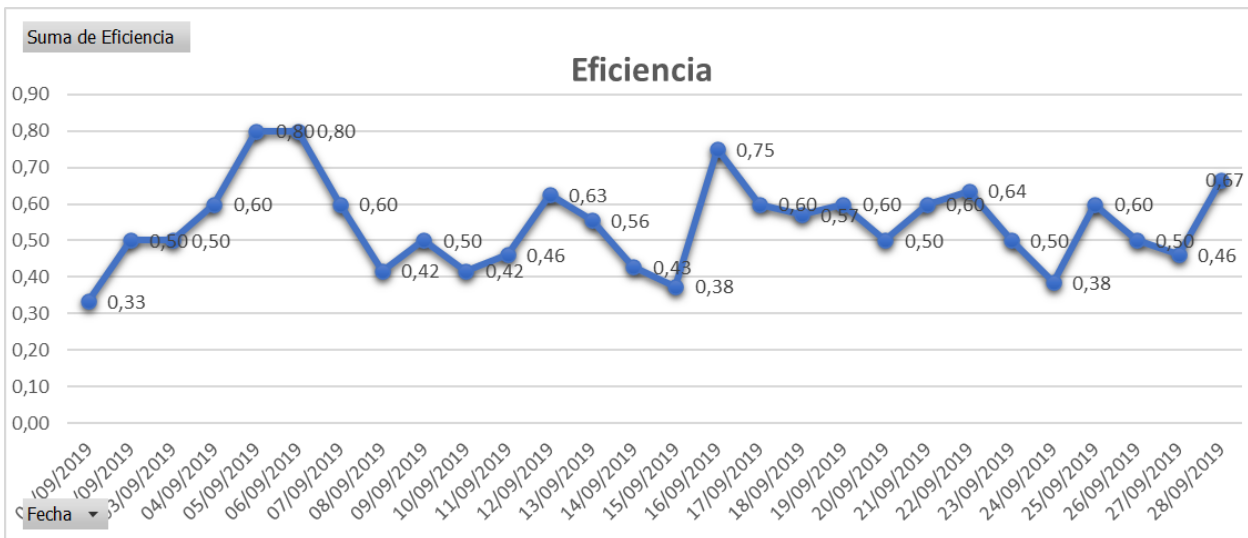


Fuente: Elaboración propia

La elaboración de este reporte (Figura1) se realizó a una data de clientes para visualizar la productividad diaria de llamadas por cartera, en un rango

tiempo de un mes, donde la línea azul equivale al número de llamadas eficientes logradas por día. La medición pretest del indicador productividad arrojó un resultado del 0.55% en un lapso de tiempo del 1 al 28 de setiembre del 2019 en la figura1 se muestra la productividad obtenida.

Figura 2: Eficiencia Diaria



Fuente: Elaboración propia

La elaboración de este cuadro de Productividad (Figura2) se realizó a una data de clientes para visualizar la eficiencia de llamadas por cartera, en un rango tiempo del 1 al 28 de setiembre y así saber la eficiencia de los colaboradores. La cual muestra que el picos bajos y altos de las llamadas tiene una mayor eficiencia. La medición pretest del indicador productividad arrojó un resultado del 0.51% en un lapso de tiempo del 1 al 28 de setiembre del 2019 en la figura1 se muestra la eficiencia obtenida. En el plano nacional, según MENDOZA RIOFRIO, Marcela (2019) en la revista El Comercio indica que, “El estudio aplicado mostro que a su vez en las áreas de comercio de las empresas dieron avances muy interesantes, pero las industrias de consumo masivo aún se mantienen con los niveles básicos para su automatización. Además, no se logra de manera la digitalización en los procesos comerciales, productivos y no se ha podido sacar provecho a tecnologías como la inteligencia artificial”. (p.1).¹

¹MENDOZA RIOFRÍO, Marcela. Avances hacia la Industria 4.0 aún son tímidos.2019.[Consulta 12 de abril del 2020].Disponible en <https://elcomercio.pe/economia/dia-1/avances-hacia-la-industria-40-aun-son-timidos-noticia/>.

La empresa Vértice Empresarial presento la siguiente problemática general: ¿De qué manera Machine Learning influye para la gestión en la Empresa Vértice Empresarial S.A.C? y los siguientes problemas específicos. 1). ¿En qué medida Machine Learning influye en la eficiencia para la gestión de ventas en la empresa Vértice Empresarial S.A.C? 2). ¿En qué medida Machine Learning influye en la productividad para la Gestión de Ventas en la Empresa Vértice Empresarial S.A.C?

Después de todo lo investigado se planteó un objetivo general: Determinar la Influencia de Machine Learning para la Gestión de ventas en la Empresa Vértice Empresarial. Los siguientes objetivos específicos:1) Determinar la Influencia de Machine Learning en la eficiencia para la gestión de ventas en la empresa Vértice Empresarial S.A.C. 2). Determinar la influencia de Machine Learning en la productividad para la Gestión de Ventas en la Empresa Vértice Empresarial S.A.C.

Los objetivos presentados permitieron aplicar la siguiente Hipótesis general: Machine Learning mejora la Gestión de Ventas en la Empresa Vértice Empresarial S.A.C. Y las siguientes hipótesis específicas: 1). Machine Learning aumenta la eficiencia para la Gestión de Ventas en la Empresa Vértice Empresarial S.A.C. 2). Machine Learning aumenta la productividad para la Gestión de Ventas en la empresa Vértice Empresarial S.A.C.

Desde el criterio en la relevancia social, la investigación pretende brindar una manera de ahorrar recursos al momento de realizar reportes o búsqueda de clientes por parte de los asesores, esta investigación beneficiara a los asesores en la realización del recurso de Gestión de Ventas de la Empresa Vértice Empresarial donde ellos puedan tener un mayor contacto con los clientes. Implicaciones prácticas, la implementación de Machine Learning ayudara a la sistematización de los reportes de llamas las cual eran entregadas fuera de tiempo. Para las cuales se usará un servicio de análisis de datos. En el valor teórico, se logrará aumentar los conocimientos sobre nuevas tecnologías de información como Machine Learning en la Pymes y la utilización CRISP DM como metodología para la preparación de los datos que existen en el entorno donde se trabajó, cuyos resultados explicaran con gran relevancia la aplicación de Machine Learning para futuras investigaciones.

II. MARCO TEÓRICO

En el año 2017, RODRIGUEZ CASTILLO, Jorge Junior realizó la investigación titulada “Desarrollo de la Aplicación informática está basada en el modelo Machine Learning donde mejora la evaluación de préstamos crediticios” (Tesis para optar grado de Ingeniero de sistemas) en la Universidad Privada Del Norte Trujillo. Esta tesis está enfocada para el estudio Machine Learning para así desarrollar la aplicación de aprendizaje, el cual va a permitir aumentar la evaluación de préstamos crediticios y además poder brindar un superior enfoque de la rentabilidad y el riesgo crediticio. Donde el problema radicaba en cómo saber que una aplicación de aprendizaje, toma un modelo de Machine Learning contribuía a la mejorar para evaluación de los préstamos crediticios brindados por la financiera. Se elaboro una interfaz guiada en un modelo de regresión logística la cual pudo permitir saber a detalle los pormenores de algunos préstamos crediticios por medio de la interfaz propuesta, en el cual se ingresan los factores necesarios para la evaluación como también; el total del monto que pidió, la tasa de interés, los meses del crédito, el estado civil y la edad del cliente. En el planteamiento se utilizó la regresión logística está se dividía en algoritmos esenciales; un algoritmo para la búsqueda de datos, un algoritmo para efectuar la predicción denominado función sigmoïdal, otra función en la cual se puede. Con la aplicación se bajó el uso de la regresión logística logrará aumentar las cantidades de dinero ganado, y minimizar la cantidad de dinero perdido y se logró reducir el porcentaje de tiempo para la evaluación y aprobación de estos préstamos crediticios. La tesis revisada tiene como tipo de prueba la investigación Pre-Experimental, en donde la población es ubicada en la ciudad de Trujillo, los instrumentos utilizados en la tesis son la encuesta y el fichaje. El sistema que se ejecutó fue, en general, para la mejora en los recursos de evaluación y aceptación de un préstamo crediticio. Por consiguiente, esta investigación nos da el aporte del uso de las la regresión logita como herramientas a base de su implementación se logró generar un modelo lineal basado en el objetivo de la variable.²

²RODRÍGUEZ CASTILLO, Jorge. Desarrollo de una Aplicación informática basada en un modelo Machine Learning para mejorar la evaluación de préstamos crediticios. 2017.[Consulta 26 de setiembre del 2019]. Disponible en en: <https://repositorio.upn.edu.pe/bitstream/handle/11537/12294/Rodr%C3%ADquez%20Castillo%2C%20Jorge%20Junior%20%20Mi%C3%B1ano%20Ochoa%2C%20Milagros%20Madeleine.pdf?sequence=1&isAllowed=y>.

En el año 2016, MEDRANO, Sandra donde investigo y planteo la tesis titulada “Modelo de minería de datos con Machine Learning para reconocimiento de patrones de síntomas y enfermedades respiratorias para las historias clínicas y realizar el diagnóstico de pacientes en la ciudad de Trujillo 2016” (Tesis para lograr el grado de Ingeniero de sistemas) en la Universidad Cesar Vallejo situada en Trujillo; Demuestra que este problema ha ido afectando a todos los que han ido buscando durante tanto tiempo poder efectuar una cultura la cual permita cambiar los tipos de trabajo y poder evaluar los acuerdos de los sistemas donde se trata de disminuir los posibles errores que se presenten de esta manera. Garantizando la recuperación del paciente por ejemplo en los Estados Unidos, México, entre otras entidades de velar por la seguridad de la salud. Donde toma el tema de Machine Learning para apoyo en los diagnósticos para el paciente. El problema se centra en no tener diagnóstico preciso que el medico entrega al paciente así dando más gastos en medicamentos y análisis innecesarios, además del desprestigio del doctor. Los Resultados de esta herramienta fueron exactos en los diagnósticos cumpliendo con el objetivo, donde se quería que los diagnósticos médicos fueran mejores, donde se abrió más puertas a otras especialidades para ampliar los resultados y no solo enfocarse en una sola especialidad del problema de las enfermedades. El tipo de investigación según lo investigado y la confirmación del procedimiento de prueba Pre-Experimental, la población son los pacientes de la ciudad de Trujillo, los instrumentos utilizados en el examen son la encuesta y la investigación narrativa. El sistema que se ejecutó fue, en general, se trata de la mejora de los hallazgos terapéuticos de un modelo programado que utiliza el cálculo del árbol de opciones. Por consiguiente, esta investigación nos aportó el uso de las redes neuronales como principal base de su implementación mediante patrones generando un valor agregado.³

³MEDRANO, Sandra. Modelo de minería de datos usando Machine Learning con reconocimiento de patrones de síntomas y enfermedades respiratorias en las historias clínicas para mejorar el diagnóstico de pacientes en la ciudad de Trujillo 2016.2016. [Consulta 12 de agosto del 2019]. Disponible en: <http://rpositorio.ucv.edu.pe/handle/UCV/9852>.

Según ASTO, Christian y VILCA, Jean 2018. Con la tesis "Búsqueda de Patrones de comportamiento usando Machine Learning, para la toma decisiones generales en la empresa Chuchuhuasi." Cusco.2018. Universidad Andina Del Departamento Cusco.

En esta tesis pudo desarrollar una aplicación la cual uso modelos de probabilidad, el cual eran parte del aprendizaje automático. En la cual permitía integrar secciones de comportamiento en procesos para la empresa como así también en distintos contextos, donde el aplicativo usaba un algoritmo de revisión de datos y podía predecir el comportamiento a futuro. Pudiendo recomendar alguna información usada con anterioridad para mejorar el servicio de venta de esta. Estos datos inducen a la mejor del comportamiento de la empresa tanto en la mejora de eficiencia y productividad de cada venta ya que a cada una en su área correspondiente con distintos tipos de procesos y requerimientos. Luego de la implementación de Machine Learning se mejoró los procesos y la toma de decisiones mediante el aplicativo web como dato adicional se mejoró el servicio de ventas por web ya que si el cliente tenía compras realizada con anterioridad se le ofrecía productos parecidos através del aplicativo web o mediante un aviso en su pantalla con algún tipo de oferta además tenía una estructura en la cual dependía mucho el horario de la compra y lo adicionales a lo que el cliente eligiera en ese momento. La investigación realizada fue cuantitativa y la recolección de datos fue mediante fichas de recolección y encuestas. Mediante el uso del algoritmo predictivo KVecino de Machine Learning, el cual realizo un promedio del punto más allegado al punto que se está buscando. El aporte para nuestra tesis fue el tipo de recolección de datos, así como el uso de datos anteriores para poder sugerirlos en procesos futuros además de las decisiones para la mejora de futuros procesos a raíz de datos ya almacenados en el aplicativo.⁴

⁴ ASTO, Christian Y VILCA, Jean. Búsqueda de Patrones de comportamiento usando Machine Learning, para la toma decisiones generales en la empresa Chuchuhuasi.2018. [Consulta 10 de abril del 2020]. Disponible en <http://repositorio.uandina.edu.pe/handle/UAC/2795>.

El año 2019 MONTIEL LOPEZ, Jacob. Realizo su investigación para optar su grado de Doctor titulada “Fast and Slow Machine Learning” en la universidad Paris-saclay.Francia. Esta tesis explora la simbiosis del aprendizaje en modo discontinuo y en flujo, tradicionalmente considerado en la literatura como antagonistas, sobre el problema de la clasificación de los flujos de datos en evolución. El aprendizaje en modo por lotes es un enfoque bien establecido basado en una secuencia finita: primero se recopilan los datos, luego se crean los modelos predictivos y finalmente se aplica el modelo. Las dos versiones del mecanismo de alerta temprana (lote y flujo) superan el rendimiento básico de la solución implementada por Groupe BPCE, la segunda institución bancaria más grande de Francia. Además, presentamos un método de imputación evolutivo basado en un modelo de datos faltantes en la clasificación. Este método presenta el problema de imputación en forma de un conjunto de tareas de clasificación / regresión resueltas gradualmente. Presentamos un marco unificado que sirve como una plataforma de aprendizaje común donde los métodos de procesamiento por lotes y flujo pueden interactuar positivamente. Mostramos que los métodos por lotes pueden ser entrenados efectivamente en la regulación del flujo bajo condiciones específicas. También ofrecemos una adaptación del algoritmo Extreme Gradient Boosting a los flujos de datos en evolución. El método adaptativo propuesto genera y actualiza el conjunto de forma incremental utilizando mini lotes de datos, El aporte para nuestro proyecto es programación un marco de Python de código abierto que llena el vacío en Python para una plataforma de desarrollo o investigación para aprender de flujos de datos en evolución.⁵

⁵MONTIEL LOPEZ, Jacob. Fast and Slow Machine Learning. 2019. [Consulta 16 de mayo del 2020]. Disponible en <https://pastel.archives-ouvertes.fr/tel-02098633/document>.

En el 2016, VOJT, Ján. Realizó su investigación de “Deep neural networks and their implementation”. En la Universidad Charles University in Prague. Donde las redes neuronales profundas son un modelo eficaz y universal para resolver una amplia variedad de tareas. Esta investigación se centra en tres diferentes tipos de redes neuronales:

El perceptron multicapa, la red neuronal convencional y la red de creencias profundas. Todos los modelos de red fueron implementados en algún sistema, y probado para distintas situaciones de la arquitectura de la red y sus parámetros. El sistema además de ser implementado fue documentado, donde detalla las decisiones optimizadas y propuestas planteadas para lograr la mejores de los procesos los cuales tenían grandes falencias. Se confirmó la eficiencia por los resultados usado en sus pruebas donde se representa en la tesis con pruebas adicionales de marcos ya existentes las cuales eran soportadas por redes neuronales profundas. Donde se ve un rendimiento superior a los marcos empleados por perceptrones multicapa y otros tipos de redes. La implementación de la red funciones de una mera ligera para capas con Rbm donde un aproximados de N neuronas ocultas, pero tiene un rendimiento inferior para una Rbm más concentrada. De esta investigación se logró obtener el uso de la toma de decisiones y la mejora y automatización de procesos a seguir en machine learning, en un rango de aceptación para los objetivos requeridos dentro de la empresa.⁶

⁶VOJT, Jan. Deep neural networks and their implementation.2016. Pp 45-250. [en línea]. [Consulta 16 de abril del 2020]. Disponible en <https://is.cuni.cz/webapps/zzp/download/120229251>.

En el año 2017 JALAL, Nouri y LARSSON, Kent. Realizo su investigación titulada “Bachelor Thesis Analytics: Using Machine Learning to Predict Dropout and Identify Performance Factors” en la Universidad del Este de Finlandia, Joensuu, Finlandia. La tesis estipula que la finalización de la tesis es un resultado deseable para estudiantes individuales, instituciones académicas y la sociedad. Desafortunadamente, muchas instituciones académicas de todo el mundo experimentan que muchos proyectos de tesis no se completan y que los estudiantes luchan con el proceso de tesis. Este trabajo aborda este problema con el objetivo de, por un lado, identificar y explicar por qué se completan o no los proyectos de tesis y, por otro lado, predecir la no finalización y finalización de los proyectos de tesis utilizando algoritmos de aprendizaje automático. Los datos se extrajeron de dos sistemas de datos diferentes utilizados para registrar datos sobre proyectos de tesis. A partir de estos sistemas, se recopilaron datos del proyecto de tesis, incluidas variables relacionadas tanto con los estudiantes como con los supervisores. El análisis estadístico tradicional (pruebas de correlación, pruebas t y análisis de factores) se realizó donde identificaron los factores que influían en la no finalización y finalización de los proyectos de tesis y se aplicaron varios algoritmos de aprendizaje automático para crear un modelo que prediga la finalización y la no terminación. Al tomar en cuenta todos los análisis mencionados anteriormente, Se puede concluir con confianza en que la capacidad y experiencia de los supervisores tuvo un rol importante en la determinación del éxito de los proyectos de tesis, lo que, por un lado, corrobora la investigación previa. Por otro lado, este estudio amplía la investigación previa al señalar factores específicos adicionales, como el tiempo que toman los supervisores para completar los proyectos de tesis y la proporción de proyectos de tesis previamente inacabados. El aporte para el trabajo de investigación fue las contribuciones más novedosas de este estudio proviene de la aplicación de algoritmos de aprendizaje automático que se utilizaron para predecir, razonablemente con precisión, la finalización y no finalización de la tesis.⁷

⁷JALAL, Nouri, LARSSON, Ken. Bachelor Thesis Analytics: Using Machine Learning to Predict Dropout and Identify Performance Factors.2017. [en línea]. [Consulta 20 de septiembre del 2019]. Disponible en: <https://onlinejour.journals.publicknowledgeproject.org/index.php/ijai/article/view/11065/5825>

La Gestión de Ventas según VEGA, Dolly (2015), indica que "La gestión de ventas tiene especialidades en su mayoría centradas en la mercadotecnia. Un proceso eficiente de las ventas se genera en un personal muy capacitado al éxito, y realiza su objetivo de manera económica y con eficiencia"(p.24).⁸

La Gestión de Ventas según CORTAGERENA, A y FREIJIDO (2015), define la gestión de ventas como: "La gestión de ventas tiene que identificar el mercado, además evaluar el segmento en el cuál actuará la empresa, establecer la misión a lograr, evaluar el mercado y a los consumidores"(p.58).⁹ Si una venta se realiza eficiente, se lleva un proceso enumerado los cuales ayuden a tener un contacto y fin eficiente con el cliente.

La Gestión de Ventas Según TORRES, Virgilio (2015) "La gestión de ventas se concentra en tener un negocio con la misión de vender, tener el mejor servicio de calidad para atender las necesidades de los mismos. Se concentra en 6 fases:

La búsqueda de clientes se realiza una búsqueda de aquel cliente el cual debería ser útil el producto o servicio que se le ofrecerá La búsqueda encuentra muchas maneras y formas de hacerse: mediante vía telefónica, eventos y capacitaciones. Además, se realiza un seguimiento del cliente a vender, ver el negocio y que puede necesitar para compararlo con nuestro servicio o producto.

⁸VEGA, Dolly. Gestión estratégica del Dpto. de ventas aplicada en una empresa comercial-farmacéutica.2014. Pp.25-100. [en línea]. Disponible en: http://sisbib.unmsm.edu.pe/bibvirtual/monografias/Ingenie/vega_sd/contenido.htm.

⁹CORTAGERENA, A y FREIJEDO. 2015. "Tecnologías de gestión". Ediciones Mcclti, Buenos Aires. Argentina. P 130 a 138.

Conocimiento de la necesidad del cliente el vendedor debe tener un mayor saber de la empresa y de sus productos o servicios. Se tiene que conocer todas las cualidades, beneficios, y características que producto ofrecerá antes de tener cualquier comunicación con el cliente, debido a que no se puede realizar una presentación eficiente sino se conoce a profundidad.

Prospección del cliente objetivo Teniendo que se conoce las necesidades de la empresa y el producto que se va a ofrecer, por consiguiente, se prospecta al cliente. Este debe saber algunos aspectos y los datos básicos del contacto incluyendo del tomador de decisiones hasta lo que se ofrece y como se le ofrece al cliente. Esto será muy importante, ya que estará informado y se presentará informándole sobre los productos en venta y algún otro requerimiento que el necesite. Esta fase se debe realizar con mucha sutileza y no darle ningún indicio que se tiene información de él, en esta etapa se logra hacer el contacto formal ya sea por llamada telefónica o correo, para solicitar una entrevista formal presentar a la empresa y sus productos.

Presentación de Ventas es la fase el cual se realizar la demostración al cliente de la empresa y de los productos que le pueden agregar valor a su negocio. Para esta presentación el gestor de ventas tiene que estar sumamente preparado adicionalmente haber realizado los pasos anteriores. Deberá contar con un material adicional el cual sirva como procedimiento en la reunión, las características técnicas y los beneficios como también el valor agregado de estos.¹⁰

¹⁰TORRES, Virgilio. Administración de ventas. 2015. [Consulta 16 de setiembre del 2019]. Disponible en <https://editorialpatria.com.mx/pdf/files/9786074384147.pdf>.

Cierre de la Venta en esta última fase donde ha desarrollado correctamente el proceso, no habrá por qué fracasar. Es donde se ayuda a tomar la mejor decisión para él y para su negocio, el gestor de venta tiene que estar muy atento con señales verbales y no verbales que emita el cliente y hacer el cierre definitivo. Basado en un contrato o una orden de pedido. Se tiene que tener presente lo que está buscando es una relación la cual dure con el cliente y que este agradezca con el servicio.

De la dimensión de búsqueda del cliente se obtuvo los siguientes indicadores. La eficiencia para WOJTCZAK A. (2002, p 67) indica que: “La Eficiencia capacidad de desempeñarse bien o lograr un resultado sin desperdiciar energía, recursos, esfuerzo, tiempo o dinero. Se logra una mayor eficiencia cuando se produce la misma cantidad y estándar de servicios por un costo menor, si se sustituye una actividad más útil por una menos útil al mismo costo o si se eliminan las actividades innecesarias.”¹¹ Por lo que la eficiencia se trabajara en Llamadas reales / Total de Llamadas.

La productividad según Zacarías Torres (2014, p 302) menciona que “La productividad es el aprovechamiento eficiente de recursos y se mide por la productividad obtenida y la cantidad de insumos utilizados en esa productividad, sin duda un medio que facilita y hasta cierto grado garantiza tanto la rentabilidad como la competitividad. El insumo es lo que utilizo para poder producir como las horas trabajadas, herramientas, etc”.¹² Por lo que la productividad se trabajara en llamadas reales / horas trabajadas.

¹¹WOJTCZAK, A., Glosario de Términos de Educación Médica,2002. [en línea]. [Consulta 15 de octubre del 2019]. Disponible <http://www.iime.org/glossary.html>

¹²TORRES, Zacarias. Administración estratégica. 2015.[en línea]. [Consulta 08 de octubre del 2019].Disponible en: https://www.academia.edu/19048745/Administraci%C3%B3n_estrat%C3%A9gica_Zacarias_Torres_Hernandez

Machine Learning Según LITSEY, Ryan (2017, p55) Machine Learning –“Son modelos de aprendizaje automáticos, las organizaciones tienen mucho una de las capacidades de predecir continuamente cambios en el negocio para son más capaces de predecir qué sigue. También los datos se van agregando seguido, los modelos de aprendizaje automático generan que la efectividad de la solución sea constantemente actualizada, usando datos más apropiadas y en mayor cambio en el ambiente del aprendizaje automático, teniendo así oportunidades de predecir el futuro.”¹³

Machine Learning según ET AL, Choi. (2018, p.53) dice que “el Machine Learning contiene todo aspecto recurso que tiene un objetivo, los algoritmos y procesos que tiene un aprendizaje continuo tiene intención de integrar datos y casos pasados para vaticinar resultados de situaciones futuras. En su interfaz, el aprendizaje automático son un grupo de técnicas, las cuales son implementadas en interfases de aprendizaje, permitiendo un proceso de minería de datos para la información, descubriendo los patrones y situaciones a partir de datos.”¹⁴

Machine Learning según MCCUE, Colleen (2015, p.40) dice “Machine Learning permite un procesamiento muy rápido de grandes cantidades de imágenes en apoyo de información en tiempo real o casi en tiempo real. Además de las claras implicaciones para una mayor conciencia situacional, la capacidad de aprovechar estas capacidades en apoyo de la detección de cambios puede facilitar la toma de decisiones rápida e informada con respecto a la asignación de recursos.”¹⁵

¹³LITSEY, Ryan. 2017. Recursos en cualquier momento y en cualquier lugar. [Consulta 10 de octubre de 2019]. ISBN 978-0-08-101989-4. Disponible en: <https://www.sciencedirect.com/book/9780081019894/resources-anytime-anywhere>.

¹⁴ ET AT, Choi. Sentiment classification of online reviews to travel destinations by supervised machine learning approaches. 2018. [en línea]. [Consulta 19 de setiembre del 2019] Disponible en <https://iopscience.iop.org/article/10.3847/1538-4365/aa9e8a>.

¹⁵MCCUE, Collen. Minería de datos y análisis predictivo. 2015 [Consulta 29 de marzo del 2020]. ISBN 978-0-12-800229-2. Disponible en la dirección web: <https://www.sciencedirect.com/book/9780128002292/data-mining-and-predictive-analysis?via=ihub=>.

Según HURWITZ, Judith y KIRSCH, Daniel (2018, p60) “La selección está influenciada en gran medida por la aplicación y los datos disponibles. Pero hay algunos algoritmos y aplicaciones de uso común: La regresión lineal y logística no dice que estos algoritmos son utilizados para comprender las relaciones en los datos de estudio. La regresión lineal mayormente es usada para anunciar el valor de una variable dependiente en función del valor de una variable independiente. La regresión logística se puede utilizar cuando la variable dependiente es de naturaleza binaria, A o B. Con la regresión lineal, por ejemplo, las ventas anuales de un vendedor (la variable dependiente) se pueden determinar por su relación con variables independientes como educación o años de experiencia. (las variables independientes).

Árboles de decisión: estos en su mayoría usan la clasificación para hacer futuros pronósticos, enfocados en un conjunto de reglas de decisión. Por ejemplo, apostar a un caballo para ganar, colocar o mostrar podría usar datos sobre el caballo (edad, porcentaje de victorias, pedigrí) y el árbol de decisión aplicaría reglas a esos factores para recomendar una acción o decisión.

Algoritmos basados en Instancia: Un ejemplo bien de un algoritmo basado en la instancia es K-vecinos más cercanos o k- nn. Utiliza la segmentación para simular la probabilidad que existe en un punto de datos, siendo miembro de un grupo u otro en función de su proximidad a otros puntos de datos.

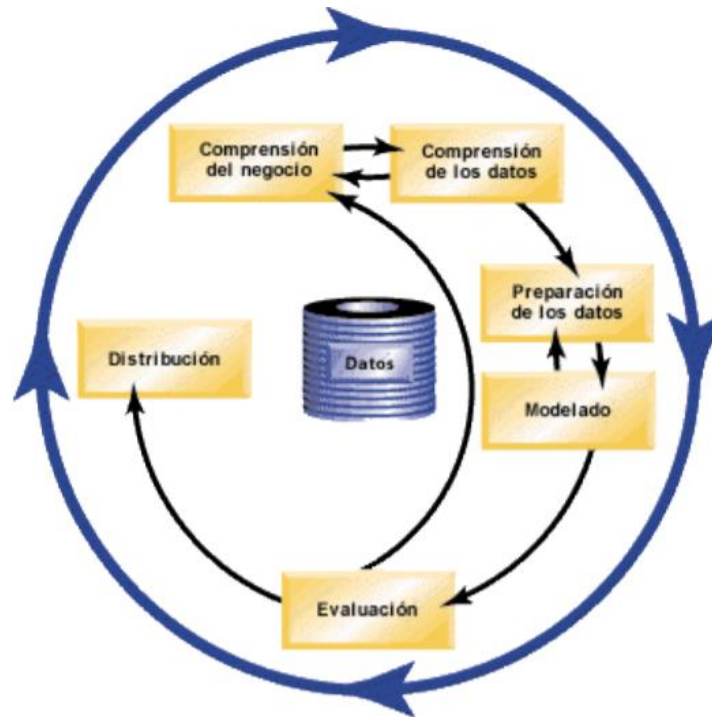
Algoritmos de agrupamiento: Esto se enfocan en los grupos. La acumulación se centra en identificar grupos de registros similares y identificar los registros de acuerdo con el grupo al que pertenecen. Esto se hace sin conocimiento previo sobre los grupos y sus características. Los tipos de algoritmos de agrupación incluyen la agrupación K-means, TwoStep y Kohonen.”¹⁶

¹⁶ HUNWITS, Judith Y KRISH, Daniel. Machine Learning for dummines. 2018.[en línea]. [Consulta 20 de abril del 2020]. ISBN 978-1-119-45495-3. Disponible en: <https://www.ibm.com/downloads/cas/GB8ZMQZ3>.

Según IBM (2012, p.1) indica que “La metodología CRISP DM ha sido una de las mejores metodologías hasta la actualidad, muy útil para la comprensión de datos y para realizar nuevos métodos de trabajo, además define esta metodología en 6 fases: 1. Comprensión del Negocio (Analiza todo los requerimientos y objetivos del negocio, donde evalúa la situación y genera un plan de proyecto para lo que se quiere llegar con la minería de datos). 2. Compresion de los Datos (Entendimiento de los datos para el objetivo del negocio, donde se quiere la recopilación inicial, descripción, exploración y verificación de calidad de datos). 3. Preparacion de los Datos (Prepara los datos para ser trabajos según los requerimientos, donde selecciona los datos, Limpieza, Construcción, Integración y formateo de datos). 4. Modelado (Se aplica las técnicas de modelo elegidas para el proyecto, donde selecciona la técnica de modelado, diseño de la evaluación, construcción del modelo, evaluación del modelado). 5. Evaluacion (Se determina si las fases anteriores fueron útiles para los objetivos del negocio, donde se evalúan los resultados, se revisar el proceso, y se establece los siguientes pasos o acciones). 6. Despliegue (Se explora nuevos modelos para integrarlos en futras decisiones, así como la predicción del despliegue, predicción de la monitorización y al generar el mantenimiento, realización del informe final y revisión del proyecto)”.¹⁷

¹⁷IBM. Manual CRISP-DM de IBM SPSS Modeler. 2012. [en línea]. [Consulta 25 de agosto del 2019]. Disponible en : <ftp://ftp.software.ibm.com/software/analytics/spss/documentation/modeler/15.0/es/CRISP-DM.pdf>.

Figura 3: Metodología CRISP DM



Fuente de IBM (2012)

Tabla 01: Selección de la metodología

Expertos	Puntuación de la Metodología			Metodología Escogida en Base a Puntos
	KDD	CRISP DM	SEMMA	
Galvez Tapia Orleans	21	35	28	CRISP DM
Rivera Crisostomo Renee	16	32	19	CRISP DM
Bermejo Terrones Henry	23	32	19	CRISP DM
Total	60	99	66	

Fuente: Elaboración propia

Se muestra en la Tabla 1, que de las metodologías planteadas para la implementación del proyecto Machine Learning para la gestión de ventas en la Empresa Vértice Empresarial S.A.C la que más se asemeja a los requerimientos para la manipulación de los datos a trabajar, es la metodología CRIPS DM y se ve reflejado en las encuestas que se realizó a los expertos mencionados en la tabla1. CRIPS DM tuvo una aceptación como metodología con un 99% aprobación con respecto a las otras metodologías.

III. METODOLOGÍA

3.1 Tipo y diseño de investigación

Según GREG GUEST, Emily y MITCHELL, Marilyn (2013, p.2) La investigación Aplicada puede, y a menudo lo hace, generar nuevos conocimientos y contribuir a la teoría, pero su objetivo principal es recopilar y generar datos para profundizar nuestra comprensión de los problemas del mundo real. Es a través de esta lente que este libro está escrito, con la intención de proporcionar a los investigadores procedimientos prácticos y herramientas para recopilar y gestionar datos cualitativos de manera rigurosa y transparente.¹⁸

Según HERNANDEZ, Rober, FERNANDEZ Carlos (2014, p141) El diseño preexperimental, “tiene un avance en el cual existe un punto inicial para ver con qué nivel se inició el almacenamiento en las variables dependientes antes del estímulo; es decir, hay un seguimiento del grupo. Actualmente, el diseño no es muy conveniente para fines de establecer algún punto técnico: no hay forma de ver lo datos, ni grupo de igualdad además que actuar varias fuentes de invalidación interna, por ejemplo, la historia. Entre 01 y 02 podrían suscitarse otros acontecimientos los cuales generan cambios, además del tratamiento experimental, y cuanto más largo sea el lapso entre ambas mediciones, mayor será también la posibilidad de que actúen tales fuentes. Por consiguiente, se tiene el riesgo de elegir a un grupo de datos atípico o que en el momento del experimento no se encuentre en su estado normal.”¹⁹

$G \quad 0_1 \quad X \quad 0_2$

¹⁸GREG GUEST, Emily y MITCHELL, Marilyn. Collecting Qualitative Data: A Field Manual for Applied Research.2013. [Consulta 28 de octubre del 2019]. ISBN 978-1-4129-8684-7. Disponible en: <https://books.google.com.pe/books?id=--3rmWYKtloC&printsec=frontcover&dq=applied+research&hl=es&sa=X&ved=0ahUKEwjeh9nm07TpAhXCI-AKHelQAG8Q6AEIXzAF#v=onepage&q=applied%20research&f=false>.

¹⁹HERNANDEZ, Rober, FERNANDEZ, Carlos. Metodología de la investigación. 6ta Edición.2014. [en línea]. [Consulta 15 de abril del 2020]. ISBN 978-1-4562-2396-0. Disponible en: https://periodicooficial.jalisco.gob.mx/sites/periodicooficial.jalisco.gob.mx/files/metodologia_de_la_investigacion_-_roberto_hernandez_sampieri.pdf.

Dónde G: Grupo experimental: Es el grupo de los tipos de servicios de atención de llamadas (muestra) al cual se aplicará la medición para evaluar la eficiencia y productividad para la gestión de ventas. X: Experimento (Machine Learning): Es la aplicación del Machine Learning para la gestión de ventas de la empresa Vértice Empresarial S.A.C., mediante dos evaluaciones (Pre – Test y Post – Test) para medir los cambios para la gestión de ventas. O1: Pre – Test: Medición del grupo experimental antes de la implementación del Machine Learning para la gestión de ventas. Esta medición se compara con la medición del Post – Test. O2: Post – Test: Medición del grupo experimental después de la implementación del Machine Learning para la gestión de ventas. Se comparan ambas mediciones para determinar la eficiencia y productividad en la gestión de ventas del antes y después de la implementación del Machine Learning.

3.2 Variables y Operacionalización

Según HUNWITS, Judith y KRISH, Daniel. (2018, p60) Machine Learning - modelos de aprendizaje y las organizaciones tienen la necesidad de predecir continuamente cambios en el negocio para que son más capaces de predecir qué sigue. Como son constantemente agregando, los modelos de aprendizaje automático los cuales dan una solución sean constantemente actualizada. El aprendizaje automático donde los factores que se han establecido para el data mining, se han hecho más conocidos hoy en día. Cosas como las variedades y volúmenes de los datos disponibles, en procesamiento analógico que es más fácil y más potente, y para que los datos sean accesibles a medida que están en crecimiento.”²⁰

²⁰ HUNWITS, Judith Y KRISH, Daniel. Machine Learning for dummines. 2018.[en línea]. [Consulta 20 de abril del 2020]. ISBN 978-1-119-45495-3. Disponible en: <https://www.ibm.com/downloads/cas/GB8ZMQZ3>.

Según TORRES, Virgilio (2015). define la gestión de ventas como: "La gestión comercial y tiene que identificar el mercado, definir en donde se establecerá en el cuál actuará la empresa, definir una meta y el posicionamiento, investigar el contexto y el comportamiento de los consumidores"(p.58).²¹

Machine Learning trata sobre el aprendizaje automático mediante factores los cuales han realizaron la minería de datos para un análisis más popular en la actualidad. Como por ejemplo la infinidad y variedades de los datos, en el proceso digital que es más accesible y más rápido, y como alberga los datos accesibles en crecimiento.

La gestión comercial necesita evaluar el mercado, definir el segmento en el cual se establecerá la empresa, definir la meta y el posicionamiento, visualizar el mercado y el comportamiento de los clientes. Tiene como dimensiones la búsqueda del cliente, conocimiento del producto o servicio, prospección del cliente, presentación de ventas, cierre de ventas.

De la dimensión de Búsqueda del Cliente se obtuvo dos indicadores:1. Eficiencia: Evaluara la eficiencia de la gestión de llamadas tanto en llamadas efectivas, perdidas y no contestadas. La fórmula para identificar la eficiencia es la siguiente $Llamadas\ reales / Total\ Llamadas$. 2.Productividad: Evalúa la productividad de gestión de llamadas, digamos cuantas llamadas realizo en un tiempo determinado o tiempo de trabajo. La fórmula para identificar la productividad es $Horas\ trabajadas / Llamadas\ Reales$.

²¹TORRES, Virgilio. Administración de ventas. 2015. [Consulta 16 de setiembre del 2019]. Disponible en <https://editorialpatria.com.mx/pdf/files/9786074384147.pdf>.

La escala de medición es nominal ya que consta de varias categorías que son en su mayoría excluyentes. Si solo hubiera dos, sería nominal dicotómica. A cada tipo se le propone en su mayoría un número de Ean sin datos cuantitativo, que facilita su incorporación en la base de datos. En cualquier aspecto, si se usara un código propio, debe tenerse definido lo que significa el Ean para cada variable.

3.3 Población muestra y muestreo

Según HERNANDEZ, Robert y FERNANDEZ Carlos (2014, p.260) dice que “la población a estudiar es un grupo de datos de los cuales tienen relación con cada una de las variables y especificaciones con la finalidad de ver cuál será la muestra. “²²En esta investigación “Machine Learning para la gestión de ventas de la empresa Vértice Empresarial S.A.C. Se considera una población de los 490 tipos de llamadas.

Según HERNANDEZ, Robert y FERNANDEZ Carlos (2014, p.196) la muestra “es un sub grupo de la población que se estudia. En la cual se recolecta la información pertinente, y será representativa con la población.” Se dice en su mayoría un subconjunto de individuos que corresponden a ese grupo determinado en sus principales factores al que llamamos población.” ²³

^{22,23}HERNÁNDEZ, Robert y FERNÁNDEZ Carlos. Metodología de la investigación. 6ta Edición 2014. [en línea]. [Consulta 15 de abril del 2020]. ISBN 978-1-4562-2396-0 Disponible en: https://periodicooficial.jalisco.gob.mx/sites/periodicooficial.jalisco.gob.mx/files/metodologia_de_la_investigacion_-_roberto_hernandez_sampieri.pdf.

Cálculo del para de la muestra:

$$n = \frac{N * Z^2 * p * q}{d^2 * (N - 1) + Z^2 * p * q}$$

Dónde:

N: Población. Z: Nivel de confianza al 95% (1.96). P: Proporción esperada (5%=0.05). I: Precisión (0.05). q: 1-p (0.95).

$$n = \frac{490 * 1.96^2 * 0.05 * 0.95}{0.05^2 * (475 - 1) + 1.96^2 * 0.05 * 0.95} = 216$$

El resultado para el cálculo de la muestra dio un resultado de 216 llamadas por gestión de venta. Esta muestra esta subdividida por 28 días de un mes de muestra estratificada.

Según HERNÁNDEZ, Rober y FERNÁNDEZ, Carlos (2014, p260) el muestreo probabilístico estratifica es “en la que la población se distribuye en segmentos y se toma una muestra para uno de manera proporcional o no proporcional.”²⁴

3.4 Técnicas e Instrumentos de Recolección de Datos

Para la investigación “Machine Learning para la Gestión de venta de la empresa Vértice Empresarial S.A.C”, se utilizará como técnica la entrevista(ver anexos) para ser usado en la muestra y para el instrumento se harán formatos (fichas de Registro), la cual es materia de investigación.

²⁴HERNÁNDEZ, Rober, FERNÁNDEZ Carlos. Metodología de la investigación. 6ta Edicion. 2014.. [en línea]. [Consulta 15 de abril del 2020]. ISBN 978-1-4562-2396-0. Disponible en: https://periodicooficial.jalisco.gob.mx/sites/periodicooficial.jalisco.gob.mx/files/metodologia_de_la_inv_estigacion_-_roberto_hernandez_sampieri.pdf.

Según HERNÁNDEZ, Rober y FERNÁNDEZ Carlos (2014, p229) refiere que la validez “es un grado en que los instrumentos miden la variable que se busca. Se logra cuando se demuestra que el instrumento refleja el concepto a través de sus indicadores empíricos.”²⁵

En la investigación “Machine Learning para la Gestión de Ventas en la Empresa Vértice Empresarial S.A.C”, se usará el juicio de expertos para la aprobación de los instrumentos de medición.

Tabla 2. Validación del indicador Productividad

N	Experto	Grado Académico	Puntaje
1	Rivera Crisóstomo Renee	Magister	88.5%
2	Bermejo Terrones Henry	Magister	92.3%
3	Gálvez Tapia Orleans	Magister	88.5%

Fuente: Elaboración propia

Se mostraron los recursos para que sean evaluados por parte de los expertos para el indicador productividad obteniendo un nivel de aceptación y confianza a los instrumentos siendo los necesarios para obtener los datos de dicho indicador.

²⁵HERNÁNDEZ, Rober y FERNÁNDEZ Carlos. Metodología de la investigación. 6ta Edición. 2014. [en línea]. [Consulta 15 de abril del 2020]. ISBN 978-1-4562-2396-0. Disponible en: https://periodicooficial.jalisco.gob.mx/sites/periodicooficial.jalisco.gob.mx/files/metodologia_de_la_investigacion_-_roberto_hernandez_sampieri.pdf.

Tabla 3. Validación del indicador Eficiencia

	Experto	Grado Académico	Puntaje
1	Rivera Crisóstomo Renee	Magister	90.8%
2	Bermejo Terrones Henry	Magister	91.0%
3	Gálvez Tapia Orleans	Magister	88.5%

Fuente: Elaboración propia

Se mostraron los recursos para que sean evaluados por parte de los expertos para el indicador Eficiencia obteniendo un nivel de aceptación y confianza a los instrumentos siendo los necesarios para obtener los datos de dicho indicador.

Según HERNÁNDEZ, Rober, FERNÁNDEZ Carlos (2014, p229) refiere que confiabilidad o fiabilidad “es un grado en que los instrumentos producen resultados consistentes y coherentes en la muestra o algún caso.”²⁶

Para la confiabilidad en la investigación “Machine Learning para la Gestión de Ventas en la Empresa Vértice Empresarial S.A.C”, se utilizará el SPSS versión 25.

²⁶HERNÁNDEZ, Rober y FERNÁNDEZ Carlos. Metodología de la investigación. 6ta Edición. 2014.[en línea]. [Consulta 15 de abril del 2020]. ISBN 978-1-4562-2396-0. Disponible en: https://periodicooficial.jalisco.gob.mx/sites/periodicooficial.jalisco.gob.mx/files/metodologia_de_la_inv_estigacion_-_roberto_hernandez_sampieri.pdf.

Tabla 4. Confiabilidad del Indicador Eficiencia

Correlaciones			
		Eficiencia_Test	Eficiencia_ReTest
Eficiencia_Test	Correlación de Pearson	1	,734**
	Sig. (bilateral)		,000
	N	28	28
Eficiencia_ReTest	Correlación de Pearson	,734**	1
	Sig. (bilateral)	,000	
	N	28	28

** . La correlación es significativa en el nivel 0,01 (bilateral).

Fuente: Elaboración propia

Debe notarse que la investigación de confiabilidad indicada por el SPSS produce un resultado similar, que muestra una dimensión aceptable de confiabilidad, esto quiere decir que es confiable.

Tabla 5. Confiabilidad del Indicador Productividad

Correlaciones			
		Productividad_Test	Productividad_ReTest
Productividad_Test	Correlación de Pearson	1	,700**
	Sig. (bilateral)		,000
	N	28	28
Productividad_ReTest	Correlación de Pearson	,700**	1
	Sig. (bilateral)	,000	
	N	28	28

** . La correlación es significativa en el nivel 0,01 (bilateral).

Fuente: Elaboración Propia

El SPSS produce un resultado similar, que demuestra una dimensión de confiabilidad aceptable, esto quiere decir que es confiable.

3.5 Procedimientos

Para la toma o recolección de datos primero se tuvo que realizar una carta de aceptación del proyecto de investigación dirigido a la empresa. Después de que el proyecto fue aceptado por la gerencia, recién se pudo realizar los procesos para la toma de algunos datos con cada uno de los implicados en el proceso que se tenía hasta ese momento. Se realizó el fichaje con el número de llamadas que realizaba cada asesor de call center durante el tiempo de un mes, para poder mapear como serian nuestros indicadores y así mismo poder ver si la implementación del proyecto mejoraría la variable independiente en este caso la Gestión de Ventas como se fue planteado desde un inicio.

3.6 Método de análisis de datos

Según VALDERRAMA, Santiago (2015) nos indica “que el propósito de la prueba T student permite comprobar las existencias o diferencias estadísticamente k hay entre grupos con respecto al conjunto de las variables consideradas, asimismo. Esta prueba se considera como una especulación para instancias específicas de variables independientes e incluso es el estadístico para terminaciones multivariante” (p.150).²⁷

Figura 6. Formula Para Hallar el T-Student

$t = \frac{\bar{X} - \mu}{\frac{s}{\sqrt{N}}}$	Donde: <ul style="list-style-type: none">• x es la media muestral,• s es la desviación estándar muestral• N es el tamaño de la muestra.
--	--

Fuente: Valderrama Santiago 2015

²⁷VALDERRAMA, Santiago. Pasos para elaborar Proyectos de Investigación Científica. Cuantitativa, Cualitativa y Mixta. 2015. Lima: San Marcos, 2013. ISBN:978-612-302-878-7

Región de Rechazo

La región de rechazo de $t = t_x$ donde t_x es tal que:

$P [t > t_x] = 0.05$ donde $t_x = \text{Valor Tabular}$

Luego Región de rechazo: $t > t_x$

Promedio:

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$$

Desviación Estándar:

$$S^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n - 1}$$

En la figura se representa en general, un ejemplo definido y gráfico de distribución T – Student, indicado la región de aceptación y posteriormente de rechazo.

Figura N° 7 Distribución según T-Student

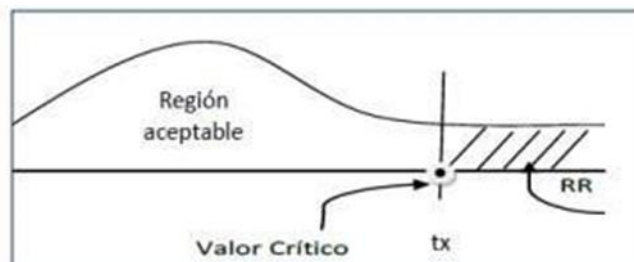


Gráfico de distribución

Fuente: Valderrama Santiago 2015

Utilizando la siguiente figura, donde la significación T se encuentra en la región, se usa la hipótesis alterna siendo de lo contrario, si se encuentra en la región positiva se pasará a utilizar la hipótesis nula.

Existen varios programas para el análisis de datos donde se incluye el análisis de las variables y a su vez explican la matriz de datos, para esta investigación se utilizó el programa SPSS -Paquete Estadísticos de Ciencias sociales, el cual contenía todos los análisis estadísticos necesarios, donde se abre el molde de datos y el investigador selecciona las opciones más adecuadas para el análisis de datos como se hacen con otros tipos de programas.

3.5 Aspectos éticos

Según Gall, Richard (2019).” El aprendizaje automático y la inteligencia artificial están llegando a la madurez. Son tecnologías que ahora, tras décadas incubadas en departamentos de informática y organizaciones de inteligencia militar, se están transformando y teniendo un impacto en una gama realmente impresionante de dominios. Con esta madurez viene más responsabilidad. Surgen preguntas éticas a medida que el aprendizaje automático afecta el cambio en todas partes y se extiende a todo, desde el marketing hasta los sistemas de justicia.”²⁸

Según GUERRA, Ángel (2018).” En España, ante la iniciativa de la Secretaría de Estado de la Sociedad de la Información y la Agenda Digital, se puso en continuidad un ‘comité de sabios’ los cuales están encargados de evaluar las inflexiones jurídicas, sociales y éticas de la Inteligencia Artificial y el Big Data y los distintos aspectos que existen en la sociedad”²⁹

²⁸GALL, Richard. Ética del aprendizaje automático: lo que necesita saber y lo que puede hacer.2019. [en línea]. [consulta: 20 abril 2019]. Disponible en: <https://hub.packtpub.com/machine-learning-ethics-what-you-need-to-know-and-what-you-can-do/>.

²⁹GUERRA, Angel. La Ética de la Data.2018. [en línea]. [consulta: 20 abril 2019]. Disponible en: <https://www.tendencias.kpmg.es/2018/04/etica-big-data/>

IV.RESULTADOS

La investigación la cual fue planteada en 2 fases de acuerdo al diseño de estudio Pre-Experimental. Donde la primera fase constituye con la determinación de la prueba del pre-Test, para luego realizarse la medición de cada indicador antes de poner en marcha el sistema planteado. Por consiguiente, luego de haberse implementado el sistema Machine Learning nuevamente se registró los indicadores de Eficiencia y Productividad.

Los resultados Descriptivos del indicador Eficiencia se visualizan en la siguiente tabla:

Tabla 6. Medidas Descriptivas del nivel de Eficiencia antes y después de la implementación de Machine Learning

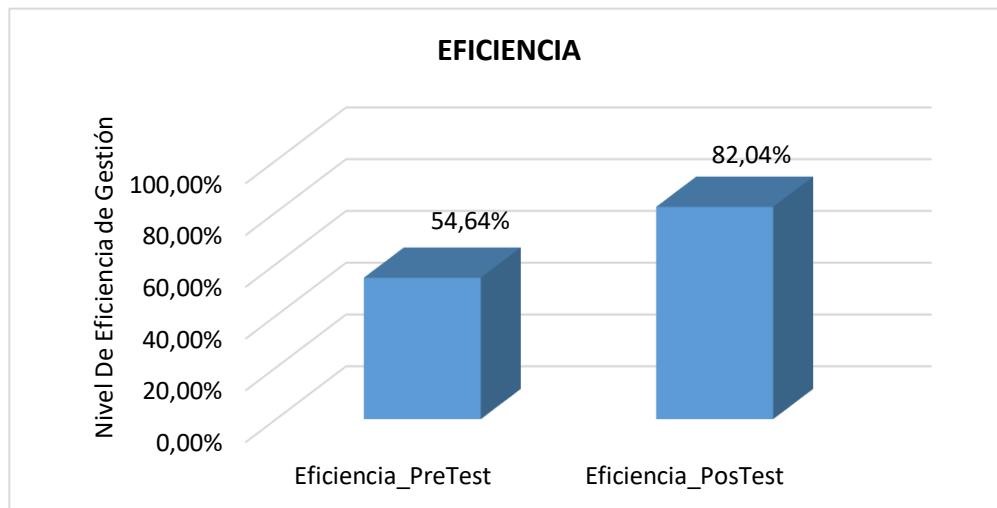
Estadísticos Descriptivos						
	N	Mínimo	Máximo	Media	Desv. Desviación	Varianza
Eficiencia_PreTest	28	0,33	0,80	0,5464	0,1209	0,015
Eficiencia_PostTest	28	0,64	1,00	0,8204	0,09179	0,008
N validos (por Lista)	28					

Fuente: Elaboración propia

Para la Eficiencia en el proceso de Gestión de Ventas donde el pre-test dio como resultado un valor de 55%, tanto que en el post-test fue de 82% tal como se logra ver en la figura; indicando una diferencia antes y después de la implementación de Machine Learning; así mismo, la Eficiencia mínima fue de 0,33 antes, y 0,64 (Tabla N°6 Descriptiva) luego de la implementación del Machine Learning.

En cuanto a la dispersión, en el pre-test se logró una variabilidad de 0,015; sin embargo, en el post-test se tuvo un valor de 0,008.

Figura 8. Índice de Nivel de Eficiencia Antes y Después de la implementación de Machine Learning



Fuente: Elaboración propia

Los resultados Descriptivos del indicador Productividad se visualiza en la siguiente tabla:

Tabla 7. Medidas Descriptivas del nivel de Productividad antes y después de la implementación de Machine Learning

Estadísticos Descriptivos						
	N	Minimo	Maximo	Media	Desv. Desviación	Varianza
Productividad_PreTest	28	0,25	0,75	0,5154	0,10592	0,011
Productividad_PostTest	28	0,56	1,00	0,8107	0,10044	0,010
N validos(por Lista)	28					

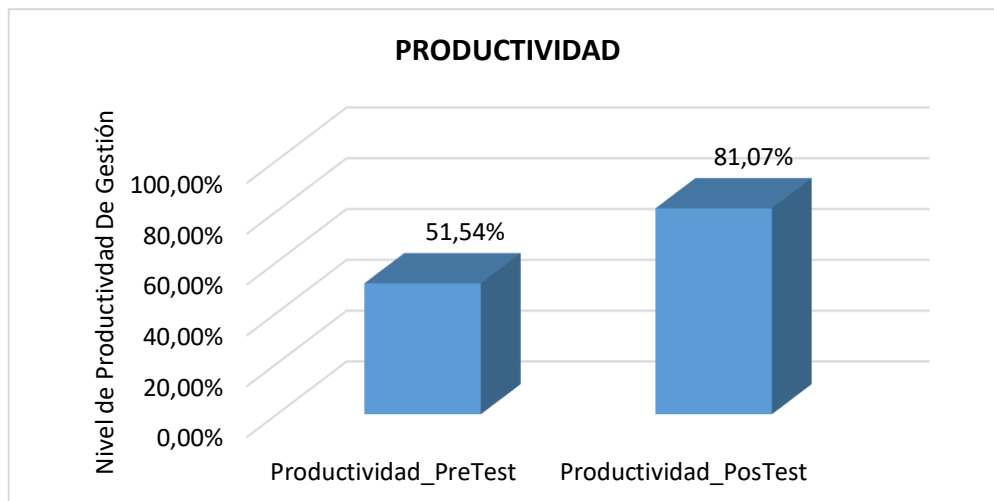
Fuente: Elaboración propia

Para la Productividad en el proceso de Gestión de Ventas en el pre-test dio como resultado un valor de 52%, tanto que en el post-test fue de 81% tal

como se aprecia en la figura; indicando una gran diferencia antes y después de la implementación de Machine Learning; así mismo, la productividad mínima fue de 0,25 antes, y 0,56 (Tabla Nª7 Descriptiva) luego de la implementación del Machine Learning.

En cuanto a la dispersión, en el pre-test se tuvo una variabilidad de 0,011; sin embargo, en el post-test se tuvo un valor de 0,010.

Figura 9. Índice de Nivel de Productividad Antes y Después de la implementación de Machine Learning



Fuente: Elaboración propia

Por Consiguiente, se realizaron las pruebas de normalidad de los indicadores de Eficiencia y Productividad a través del método Shapiro-Wilk, debido a que el tamaño de muestra estratificada está conformado por 28 días y es menor a 50, tal como lo indica Hernández, Fernández y Baptista (2014, p. 376). Esta prueba se realiza ingresando los datos del indicador en el programa estadístico SPSS 25.0, para así tener un nivel de confiabilidad del 95%, mediante las siguientes condiciones:

Si: Sig. < 0.05 adopta una distribución no normal.

Sig. \geq 0.05 adopta una distribución normal.

Dónde:

Sig.: P-valor o nivel crítico del contraste.

Los valores obtenidos fueron los siguientes:

Con la finalidad de seleccionar la prueba de hipótesis; los datos fueron sometidos a la comprobación de su distribución, específicamente si los datos de la Eficiencia contaban con distribución normal.

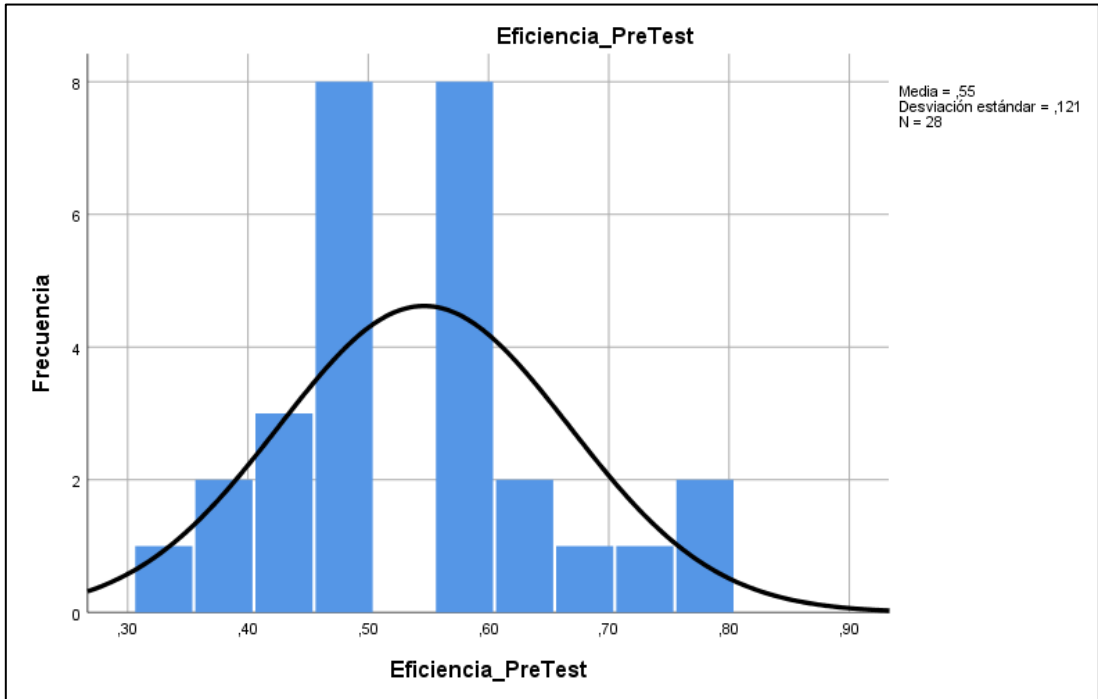
Tabla 8. Prueba Normalidad Eficiencia

Pruebas de Normalidad			
Shapiro - Willk			
	Estadístico	gl	Sig.
Eficiencia_PreTest	0,150	28	0,303
Eficiencia_PostTest	0,127	28	0,483

Fuente: Elaboración propia

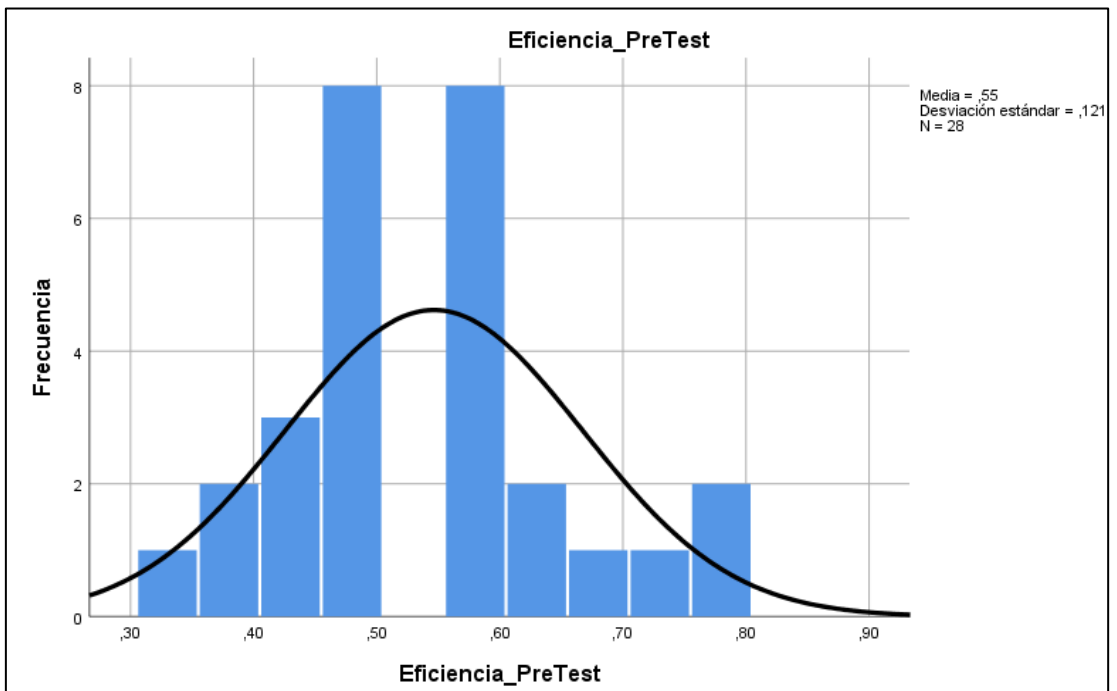
Como puede observar en la Tabla N^o8, los resultados de la prueba indican que el Sig. De Eficiencia generados en el Pre-Test fue de 0.303, valor el cual es mayor que 0.05. Por consiguiente, la eficiencia se distribuye normalmente. Los datos obtenidos de la prueba del Post-Test indican que el Sig. de eficiencia fue de 0.483, cuyo valor es mayor que 0.05, por lo que indica que Eficiencia se distribuye normalmente.

Figura 10. Eficiencia PreTest



Fuente: Elaboración propia

Figura 11. Eficiencia PostTest



Fuente: Elaboración propia

Con la finalidad de seleccionar la prueba de hipótesis; los datos fueron sometidos a la comprobación de su distribución, específicamente si los datos de Productividad contaban con distribución normal.

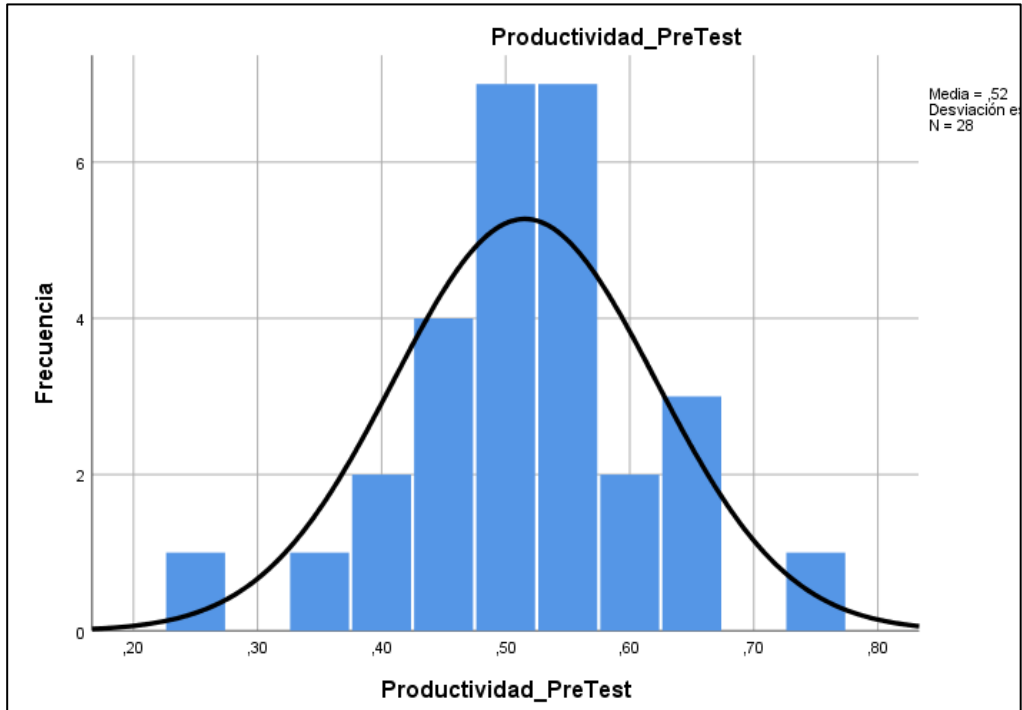
Tabla 9. Prueba Normalidad Productividad

Pruebas de Normalidad			
Shapiro - Willk			
	Estadístico	gl	Sig.
Productividad_PreTest	0,151	28	0,661
Productividad_PostTest	0,13	28	0,223

Fuente: Elaboración propia

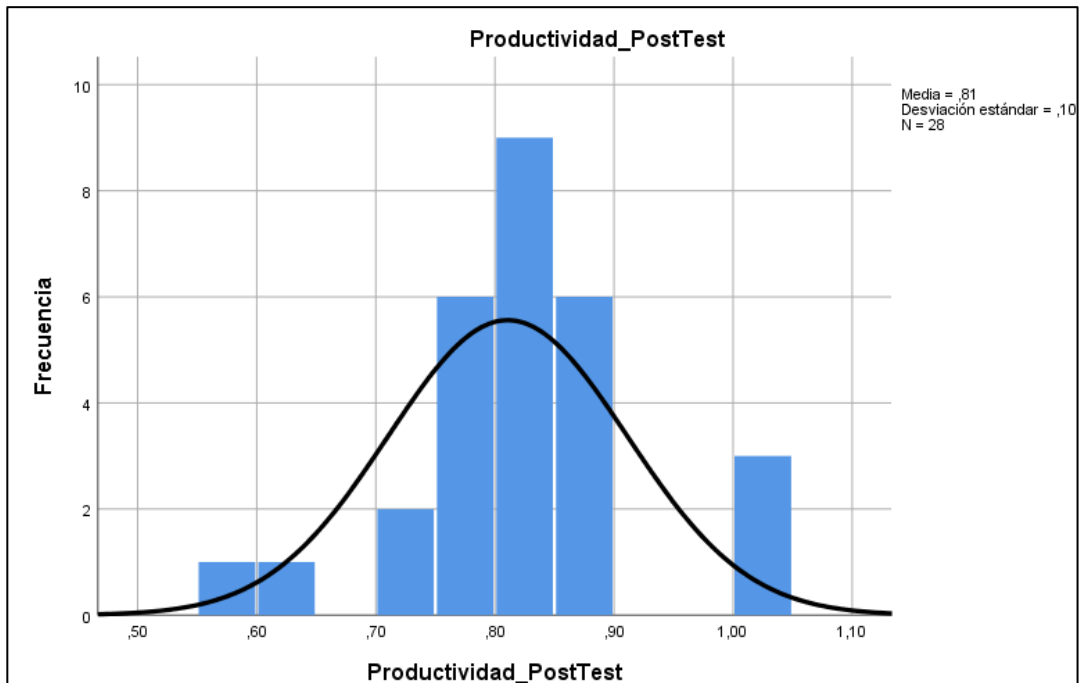
Como se puede observar en la Tabla N^o9, los resultados de la prueba indican que el Sig. De Productividad generados en el Pre-Test fue de 0.661, valor el cual es mayor que 0.05. Por consiguiente, la productividad se distribuye normalmente. Los resultados de la prueba del Post-Test indican que el Sig. De la productividad fue de 0.223, cuyo valor es mayor que 0.05, por lo que indica que Productividad se distribuye normalmente.

Figura 12. Productividad PreTest



Fuente: Elaboración propia

Figura 13. Productividad PostTest



Fuente: Elaboración propia

H1: Machine Learning aumenta la eficiencia para la Gestión de Ventas en la Empresa Vértice Empresarial S.A.C.

indicador: Eficiencia

Hipótesis Estadísticas

Definiciones de Variables:

- **EFa:** Eficiencia antes de utilizar Machine Learning.
- **EFd:** Eficiencia después de utilizar Machine Learning.

H0: Machine Learning no aumenta la eficiencia para la Gestión de Ventas en la Empresa Vértice Empresarial S.A.C.

$$H_0: EF_a \geq EF_d$$

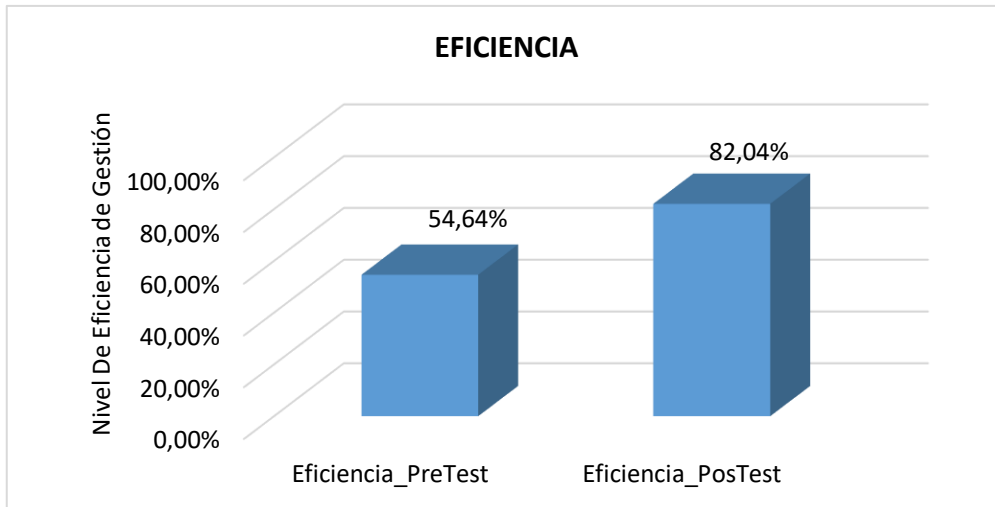
El indicador sin Machine Learning es mejor que el indicador con Machine Learning.

HA: Machine Learning aumenta la eficiencia para la Gestión de Ventas en la Empresa Vértice Empresarial S.A.C.

$$H_a: EF_a < EF_d$$

El indicador sin Machine Learning es mejor que el indicador con Machine Learning.

Figura 14. Nivel De Eficiencia



Fuente: Elaboración propia

Se visualiza en la Figura 14 donde se realiza un aumento en la eficiencia, se puede notar al comparar las mediciones respectivas, que asciende de 54% al valor de 82%.

En cuanto al resultado obtenido del contraste de hipótesis se aplicó la Prueba T-Student, debido a que los datos durante la investigación (Pre-Test y Post-Test) se distribuyen normalmente. Donde el valor de T contraste es de -8.711, el cual es claramente menor que 1.7033.(Ver Tabla Nª7).

Tabla 10. T Student Eficiencia

Media	Prueba De T-Student		
	T	gl	Sig. (Bilateral)
Eficiencia_PreTest	-8,711	27	0,000
Eficiencia_PostTest			

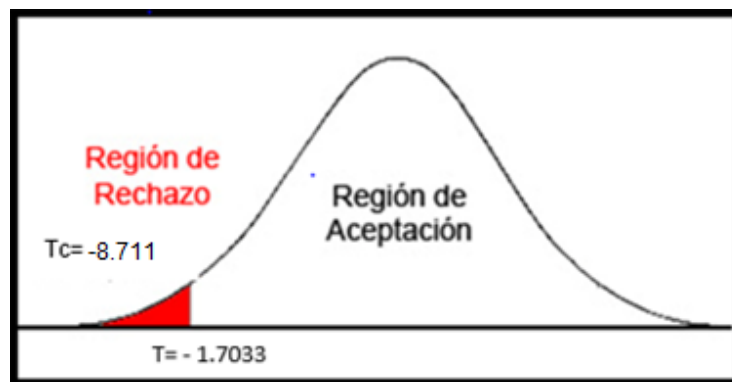
Fuente: Elaboración propia

Aplicando la formula T Student.

$$\frac{x - u}{S / \sqrt{n}}$$
$$T_c = \frac{0,5464 - 0,8204}{0,16641\sqrt{27}}$$
$$T_c = \frac{-0.274}{0,16641 / 5.196}$$
$$T_c = \frac{-0.274}{0.0314}$$
$$T_c = -8.711$$

Por lo tanto, se rechaza la hipótesis nula, aceptando la hipótesis alterna con un 95.00% de confianza. Donde, el valor T obtenido, como se muestra en la Figura 15, ubicándose en la zona de rechazo. Por lo tanto, Machine Learning aumenta la Eficiencia para la gestión de ventas en la empresa Vértice Empresarial S.A.C.

Figura 15. T Student Eficiencia



Fuente: Elaboración propia

Figura 16. Distribución T Student

	0'9995	0'995	0'9875	0'975	0'95	0'875	0'85	0'8	0'75	0'7
1	636'58	63'656	25'452	12'706	6'3137	2'4142	1'9626	1'3764	1'0000	0'7265
2	31'600	9'9250	6'2054	4'3027	2'9200	1'6036	1'3862	1'0607	0'8165	0'6172
3	12'924	5'8408	4'1765	3'1824	2'3534	1'4226	1'2498	0'9785	0'7649	0'5844
4	8'6101	4'6041	3'4954	2'7765	2'1818	1'3444	1'1896	0'9410	0'7407	0'5686
5	6'8685	4'0321	3'1634	2'5706	2'0150	1'3009	1'1558	0'9195	0'7267	0'5594
6	5'9587	3'7074	2'9687	2'4469	1'9432	1'2733	1'1342	0'9057	0'7176	0'5534
7	5'4081	3'4995	2'8412	2'3646	1'8946	1'2543	1'1192	0'8960	0'7111	0'5491
8	5'0414	3'3554	2'7515	2'3060	1'8595	1'2403	1'1081	0'8889	0'7064	0'5459
9	4'7809	3'2498	2'6850	2'2622	1'8331	1'2297	1'0997	0'8834	0'7027	0'5435
10	4'5868	3'1693	2'6338	2'2281	1'8125	1'2213	1'0931	0'8791	0'6998	0'5415
11	4'4369	3'1058	2'5931	2'2010	1'7959	1'2145	1'0877	0'8755	0'6974	0'5399
12	4'3178	3'0545	2'5600	2'1788	1'7823	1'2089	1'0832	0'8726	0'6955	0'5386
13	4'2209	3'0123	2'5326	2'1604	1'7709	1'2041	1'0795	0'8702	0'6938	0'5375
14	4'1403	2'9768	2'5096	2'1448	1'7513	1'2001	1'0763	0'8681	0'6924	0'5366
15	4'0728	2'9467	2'4899	2'1315	1'7531	1'1967	1'0735	0'8662	0'6912	0'5357
16	4'0149	2'9208	2'4729	2'1199	1'7459	1'1937	1'0711	0'8647	0'6901	0'5350
17	3'9651	2'8982	2'4581	2'1098	1'7396	1'1910	1'0690	0'8633	0'6892	0'5344
18	3'9217	2'8784	2'4450	2'1009	1'7341	1'1887	1'0672	0'8620	0'6884	0'5338
19	3'8833	2'8609	2'4334	2'0930	1'7291	1'1866	1'0655	0'8610	0'6876	0'5333
20	3'8496	2'8453	2'4231	2'0860	1'7247	1'1848	1'0640	0'8600	0'6870	0'5329
21	3'8193	2'8314	2'4138	2'0796	1'7207	1'1831	1'0627	0'8591	0'6864	0'5325
22	3'7922	2'8188	2'4055	2'0739	1'7171	1'1815	1'0614	0'8583	0'6858	0'5321
23	3'7676	2'8073	2'3979	2'0687	1'7139	1'1802	1'0603	0'8575	0'6853	0'5317
24	3'7454	2'7970	2'3910	2'0639	1'7109	1'1789	1'0593	0'8569	0'6848	0'5314
25	3'7251	2'7874	2'3846	2'0595	1'7081	1'1777	1'0584	0'8562	0'6844	0'5312
26	3'7067	2'7787	2'3788	2'0555	1'7056	1'1766	1'0575	0'8557	0'6840	0'5309
27	3'6895	2'7707	2'3734	2'0518	1'7033	1'1756	1'0567	0'8551	0'6837	0'5306
28	3'6739	2'7633	2'3685	2'0484	1'7011	1'1747	1'0560	0'8546	0'6834	0'5304
29	3'6595	2'7564	2'3638	2'0452	1'6991	1'1739	1'0553	0'8542	0'6830	0'5302
30	3'6460	2'7500	2'3596	2'0423	1'6973	1'1731	1'0547	0'8538	0'6828	0'5300

Fuente: Elaboración propia

H2: Machine Learning aumenta la Productividad para la gestión de ventas en la empresa Vértice Empresarial S.A.C.

Indicador: Productividad

Hipótesis Estadísticas

Definiciones de Variables:

- **PRa:** Productividad antes de utilizar Machine Learning.
- **PRd:** Productividad después de utilizar Machine Learning.

H0: Machine Learning no aumenta la Productividad para la gestión de ventas en la empresa Vértice Empresarial S.A.C.

$$H0: PRa \geq PRd$$

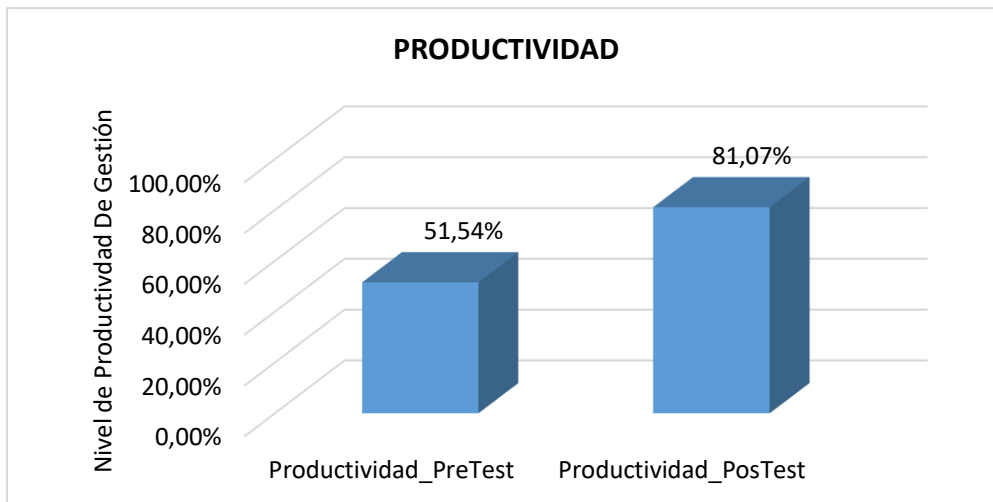
El indicador sin Machine Learning es mejor que el indicador con Machine Learning.

HA: Machine Learning aumenta la Productividad para la gestión de ventas en la empresa Vértice Empresarial S.A.C.

HA: $PRa < PRd$

El indicador sin Machine Learning es mejor que el indicador con Machine Learning.

Figura 17. Nivel De Productividad



Fuente: Elaboración propia

Se visualiza en la Figura 17, donde se realiza un incremento en la productividad, se puede notar al comparar las medias respectivas, que asciende de 52% al valor de 81%.

En cuanto al resultado obtenido del contraste de hipótesis se aplicó la Prueba T-Student, debido a que los datos durante la investigación (Pre-Test y Post-Test) se distribuyen normalmente. Donde el valor de T contraste es de -11,807, el cual es claramente menor que 1.7033.(Ver Tabla N°8).

Tabla 11. T Student Productividad

Media	Prueba De T-Student		
	T	gl	Sig. (Bilateral)
Productividad_PreTest	-11,807	27	0,000
Productividad_PostTest			

Fuente: Elaboración propia

Aplicando la formula T Student.

$$T_c = \frac{x - u}{S / \sqrt{n}}$$

$$T_c = \frac{0,5154 - 0,8107}{0,13237 / \sqrt{27}}$$

$$T_c = \frac{-0.2953}{0,13237 / 5.196}$$

$$T_c = \frac{-0.2953}{0.0254}$$

$$T_c = -11.807$$

Por lo tanto, se rechaza la hipótesis nula, aceptando la hipótesis alterna con un 95.00% de confianza. Donde, el valor T obtenido, como se muestra en la Figura, ubicándose en la zona de rechazo. Por lo tanto, Machine Learning aumenta la Productividad para gestión de ventas en la empresa Vértice Empresarial S.A.C.

Figura 18. T Student Productividad



Fuente: Elaboración propia

Figura 19. Distribución T Student

	0'9995	0'995	0'9875	0'975	0'95	0'875	0'85	0'8	0'75	0'7
1	636'58	63'656	25'452	12'706	6'3137	2'4142	1'9626	1'3764	1'0000	0'7265
2	31'600	9'9250	6'2054	4'3027	2'9200	1'6036	1'3862	1'0607	0'8165	0'6172
3	12'924	5'8408	4'1765	3'1824	2'3534	1'4226	1'2498	0'9785	0'7649	0'5844
4	8'6101	4'6041	3'4954	2'7765	2'1318	1'3444	1'1896	0'9410	0'7407	0'5686
5	6'8685	4'0321	3'1634	2'5706	2'0150	1'3009	1'1558	0'9195	0'7267	0'5594
6	5'9587	3'7074	2'9687	2'4469	1'9432	1'2733	1'1342	0'9057	0'7176	0'5534
7	5'4081	3'4995	2'8412	2'3646	1'8946	1'2543	1'1192	0'8960	0'7111	0'5491
8	5'0414	3'3554	2'7515	2'3060	1'8595	1'2403	1'1081	0'8889	0'7064	0'5459
9	4'7809	3'2498	2'6850	2'2622	1'8331	1'2297	1'0997	0'8834	0'7027	0'5435
10	4'5868	3'1693	2'6338	2'2281	1'8125	1'2213	1'0931	0'8791	0'6998	0'5415
11	4'4369	3'1058	2'5931	2'2010	1'7959	1'2145	1'0877	0'8755	0'6974	0'5399
12	4'3178	3'0545	2'5600	2'1788	1'7823	1'2089	1'0832	0'8726	0'6955	0'5386
13	4'2209	3'0123	2'5326	2'1604	1'7709	1'2041	1'0795	0'8702	0'6938	0'5375
14	4'1403	2'9768	2'5096	2'1448	1'7613	1'2001	1'0763	0'8681	0'6924	0'5366
15	4'0728	2'9467	2'4899	2'1315	1'7531	1'1967	1'0735	0'8662	0'6912	0'5357
16	4'0149	2'9208	2'4729	2'1199	1'7459	1'1937	1'0711	0'8647	0'6901	0'5350
17	3'9651	2'8982	2'4581	2'1098	1'7396	1'1910	1'0690	0'8633	0'6892	0'5344
18	3'9217	2'8784	2'4450	2'1009	1'7341	1'1887	1'0672	0'8620	0'6884	0'5338
19	3'8833	2'8609	2'4334	2'0930	1'7291	1'1866	1'0655	0'8610	0'6876	0'5333
20	3'8496	2'8453	2'4231	2'0860	1'7247	1'1848	1'0640	0'8600	0'6870	0'5329
21	3'8193	2'8314	2'4138	2'0796	1'7207	1'1831	1'0627	0'8591	0'6864	0'5325
22	3'7922	2'8188	2'4055	2'0739	1'7171	1'1815	1'0614	0'8583	0'6858	0'5321
23	3'7676	2'8073	2'3979	2'0687	1'7139	1'1802	1'0603	0'8575	0'6853	0'5317
24	3'7454	2'7970	2'3910	2'0639	1'7109	1'1789	1'0593	0'8569	0'6848	0'5314
25	3'7251	2'7874	2'3846	2'0595	1'7081	1'1777	1'0584	0'8562	0'6844	0'5312
26	3'7067	2'7787	2'3788	2'0555	1'7056	1'1766	1'0575	0'8557	0'6840	0'5309
27	3'6895	2'7707	2'3734	2'0518	1'7033	1'1756	1'0567	0'8551	0'6837	0'5306
28	3'6739	2'7633	2'3685	2'0484	1'7011	1'1747	1'0560	0'8546	0'6834	0'5304
29	3'6595	2'7564	2'3638	2'0452	1'6991	1'1739	1'0553	0'8542	0'6830	0'5302
30	3'6460	2'7500	2'3596	2'0423	1'6973	1'1731	1'0547	0'8538	0'6828	0'5300

Fuente: Elaboración propia

V.DISCUSIÓN

Esta investigación, obtuvo como resultados que Machine Learning aumenta la eficiencia de un 55% a un 82% aumentando en un 27% y en la productividad de un 52% a un 81% aumentando en 29% en un plazo establecido por el proceso de gestión de ventas.

De la misma manera Rodríguez Castillo, Jorge Junior y Miñano Ochoa, Milagros Madeleine, en su investigación Desarrollo De Una Aplicación Informática Basada en un modelo Machine Learning Para Mejorar La Evaluación de Prestamos Crediticios, llego a concluir que el Machine Learning influye formidablemente en el aumento la evaluación de préstamos crediticios en la ciudad de Lima con una eficiencia de 87.17% por lo tanto paso de un 52.05% a 87.17%.

De la misma manera Barroso Salgado, Javier, en su investigación Modelo Predictivo Basado en Machine Learning De Ordenes de Trabajo Riesgosas para Mantenimiento de Equipos Mineros, llego a concluir que el Machine Learning influye formidablemente en el aumento de la productividad de órdenes para mantenimiento de equipos mineros. Estas Ordenes Aumentaron la productividad en un 82,2% por lo tanto se pasó de un 48,20% a 82,2%.

Los resultados que fueron obtenidos mediante la investigación dan conformidad que usar Machine Learning dio resultados positivos en la mejora del proceso de la gestión de ventas, para así poder ampliar la investigación con otras áreas y procesos de alta demanda por la gerencia de la empresa.

VI. CONCLUSIONES

Las conclusiones adquiridas en la presente tesis de investigación fueron las siguientes:

PRIMERO: Se concluye que el Machine Learning aumento el porcentaje de eficiencia a un 82,04% Teniendo inicialmente un 54,64%. Por lo tanto, se afirma que el Machine Learning aumenta la eficiencia para la gestión de ventas en la empresa Vértice Empresarial S.A.C.

SEGUNDO: Se concluye que el Machine Learning aumento el porcentaje de Productividad a un 81,07% Teniendo inicialmente un 51,54%. Por lo tanto, se afirma que el Machine Learning aumenta la Productividad para la gestión de ventas en la empresa Vértice Empresarial S.A.C.

TERCERO: Se concluye que el Machine Learning mejoro el proceso para la gestión de ventas en la empresa Vértice Empresarial S.A.C. Pudiendo así alcanzar los objetivos planteados para esta investigación.

VII. RECOMENDACIONES

Referente a las recomendaciones para la empresa, se recomienda lo siguiente:

- Renovar el equipo de trabajo (computadoras), por equipo más actual y así se pueda aprovechar más el funcionamiento de los programas instalados para Machine Learning.
- Modificar el proceso de Gestión de ventas el cual fue desarrollado como parte de esta investigación.
- Capacitación constante al personal del área Gestión para una óptima realización del proceso.

Recomendaciones para la competencia con otras entidades desarrolladoras de software que abarcan tecnológicamente los procesos de las Pymes o grandes empresas se recomienda lo siguiente:

- Revisar detalladamente los procesos en los cuales se va a trabajar o desarrollar la investigación dentro de la entidad o empresa.
- De acuerdo a los datos obtenidos, establecer la metodología que más se adapte a las necesidades de los procesos a mejorar.
- Consolidar un equipo de trabajo responsable y comunicativo para el desarrollo y cumplimiento de las tareas o roles asignados en el proyecto.
- Analizar el impacto de una mejora para el algoritmo de optimización en la fase de construcción para obtener una mejor eficacia del modelo de Machine Learning.
- Evaluar y Agregar nuevas características de acuerdo a los requerimientos de la empresa en donde se implementó Machine Learning.

REFERENCIAS

CORTAGERENA, A Y FREÍJEDO, C. 2015. "Tecnologías de gestión". Ediciones Mcclti, Buenos Aires. Argentina. P 130 a 138.

ET AT, Choi. Sentiment classification of online reviews to travel destinations by supervised machine learning approaches.2018. [en línea]. [Consulta 19 de setiembre del 2019]. Disponible en <https://iopscience.iop.org/article/10.3847/1538-4365/aa9e8a>.

GALL, Richar. Ética del aprendizaje automático: lo que necesita saber y lo que puede hacer.2019. [en línea]. [consulta: 20 abril 2019]. Disponible en: <https://hub.packtpub.com/machine-learning-ethics-what-you-need-to-know-and-what-you-can-do/>

GUERRA, Angel. La Ética de la Data.2018. [en línea]. [consulta: 20 abril 2019]. Disponible en: <https://www.tendencias.kpmg.es/2018/04/etica-big-data/>

GREG GUEST, Emily y MITCHELL, Marilyn. Collecting Qualitative Data: A Field Manual for Applied Research.2013. [Consulta 28 de octubre del 2019]. ISBN 978-1-4129-8684-7. Disponible en: <https://books.google.com.pe/books?id=--3rmWYKtloC&printsec=frontcover&dq=applied+research&hl=es&sa=X&ved=0ahUKEwjeh9nm07TpAhXCI-AKHelQAG8Q6AEIXzAF#v=onepage&q=applied%20research&f=false>.

HERNÁNDEZ, Rober y FERNÁNDEZ Carlos. Metodología de la investigación. 6ta Edición.2014. [en línea]. [Consulta 15 de abril del 2020]. ISBN 978-1-4562-2396-0. Disponible en: https://periodicooficial.jalisco.gob.mx/sites/periodicooficial.jalisco.gob.mx/files/metodologia_de_la_investigacion_-_roberto_hernandez_sampieri.pdf.

IBM. Manual CRISP-DM de IBM SPSS Modeler .2012. [en línea]. [Consulta 28 de octubre del 2019]. Disponible en : <ftp://ftp.software.ibm.com/software/analytics/spss/documentation/modeler/15.0/es/CRISP-DM.pdf>.

HUNWITS, Judith y KRISH, Daniel. Machine Learning for dummines. 2018.[en línea]. [Consulta 20 de abril del 2020]. ISBN 978-1-119-45495-3. Disponible en: <https://www.ibm.com/downloads/cas/GB8ZMQZ3>.

LITSEY, Ryan. 2017.Recursos en cualquier momento y en cualquier lugar. [Consulta 10 de octubre de 2019]. ISBN 978-0-08-101989-4. Disponible en: <https://www.sciencedirect.com/book/9780081019894/resources-anytime-anywhere>.

MCCUE, Collen. Minería de datos y análisis predictivo.2015[Consulta 29 de marzo del 2020]. ISBN 978-0-12-800229-2 Disponible en la dirección web: <https://www.sciencedirect.com/book/9780128002292/data-mining-and-predictive-analysis?via=ihub=>.

MEDRANO PARADO, Sandra. Modelo de minería de datos usando Machine Learning con reconocimiento de patrones de síntomas y enfermedades respiratorias en las historias clínicas para mejorar el diagnóstico de pacientes en la ciudad de Trujillo 2016.2016 [Consulta 12 de setiembre del 2019]. Disponible en: <http://rpositorio.ucv.edu.pe/handle/UCV/9852>.

MENDOZA RIOFRÍO, Marcela. Avances hacia la Industria 4.0 aún son tímidos.2019. [Consulta 12 de abril del 2020]. Disponible en <https://elcomercio.pe/economia/dia-1/avances-hacia-la-industria-40-aun-son-timidos-noticia/>.

MONTIEL LOPEZ, Jacob. Fast and Slow Machine Learning.2019. [En Línea]. [Consulta 16 de abril del 2020]. Disponible en <https://pastel.archives-ouvertes.fr/tel-02098633/document>.

NINA ASTO, Christian y VILCA ALPARTIDA, Jean. Búsqueda de Patrones de comportamiento usando Machine Learning, para la toma de decisiones generales en la empresa Chuchuhuasi.2018. [Consulta 10 de abril del 2020]. Disponible en <http://repositorio.uandina.edu.pe/handle/UAC/2795>.

NOURI, Jalal y LARSSON, Ken. Bachelor Thesis Analytics: Using Machine Learning to Predict Dropout and Identify Performance Factors.2017. [en línea]. [Consulta 20 de septiembre del 2019]. Disponible en: <https://onlinejour.journals.publicknowledgeproject.org/index.php/ijai/article/view/11065/5825>.

RODRÍGUEZ CASTILLO, Jorge. Desarrollo de una Aplicación informática basada en un modelo Machine Learning para mejorar la evaluación de préstamos crediticios. 2017.[Consulta 26 de setiembre del 2019]. Disponible en la página de la universidad en: <https://repositorio.upn.edu.pe/bitstream/handle/11537/12294/Rodr%C3%ADguez%20Castillo%2C%20Jorge%20Junior%20%20Mi%20%20Mi%C3%B1ano%20Ochoa%2C%20Milagros%20Madeleine.pdf?sequence=1&isAllowed=y>.

TORRES, Virgilio. Administración de ventas. 2015. [Consulta 16 de setiembre del 2019]. Disponible en: <https://editorialpatria.com.mx/pdffiles/9786074384147.pdf>.

TORRES, Zacarias. Administración estratégica. 2015.[en línea]. [Consulta 08 de octubre del 2019]. Disponible en: https://www.academia.edu/19048745/Administraci%C3%B3n_estrat%C3%A9gica_Zacarias_Torres_Hernandez.

VEGA SÁNCHEZ, Dolly. Gestión estratégica del Dpto. de ventas aplicada en una empresa comercial-farmacéutica.2014. Pp.25-100. [en línea]. Disponible en: http://sisbib.unmsm.edu.pe/bibvirtual/monografias/Ingenie/vega_sd/contenido.htm.

VOJT, Jan. Deep neural networks and their implementation.2016. Pp 45-250. [En Línea]. [Consulta 16 de abril del 2020]. Disponible en: <https://is.cuni.cz/webapps/zzp/download/120229251>.

WOJTCZAK, A. Glosario de Términos de Educación Médica,2002. [en línea]. [Consulta 15 de octubre del 2019]. Disponible <http://www.iime.org/glossary.html>.

ANEXOS

**Aplicación de la
Metodología CRISP-DM
Al Problema**

1. Comprensión del Negocio

A continuación, iremos siguiendo cada una de las tareas de las que consta esta primera fase en el proceso de la minería de datos, cuya finalidad es determinar los objetivos y requisitos del proyecto desde una perspectiva de negocio, para más adelante poder convertirlos en objetivos desde el punto de vista técnico y en un plan de proyecto.

1.1. Determinar los Objetivos del Negocio

El objetivo de machine learning que se va a aplicar en este proyecto es el de hacer predicciones lo más fiables posible a partir de los datos de los que ya se disponen de las llamadas telefónicas y los números de los clientes ya contactados

Contexto

En referencia a la situación de negocio en la organización al principio de este proyecto se puede decir que se cuenta con una base de datos de clientes que actualmente fueron contactados y se tuvo una gestión Sin embargo no se tiene los procesos sistematizados para tener una gestión más rápida y eficiente.

Objetivos del Negocio

Los objetivos del negocio como ya se ha mencionado son la predicción de llamadas por asesores de call de tal manera que se pueda hacer una estimación fiable partiendo de los datos que ya tenemos de dichos clientes contactados anteriormente. Se podrían hacer muchas predicciones según las necesidades de la empresa en cada momento, pero en este proyecto se han definido los siguientes objetivos:

- Hacer predicciones acerca del contacto eficiente de los asesores de call que hacen en las llamadas en un tiempo determinado.

- Validar la productividad del asesor de call que hacen en las llamadas en un tiempo determinado.

Estos informes pueden ser muy útiles para los coordinadores a la hora de validación de producción y eficiencia, así como validar en que horarios se tuvo mayor afluencia de llamadas tiempos y todo. Todo esto permitirá a la empresa mejorar la calidad de los servicios ofrecidos a los sus clientes.

Criterios de éxito del negocio

Desde el punto de vista del negocio se establece como criterio de éxito la posibilidad de realizar predicciones sobre futuras llamadas y las que se realizaron durante un tiempo determinado. Sobre el contacto con nuevos clientes con una elevada fiabilidad, así como también la posibilidad de identificar en días durante el mes se tiene mayor eficiencia en las llamadas realizadas para así poder enfocarse en la productividad de las llamadas.

1.2. Valuación de la Situación

Se cuenta con información de gestiones eficiente de anteriores clientes, que son guardados en archivos de Excel y una base de datos de los números telefónicos y a la persona que es la titular de esta entre otros datos personales del cliente que puede ser útiles al realizar el machine learning.

Inventario de recursos

En cuanto a recursos de software disponemos del programa Sql Server que proporciona herramientas para realizar tareas de minería de datos sobre una base de datos que se creara que es con la que contamos para el almacenamiento de los datos.

Los recursos de hardware de los que disponemos son un ordenador de escritorio de las siguientes características:

- Marca: DELL
- Modelo: Dell Optiplex 9020
- Procesador: Intel © Core i7 4ta Gn 3.00 GHz
- Memoria RAM: 8,00 GB
- Capacidad de almacenamiento: HDD 1TB
- Tarjeta gráfica: Video integrado Intel HD Graphics 4GB
- Sistema operativo: Microsoft Windows 10 Professional ©
- Monitor TFT: Philips © 190 S de 19”

1.3. Determinar los Objetivos del proyecto

Los objetivos en términos de minería de datos son:

- Predecir la eficiencia del asesor en las llamadas durante el día de trabajo.
- Verificar la productividad del asesor en las llamadas durante el día de trabajo.
- Sistematizar los reportes que son enviados los coordinadores o gerencia.

Desde el punto de vista de la minería de datos se establece como criterio de éxito la posibilidad de realizar predicciones sobre nuevos clientes contactados eficientemente con un elevado porcentaje de fiabilidad. El grado de fiabilidad lo determinará el algoritmo específico que se emplee a la hora de conseguir el modelo de la minería de datos.

1.4. Realizar el Plan del Proyecto

El proyecto se planificará en fases para así facilitar su y estimar su tiempo

- **Fase 1:** Análisis de la estructura de los datos y la información de la base de datos.
 - Tiempo estimado: 2 semanas.
- **Fase 2:** Ejecución de consultas para tener muestras representativas de los datos.
 - Tiempo estimado: 1 semana.
- **Fase 3:** Preparación de los datos (selección, limpieza, conversión y formateo, si Fuera necesario) para facilitar la minería de datos sobre ellos
 - Tiempo estimado: 1 semana.
- **Fase 4:** Elección de las técnicas de modelado y ejecución de las mismas sobre los datos.
 - Tiempo estimado: 1 semana.
- **Fase 5:** Análisis de los resultados obtenidos en la etapa anterior, si fuera necesario repetir la etapa 4.
 - Tiempo estimado: 1 semana.
- **Fase 6:** Producción de informes con los resultados obtenidos en función de los objetivos de negocio y los criterios de éxito establecidos.
 - Tiempo estimado: 1 semana.
- **Fase 7:** Presentación de los resultados finales.
 - Tiempo estimado: 1 semana.

2. Comprensión de los Datos

En esta fase realiza la recolección inicial de los datos para poder establecer un primer contacto con el problema, familiarizarse con los datos y averiguar su calidad.

2.1. Recolección de los datos

A continuación, listamos los datos que usaremos a la largo del proyecto para su desarrollo.

ASESOR:

- ✓ ID ASESOR
- ✓ ASESOR

GESTIONES:

- ✓ IDGESTIONES
- ✓ IDASESOR
- ✓ IDSEDE
- ✓ IDLIDER_CARTERA
- ✓ IDCOORDINADOR
- ✓ IDJEFE_OPERACIONES
- ✓ IDSERVICIO
- ✓ IDAREA
- ✓ IDCUENTA
- ✓ IDCARGO
- ✓ TIEMPO ATENTIDAS
- ✓ TIEMPO DE CONVERSACION
- ✓ TIEMPO ESPERA
- ✓ TIEMPO DESPUES
- ✓ TIEMPO LOGUEO
- ✓ TIEMPO INICIO
- ✓ TIEMPO NO INICIO
- ✓ TIEMPO
- ✓ TIEMPO PROGRAMADO
- ✓ TIEMPO AUSENTE

SERVICIO:

- ✓ IDSERVICIO
- ✓ SERVICIO

COORDINADOR:

- ✓ IDCOORDINADOR
- ✓ COORDINADOR

SEDE:

- ✓ IDSEDE
- ✓ SEDE

ÁREA:

- ✓ IDÁREA
- ✓ ÁREA

JEFE OPERACIONES:

- ✓ IDJEFE OPERACIONES
- ✓ JEFE OPERACIONES

CUENTA:

- ✓ IDCUENTA
- ✓ CUENTA

CARGO:

- ✓ IDCARGO
- ✓ CARGO

LIDER CARTERA:

- ✓ IDLIDER CARTERA
- ✓ LIDER CARTERA

Las tablas de los cuales se recoge la información necesaria para la implementación del proyecto son:

2.2. Descripción de los datos

Los datos se encuentran almacenados en un almacén de datos con esquema dimensional en estrella. El cual es un esquema relacional de esta base de datos, para generar esta figura se ha utilizado la herramienta proporcionada por SQL server que sirve para generar modelos de las bases de datos Oracle.

Tabla_Asesor: En esta tabla se registran todos los asesores que ingresaran a la plataforma. Tiene como llave primaria: IDAsesor.

- **IDAsesor:** Tipo numérico. Este campo es un número que identifica a cada asesor que es único para cada uno.

- **Asesor:** Tipo alfanumérico Este campo representa la identificación a cada asesor y que es único para cada una.

Tabla_Servicio: En esta tabla se registran todos los servicios a brindar o existentes en ese momento. Tiene como llave primaria: IDServicio.

- **IDServicio:** Tipo numérico. Este campo es un número que identifica a cada servicio que es único para cada uno.
- **Servicio:** Tipo alfanumérico Este campo representa el tipo de servicio brindado que es único para cada uno.

Tabla_Coordinador: En esta tabla se registran todos los coordinadores existentes de cada cartera. Tiene como llave primaria: IDCoordinador.

- **IDCoordinador:** Tipo numérico. Este campo es un número que identifica a cada coordinador de cartera que es único para cada uno.
- **Coordinador:** Tipo alfanumérico Este campo representa a cada coordinador que es único para cada uno.

Tabla_Sede: En esta tabla se registran todos los coordinadores existentes de cada cartera. Tiene como llave primaria: IDSede.

- **IDSede:** Tipo numérico. Este campo es un número que identifica a cada sede y que es único para cada uno.
- **Sede:** Tipo alfanumérico Este campo representa a cada sede y que es único para cada uno.

Tabla_Área: En esta tabla se registran todos los coordinadores existentes de cada cartera. Tiene como llave primaria: IDÁrea.

- **IDÁrea:** Tipo numérico. Este campo es un número que identifica a cada área dentro de la empresa y que es único para cada uno.
- **Área:** Tipo alfanumérico Este campo representa a cada área que es único para cada una.

Tabla_Jefe_Operaciones: En esta tabla se registran todos los jefes de operaciones existentes. Tiene como llave primaria: IDJefe_Operaciones.

- **IDJefe_Operaciones:** Tipo numérico. Este campo es un número que identifica a el jefe de operaciones dentro de la empresa y que es único para cada uno.
- **Jefe_Operaciones:** Tipo alfanumérico Este campo representa a cada al jefe de operaciones y que es único para cada una.

Tabla_Lider_Cartera: En esta tabla se registran todos los líderes de una cartera en su mayoría los que están a cargo de los asesores y bajo la supervisión de los coordinadores de operaciones existentes. Tiene como llave primaria: IDLider_Cartera.

- **IDLider_Cartera:** Tipo numérico. Este campo es un número que identifica a los líderes de cartera dentro de la empresa y que es único para cada uno.
- **Lider_Cartera:** Tipo alfanumérico Este campo representa a cada el nombre de líder de cartera y que es único para cada una.

Tabla_Cargo: En esta tabla se registran todos los cargos existentes. Tiene como llave primaria: IDCargo.

- **IDCargo:** Tipo numérico. Este campo es un número que los cargos dentro de la empresa con valores únicos.

- **Cargo:** Tipo alfanumérico Este campo representa a cada cargo dentro de la empresa y que es único para cada una.

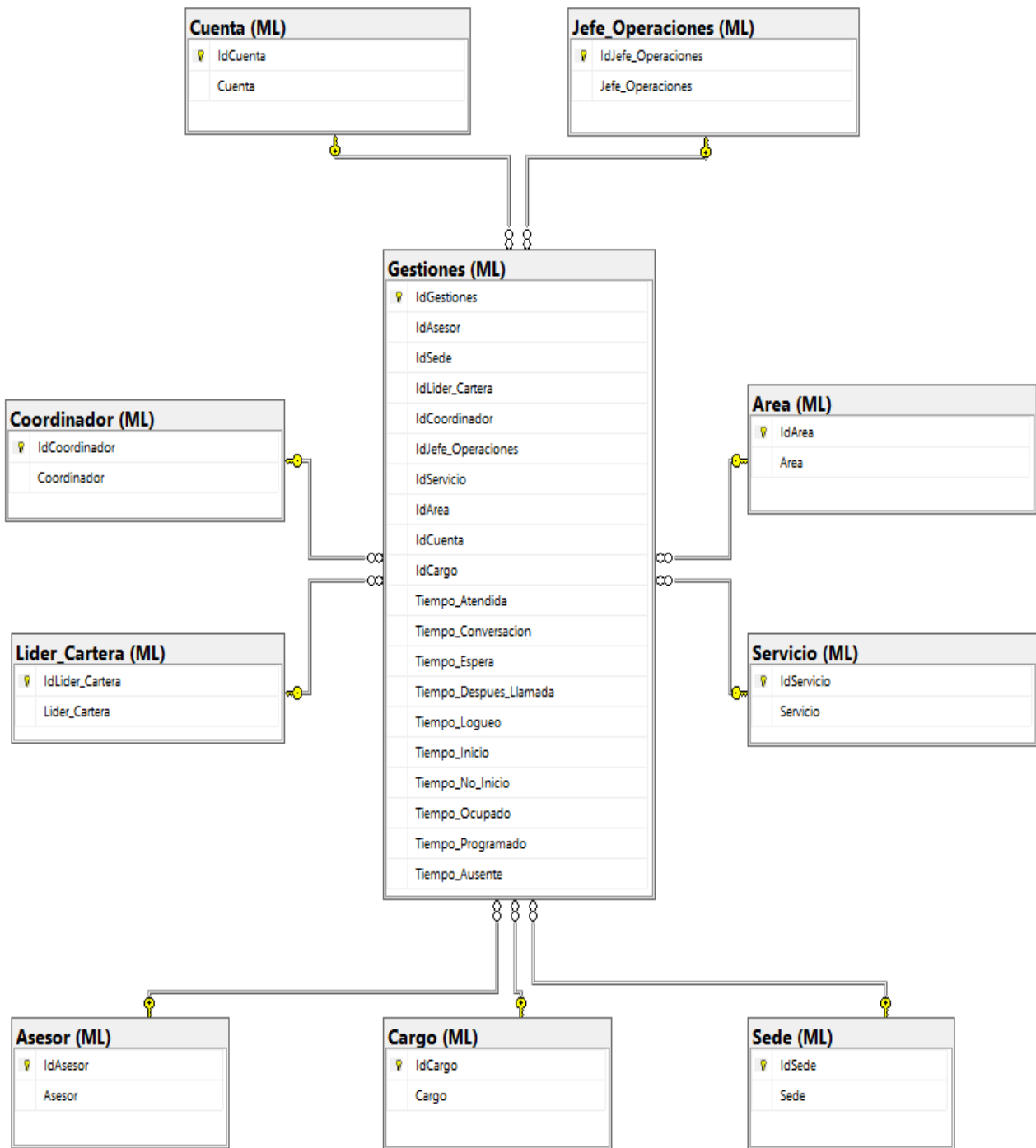
Tabla_Cuenta: En esta tabla se registran todas las cuentas existentes. Tiene como llave primaria: IDCuenta.

- **IDCuenta:** Tipo numérico. Este campo es un número que se asigna a las cuentas dentro de la empresa con valores únicos.
- **Cuenta:** Tipo alfanumérico Este campo representa a cada cuenta dentro de la empresa y que es único para cada una.

Tabla_Gestiones: Esta tabla es la tabla central del almacén de datos, también gestiones “tabla de hechos”, ya que es en esta tabla en la que se registra toda la información de las llamadas realizadas. Esta tabla al ser la tabla central tiene como clave primaria una combinación de cada una de las claves principales del resto de las tablas, llamadas tablas dimensionales, estas claves son: IDGestiones, IDAsesor, IDSede, IDLider_cartera, IDCoordinador, IDJefe_Operaciones, IDServicio, IDArea, IDCuenta e IDCargo. Estas claves son a la vez claves foráneas (foreign keys). Esta tabla tiene la información de las llamadas realizadas y los campos de cada registro son:

- **IDGestiones:** Tipo numérico. Este campo es un número que identifica a cada llamada y que es único para cada una.
- **IDAsesor:** Tipo numérico. Este campo es un número que identifica a cada empleado registrado o conectado y que es único para cada una.
- **IDSede:** Tipo numérico. Este campo es un número que identifica a cada sede o lugar de y que es único para cada una.
- **IDLider_Cartera:** Tipo numérico. Este campo es un número que identifica a cada Lider de Cartera y que es único para cada una.
- **IDCoordinador:** Tipo numérico. Este campo es un número que identifica a cada coordinador por servicio y que es único para cada una.

- **IDJefe_Operaciones:** Tipo numérico. Este campo es un número que identifica a cada jefe de área y que es único para cada una.
- **IDServicio:** Tipo numérico. Este campo es un número que identifica a cada servicio brindado o disponible y que es único para cada una.
- **IDArea:** Tipo numérico. Este campo es un número que identifica la Área donde se realiza la llamada telefónica y que es único para cada una.
- **IDCuenta:** Tipo numérico. Este campo es un número que identifica a cada cuenta registrada dentro de la plataforma y que es único para cada una.
- **IDCargo:** Tipo numérico. Este campo es un número que identifica a cada cargo y que es único para cada una.
- **Tiempo_Atendidas:** Tipo time que guarda la información de las llamadas que fueron atendidas correctamente.
- **Tiempo_Conversacion:** Tipo time que guarda la información de la duración de las llamadas que fueron atendidas correctamente.
- **Tiempo_Espera:** Tipo time que guarda la información de la duración de espera de las llamadas.
- **Tiempo_Despues_Llamada:** Tipo time que guarda la información del tiempo después de la llamada.
- **Tiempo_Logueo:** Tipo time que guarda la información del último logueo del asesor.
- **Tiempo_Inicio:** Tipo time que guarda la información del inicio de sesión del asesor.
- **Tiempo_No_Inicio:** Tipo time que guarda la información del no inicio de sesión del asesor.
- **Tiempo_Ocupado:** Tipo time que guarda la información del tiempo en que el asesor se mantuvo ocupado en una llamada.
- **Tiempo_Programado:** Tipo time que guarda la información de las llamadas que fueron programadas correctamente.
- **Tiempo_Ausente:** Tipo time que guarda la información del tiempo en el que el asesor no estuvo conectado o realizando las llamadas correctamente.



2.3. Exploración de los Datos

Una vez que se han descrito los datos, se procede a explorarlos, esto implica aplicar pruebas estadísticas básicas que revelarán propiedades de los datos, y crear tablas de frecuencia y gráficos de distribución de los datos. Este

informe sirve principalmente para determinar la consistencia y completitud de los datos.

" Acceso a la base de datos"

```
server = "DESKTOP-8BSFAVP"
```

```
database = "Vertice_Empresarial"
```

```
user = "sa"
```

```
password = "miguel147258"
```

```
select = 'SELECT * FROM Setting.LoadData Where Id = 0'
```

```
cnx = Connection_String(server, database, user, password)
```

" Conexión a la Base de Datos"

```
table_configuration = settingloaddata(select, cnx)
```

```
pd.set_option('display.expand_frame_repr', False)
```

" Visualización de columnas"

```
query_consolidado.columns
```

```
query_consolidado.shape
```

3. Preparación de los Datos

En esta fase de la metodología se trata de preparar los datos para adecuarlos a las técnicas de minería de datos que se van a emplear sobre ellos. Esto

implica seleccionar el subconjunto de datos que se va a utilizar, limpiarlos para mejorar su calidad, añadir nuevos datos a partir de los existentes y darles el formato requerido por la herramienta de modelado.

“Verificación del tipo de dato de las columnas de la tabla consolidado”

```
query_consolidado.dtypes
```

“Verificación del tipo de datos entero de las columnas de la tabla consolidado”

```
query_consolidado.dtypes == float
```

“Verificación del tipo de datos numérico de las columnas de la tabla consolidado”

```
(query_consolidado.dtypes == float)|(query_consolidado.dtypes == int)
```

“Verificación del tipo de datos texto de las columnas de la tabla consolidado”

```
query_consolidado.dtypes == object
```

“Verificación de llamadas atendidas mediante un gráfico de barras”

```
features = ['Atendidas']
```

```
query_consolidado[features].hist(figsize=(8, 8));
```

“Verificación de llamadas atendidas mediante un diagrama de desing”

```
query_consolidado[features].plot(kind='density', subplots=True, layout=(1, 2), sharex=False, figsize=(20, 8));
```

“Verificación de llamadas atendidas mediante un diagrama de boxplot”

```
sb.boxplot(x='Atendidas', data=query_consolidado);
```

“Verificación de llamadas atendidas mediante un diagrama de tabla cruzada”

```
pd.crosstab(query_consolidado['Atendidas'],  
query_consolidado['Atendidas'], margins=True)
```

4. Modelado

En esta fase de la metodología se escogerá la técnica (o técnicas) más apropiadas para los objetivos marcados. A continuación, y una vez realizado un plan de prueba para los modelos escogidos, se procederá a aplicar dichas técnicas sobre los datos para generar el modelo y por último se tendrá que evaluar si dicho modelo ha cumplido los criterios de éxito o no.

4.1. Regresión Lineal

```
from sklearn.linear_model import LinearRegression
from sklearn.metrics import mean_squared_error
while True:
    y = query_consolidado[['Atendidas']]
    x = query_consolidado[['Ndia']]

    X_train, X_test, y_train, y_test = train_test_split(x,y,test_size = 0.9)

    #Defino el algoritmo a utilizar
    lr = linear_model.LinearRegression()
    #Entreno el modelo
    lr.fit(X_train, y_train)
    #Realizo una predicción
    Y_pred = lr.predict(X_test)
    #Graficamos los datos junto con el modelo
    plt.scatter(X_test, y_test)
    plt.plot(X_test, Y_pred, color='red', linewidth=3)
    plt.title('Modelo de Regresión Lineal')
    plt.xlabel('Ndia')
    plt.ylabel('Número de atendidas')
    plt.show()
    print()
    print('Datos del Modelo Regresión Lineal')
    print()
    print('Valor de la pendiente "a":')
    print(lr.coef_)
    print('Valor de la intersección "b":')
    print(lr.intercept_)
    print()
    print('La ecuación del modelo es igual a:')
    print('y = ', lr.coef_, 'x ', lr.intercept_)
    print()
    print('Exactitud del modelo:')
    print(lr.score(X_train, y_train))
    if lr.score(X_train, y_train) > 0.10:
        break
    print("La predicción es: " )
    print(Y_pred)
    print(X_test)
    Y_pred = pd.DataFrame(lr.predict(Y_pred))
    llamadas_reales = pd.DataFrame(X_test)
prediction = pd.DataFrame(Y_pred)
prediction
```

4.2. Regresión Lineal Simple

```
from sklearn.linear_model import LinearRegression
from sklearn.metrics import mean_squared_error
while True:
    y = query_consolidado[['Atendidas']]
    x = query_consolidado[['Ndia']]

    X_train, X_test, y_train, y_test = train_test_split(x,y,test_size = 0.7)

    #Definir el algoritmo
    lr = linear_model.LinearRegression()
    #Entreno el modelo
    lr.fit(X_train, y_train)
    #Realizo una predicción
    Y_pred = lr.predict(X_test)
    #Graficamos los datos junto con el modelo
    plt.scatter(X_test, y_test)
    plt.plot(X_test, Y_pred, color='red', linewidth=3)
    plt.title('Regresión Lineal Múltiple')
    plt.xlabel('Ndia')
    plt.ylabel('Número de atendidas')
    plt.show()
    print()
    print('Datos del modelo regresion Lineal Multiple')
    print()
    print('Valor de la pendiente "a":')
    print(lr.coef_)
    print('Valor de la intersección "b":')
    print(lr.intercept_)
    print()
    print('La ecuación del modelo es igual a:')
    print('y = ', lr.coef_, 'x ', lr.intercept_)
    print()
    print('Exactitud del modelo:')
    print(lr.score(X_train, y_train))
    if lr.score(X_train, y_train) > 0.40:
        break
    print("La prediccion es: ")
    print(Y_pred)
    print(X_test)

    print('Datos del Modelo de Regresión Lineal Multiple')
    print()
    print('Valor de las pendientes "a":')
    print(lr.coef_)
    print('Valor de la intersección "b":')
    print(lr.intercept_)
    print('Exactitud del modelo:')
    print(lr.score(X_train, y_train))
    Y_pred = pd.DataFrame(lr.predict(Y_pred))
    llamadas_reales = pd.DataFrame(X_test)
    prediction = pd.DataFrame(Y_pred)
    prediction
```

4.3. Red Neuronal

```
from sklearn.neural_network import MLPRegressor
from sklearn.model_selection import train_test_split

startprocess = datetime.now()
time.sleep(2)

""" Bloque de Funciones """

def validationdataframe(d):
    if len(d) <= 0:
        print("El Conjunto de datos esta vacío")
    else:
        d = d
    return d

def settingloaddata(obj, cn):
    try:
        df = pd.read_sql(obj, cn, index_col='Id')
        df = df[df['Cancelar'] == 0]
    except ValueError:
        er = os.sys.exc_info()
        er_one = 'No se realizo la conexión'
        return er, er_one
    else:
        return df

""" Credenciales del servidor de origen de la tabla de configuración """
server = "localhost"
database = "Vertice_Empresarial"
user = "sa"
password = "miguel147258"
select = 'SELECT * FROM Setting.LoadData Where Id = 0'
cnx = Connection_String(server, database, user, password)
```

```
""" Conexión a base de datos """
table_configuration = settingloaddata(select, cnx)
pd.set_option('display.expand_frame_repr', False)

try:
    for regconf in table_configuration.itertuples():
        """ Obtener datos en variables de origen """
        driver_source = regconf[1]
        server_source = regconf[2]
        user_source = regconf[3]
        password_source = regconf[4]
        database_source = regconf[5]
        scheme_source = regconf[6]
        objeto_source = regconf[7]
        objeto_dest = regconf[15]
        user_dest = regconf[11]
        password_dest = regconf[12]
        server_dest = regconf[10]
        database_dest = regconf[13]
        driver_dest = regconf[16]
        scheme_dest = regconf[14]
        query_delete = regconf[17]
```

```

""" Abrir conexión Objeto Origen """
cnx_source = Connection_String(server_source, database_source, user_source, password_source)

""" Abrir conexión Objeto Destino"""
cnx_dest = Connection_String(server_dest, database_dest, user_dest, password_dest)

""" Consultar objeto """
query = objeto_source

try:
    print('Consultando: %s' % query, ' ...')
    query = pd.read_sql_query(query, cnx_source)
    print(query)
    y = np.array(query['Atendidas'])
    x = np.array(query['Ndia'])

    X = x[:, np.newaxis]

    while True:

        X_train, X_test, y_train, y_test = train_test_split(X,y, test_size = 0.8) #, test_size = 0.8

        mlr = MLPRegressor(solver="lbfgs", alpha=1e-5, hidden_layer_sizes=(3,3), random_state=1)
        mlr.fit(X_train, y_train)
        print("El score es " ,mlr.score(X_train, y_train))
        if mlr.score(X_train, y_train) >= 0.75:
            break

        Y_pred = pd.DataFrame(mlr.predict(X_test))
        llamadas_reales = pd.DataFrame(X_test)
        prediction = pd.DataFrame(Y_pred)

        pandas = llamadas_reales.join(prediction, lsuffix='Dia', rsuffix='PredicctionAtendida')

        predicctionAtendida = pd.DataFrame(pandas['0PredicctionAtendida'].unique())
        numeroDia = pd.DataFrame(pandas['0Dia'].unique())
        pandasSQL = numeroDia.join(predicctionAtendida, lsuffix='Dia', rsuffix='PredicctionAtendida')

```


5. Evaluación

En esta fase de la metodología se intentan evaluar los modelos generados, pero en esta ocasión la evaluación se hace desde el punto de vista de los objetivos de negocio en lugar de los objetivos de minería de datos. Una vez realizada esta evaluación, se debe decidir si los objetivos han sido cumplidos y de ser así se puede avanzar a la fase de implantación, de lo contrario se tendría que identificar cualquier factor que se haya podido pasar por alto y hacer una revisión del proceso.

5.1. Evaluar los Resultados

Desde el punto de vista del negocio, se había establecido como criterio de éxito principal el poder realizar predicciones con un porcentaje de fiabilidad “aceptable”, este criterio puede ser algo subjetivo, por lo que es inevitable apoyarse principalmente en los criterios de éxito desde el punto de vista que son mucho más específicos y precisos. Además, para poder calificar como aceptable o no las predicciones que se van a realizar es necesario tener una base objetiva, como lo son los indicadores estadísticos que se han obtenido al ejecutar los modelos. También sería conveniente la evaluación de los resultados por parte de un grupo de expertos, si se contara con ellos. En cualquier caso, basándonos en los indicadores obtenidos mediante la herramienta de minería de datos, a continuación, podemos hacer una evaluación de cada modelo para así descartar aquel que no cumpla con unos requisitos mínimos.

- **Modelo 1 para el objetivo.**

Este modelo no es viable ya que no nos ofrece suficientes garantías con tan sólo un 12% de confianza predictiva para poder realizar predicciones fiables acerca de las llamadas que se realizaran.

- **Modelo 2 para el objetivo**

Este modelo no es viable ya que no nos ofrece suficientes garantías con tan sólo un 22% de confianza predictiva para poder realizar predicciones fiables acerca de las llamadas que realizara el asesor para la gestión de ventas.

• **Modelo 3 para el objetivo**

Este modelo es factible ya que se pueden hacer predicciones acerca de cuántas llamadas tendrá un asesor con un porcentaje de fiabilidad de un 68%, el cual consideramos aceptable desde el punto de vista de los objetivos de negocio.

Modelo aprobado

Por las razones explicadas en este apartado y en el apartado 4.4 (evaluar el modelo) los modelos aprobados son el modelo 3 que cumple con los criterios de éxito de negocio, mientras que el modelo 1 y 2 será descartado por no cumplir con los requisitos de negocio ni machine Learning.

6. Distribución

Esta es la última fase de la metodología CRISP-DM y el objetivo de la misma es el de explicar al cliente como poner en funcionamiento el proyecto que se ha construido en las fases anteriores, así como exponer los resultados obtenidos al cliente de forma que lo pueda entender fácilmente. Otro objetivo de esta fase es el de crear una estrategia para el mantenimiento del proyecto y producir un informe en el que se incluyan posibles mejoras para el futuro y un listado de las dificultades encontradas a la hora de realizarlo.

6.1 Planear la Implantación

Para poder implantar este proyecto en el negocio real sería necesario en primer lugar tener acceso a la base de datos real del negocio, es decir la base de datos que contiene toda la información relativa a las llamadas del negocio. A partir de ahí, los pasos a seguir serían los mismos que se han seguido en este documento desde la comprensión del negocio hasta la implantación. Si bien, cabe decir que habrá algunas fases, como la de comprensión y preparación de los datos, que en el negocio real probablemente sean más complejas y llevarán más tiempo que en este proyecto ya que se puede esperar que en la base de datos real se tengan

muchos más registros y estos mismos contengan más ruido que en nuestra base de datos ficticia creada específicamente para este uso.

6.2 Planear la Monitorización y Mantenimiento

La supervisión y mantenimiento de la implementación del presente proyecto es una fase importante del mismo debido a que los datos que se procesan con mucha frecuencia pueden ser modificados por el personal de la Call center. Los datos pueden ser modificados por diferentes motivos como haber realizado una codificación incorrecta, haber asignado una nota incorrecta al alumno, etc. El volumen de estos datos en movimiento es grande motivo por el cual la extracción de las muestras debe ser realizada cuidadosamente y realizando siempre backups de los datos explotados en cada proceso. La minería de datos debería ser realizada en periodos de cuatro meses (cuatrimestres) ya que esta es la medida de tiempo utilizada en el call Center para realizar los exámenes y asignar las notas finales a los alumnos, sin embargo, esta medida podría variar en cualquier momento en función del plan de estudios que esté vigente en cada momento.

Como plan de supervisión y mantenimiento se podría establecer los siguientes procesos:

- Extracción y almacenamiento cuatrimestral de los datos guardando la información obtenida en formato de hoja de cálculo
- Distribución de los datos en función de los modelos de software de minería de datos a trabajar.
- Los archivos de la explotación de datos deberán ser guardados en soporte magnético en la propia universidad, almacenándolos por ejemplo en carpetas ordenadas por procesos cuatrimestrales.
- Los resultados obtenidos en cada explotación de datos deberán ser llevados a formato de hoja de cálculo y generar gráficas de distintos tipos para una mejor visualización e interpretación de los resultados obtenidos en cada periodo.

6.3. Producir el Informe Final

En este paso se debe presentar un informe resumiendo los puntos importantes del proyecto y la experiencia adquirida durante su desarrollo. El público al que va dirigido este informe sería el personal de sistemas encargado del registro de llamadas de tal manera que se pueda estudiar la situación actual y tomar medidas correctivas para la mejora del servicio académico. Cabe decir que parte de este informe final será presentado de manera oral con una presentación, por lo que en este apartado solamente haremos un breve resumen.

El uso de la metodología CRISP-DM en este proyecto ha permitido encontrar un comportamiento predictivo a la hora de estimar la duración de la carrera de los alumnos y la nota media de los mismos. Se ha podido encontrar un plan de extracción, normalización, y codificación de datos para la realización de procesos de datos cuatrimestrales.

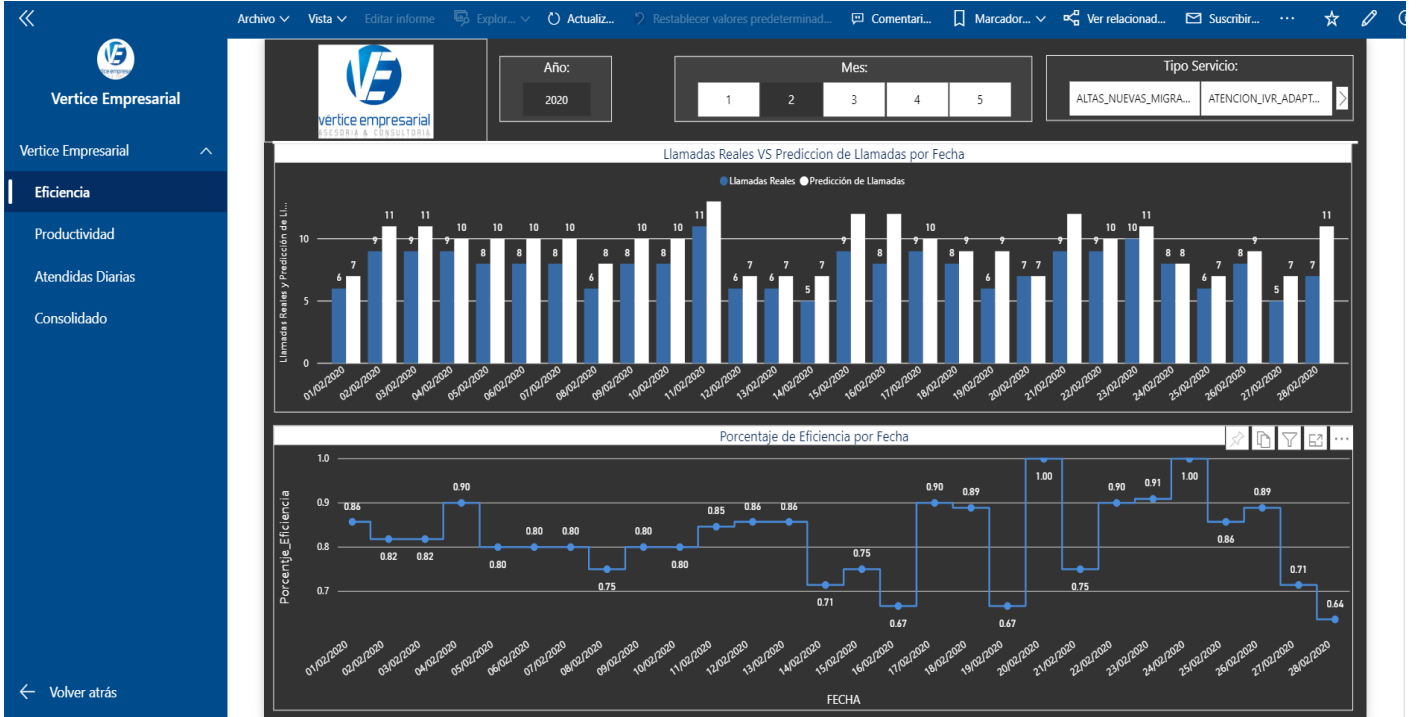
6.4. Revisar el Proyecto

En esta última etapa de la metodología se debe hacer una evaluación de aquellas cosas que se hicieron correctamente y aquellas que no, así como posibles mejoras para que en las futuras ejecuciones de la minería de datos se vayan puliendo los fallos y se obtengan mejores resultados.

En primer lugar, y como ya se ha comentado anteriormente en otros apartados, el mayor lastre que se ha ido arrastrando a lo largo de este proyecto es el de no disponer de una base de datos real sobre la que actuar ya que esto condiciona en gran medida los resultados obtenidos.

Aplicativo complementado con Power BI para la visualización de los datos de manera Instantánea y real.

- Eficiencia



- Productividad



- **Atendidas Diarias**



Consolidado LLamadas



Anexo Entrevista

Entrevista a La Coordinadora de Sistemas

Entrevista

La siguiente entrevista muestra las preguntas que se realizaron a la Coordinadora de Sistemas Viviana Rosell Carreño de la empresa Vértice empresarial, para así recolectar información la cual nos pueda mostrar un mejor panorama del estado de la empresa.

¿El área de sistemas está cumpliendo con los objetivos planteados?, ¿Por qué?

En la actualidad el área no está cumpliendo con los objetivos planteados, debido a que se deben enviar indicadores diariamente (dos veces al día), semanal y mensualmente. Los cuales son sumamente importantes para la toma de decisiones, estos no están siendo enviados según los requerimientos pedidos, además que existen datos que o se asemejan a la realidad pudiendo ocasionar una mala toma de decisiones para el área de operaciones.

¿Usted cree que el personal de sistemas es el adecuado? ¿Por qué?

Si, Por qué el personal que tiene la labor de realizar los indicadores pasa por un proceso de capacitación en el cual deben cumplir las funciones como asistentes de sistemas.oder cumplir con sus funciones como asistentes de sistemas.

¿Se están llegando a los objetivos puestos con el uso de indicadores? ¿Por qué?

No se está cumpliendo con los objetivos de los indicadores, debido a la demora de envíos en los datos, también porque son generados manualmente y hace que estos contengan un margen de error no especificado.

¿Usted cree que cuenta con los recursos necesarios para tener un buen control del negocio?

Actualmente no se tienen todos los recursos que se necesitan, ya que solo manejan herramientas como Excel para el cruce de datos y de tal forma obtener mediante fórmulas los datos establecidos o requeridos.

¿Cuál es el negocio de la empresa?

La venta y cobranza de productos establecidos por un Empleador mediante las telecomunicaciones, tecnología de información a todas las partes del país.

¿Qué servicios presta? ¿A quienes presta Servicios? ¿Para qué lo hace?

Servicios de Atención al Cliente, cobranzas y Ventas.

Presta servicios a empresas de sectores del rubro financiero, seguros. Lo hace para que puedan ver cómo van los objetivos que tienen con cada cliente.

¿Cómo se mide la productividad de los asesores en el área?

Mediante las llamadas realizadas durante el día, efectividad y cierre de venta.

Además tiempo de horas trabajas por asesor con el total de llamadas o la contactabilidad.

¿Cumplen sus funciones correctamente?

La productividad de los asesores se mide mediante el avance de las bases asignadas y los datos planificados por los informes enviados (indicadores) en el tiempo establecido por operaciones. Algunas veces se cumplen con las funciones que ya están establecidas y en otras depende seriamente de los recursos y herramientas que tienen para la elaboración de los datos.

¿Cómo elaboran los datos planificados en el área de operaciones?

Los datos planificados se elaboran de forma manual en hojas de cálculo Excel, en base a los avances por datas según los indicadores presentados.

¿Qué herramientas tecnológicas utilizan para la elaboración de informes?

Utilizan Excel para el cálculo, como gestor de base de datos MySQL, entre otros para descargas datos de la gestión diaria, mensual y anual.

¿Con que tipo de dificultades o limitaciones se cuenta para la generación de informes?

Algunas limitaciones en los recursos tecnológicos para la obtención de datos como por ejemplo la velocidad para procesar los datos en las computadoras, los discos duros, memorias RAM de acuerdo a la cantidad de datos que se quieran manejar y que resultados se quieran obtener.



Miguel Angel David Castañeda Rojas

Estudiante

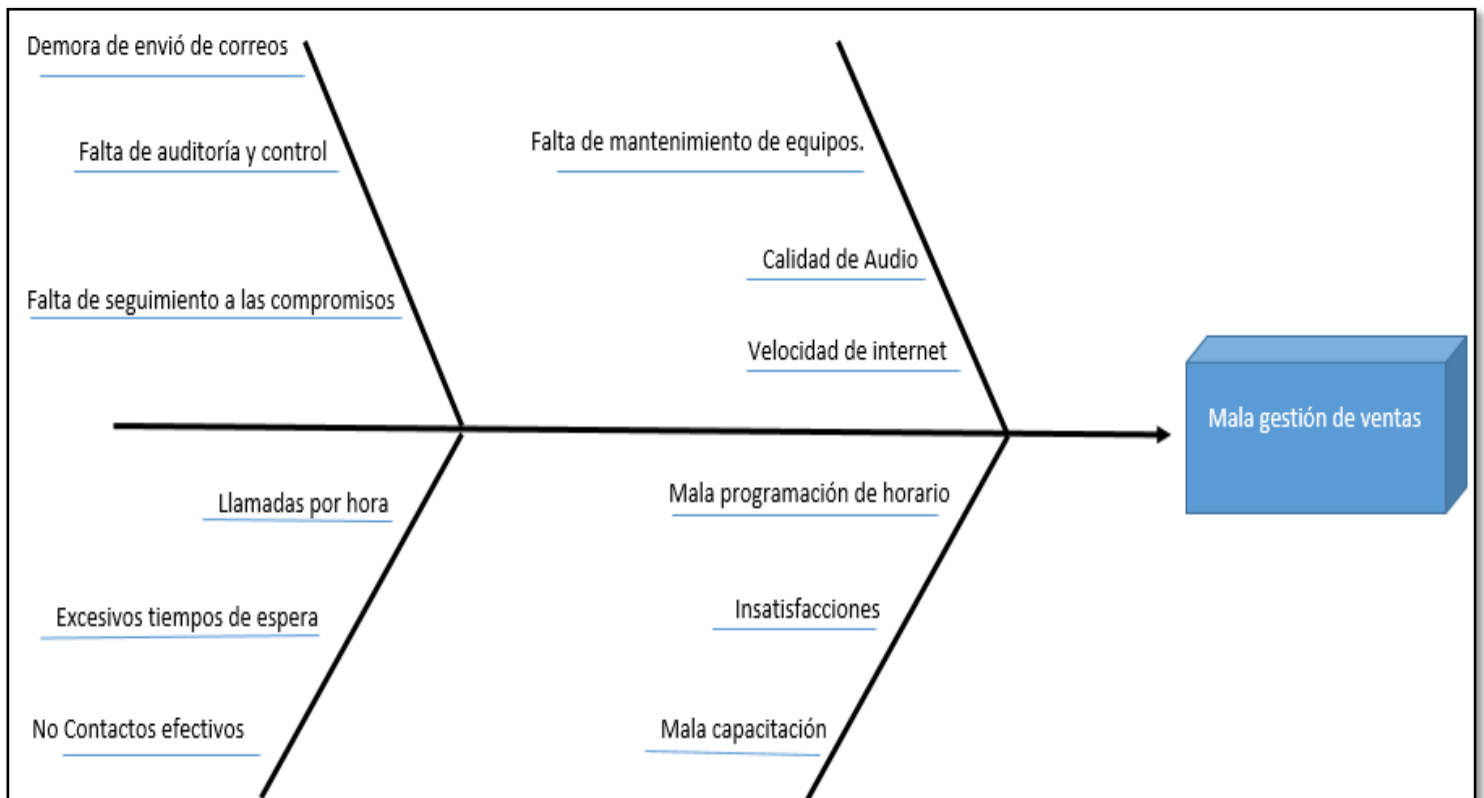
Anexo Matriz de Consistencia

PROBLEMA		OBJETIVOS		HIPOSTESIS		OPERACIONALIZACION DE VARIABLES			
General	General	General	General	General	VARIABLE	DIMENSION	INDICADOR	METODOLOGIA	
De que medida Machine Learning influye para la gestión de ventas en la empresa Vértice Empresarial S.A.C	Determinar la influencia de Machine Learning para la gestión de ventas en la empresa Vértice Empresarial S.A.C	General	Machine Learning mejora la gestión de ventas en la empresa Vértice Empresarial S.A.C.	Independiente				Tipo de Investigación: Experimental - Aplicada. Diseño de Estudio: Pre Experimental Población : I=490 I2=490	
En que medida Machine Learning influye en la eficiencia para la gestión de ventas en la empresa Vértice Empresarial S.A.C	Determinar la influencia de Machine Learning en la eficiencia para la gestión de ventas en la empresa Vértice Empresarial S.A.C	Específicos	Específicos Machine Learning aumenta la eficiencia para la gestión de ventas en la empresa Vértice Empresarial.	Dependiente			EFICIENCIA	Muestra: I=216 I2=216 Estratificada en 28 días. Técnica e Instrumentos de recolección de datos	
En que medida Machine Learning influye en la productividad para la gestión de ventas en la empresa Vértice Empresarial S.A.C	Determinar la influencia de Machine Learning en la producción para la gestión de ventas en la empresa Vértice Empresarial S.A.C	Específicos	Específicos Machine Learning aumenta la producción para la gestión de ventas en la empresa Vértice Empresarial.	GESTIÓN DE VENTAS	Busqueda de Clientes		PRODUCTIVIDAD	Técnicas: -Fichaje Instrumentos: -Ficha de Registro	

Anexo Operacionalización de Variables

OPERACIONALIZACIÓN DE VARIABLES				
Tipo	Variable	Definición Conceptual	Definición Operacional	Indicador
INDENDE	MACHINE LEARNING	Hurwitz, Judith y Kirsch, Daniel (2018, p60) modelos de aprendizaje y las organizaciones tienen la necesidad de predecir continuamente cambios en el negocio para que son más capaces de predecir qué sigue. Como son constantemente agregando, los modelos de aprendizaje automático los cuales dan una solución sean constantemente actualizada.	Machine Learning es un paso previo a la Inteligencia Artificial para que este se desarrolle en su totalidad. Concretamente, es un aprendizaje de máquinas que utilizan muchos deatos con el objetivo de ser cada vez mas inteligentes a la hora de la toma de decisiones, donde se alimenta de informacion, es algo como llevar la inteligencia del dato.	
DEPENDE	GESTIÓN DE VENTAS	Según TORRES, Virgilio (2015). define la gestión de ventas como: "La gestión comercial y tiene que identificar el mercado, definir en donde se establecerá en el cuál actuará la empresa, definir una meta y el posicionamiento, investigar el contexto y el comportamiento de los consumidores.	Las gestiones de venta consisten en ofrecer ciertos productos, los cuales tiene una sola finalidad tener un cierre de venta o comunicación efectiva para así obtener futuros ingresos y mejorar la calidad del servicio de esta.	<p>Eficiencia</p> <p>Busqueda del Cliente</p> <p>Productividad</p>

Anexo Diagrama de Ishikawa



Anexo Carta Aceptación de la Empresa



CARTA DE ACEPTACION DE LA EMPRESA VERTICE EMPRESARIAL S.A.C

Lima 20 de noviembre del 2019

Señores

COMITÉ DE PROGRAMA DE INVESTIGACION

Universidad César Vallejo

De manera atenta manifiesto nuestro interés y conocimiento de la propuesta del proyecto de investigación titulada.

“MACHINE LEARNING PARA LA GESTION DE VENTAS EN LA EMPRESA VERTICE EMPRESARIAL S.A.C.”

Elaborado por el estudiante:

Miguel Angel David Castañeda Rojas con Código de estudiante N°6700237990

En este sentido, me comprometo a participar en este proceso ofreciendo la información, el apoyo necesario de la UCV, como elemento de consulta para el público.

Conociendo y aceptando el reglamento y disposición sobre la realización de opciones de grado de la UCV.

Cordialmente.

Representante Legal

Antonio Llanos Cardenas
Gerente General
Dirección: Jr. Camaná N°380
Oficina 510-Cercado de Lima
T (511) 700-8500 - Anexo 500



Anexo Carta de Implementación de la Empresa



CARTA DE IMPLEMENTACION DE LA EMPRESA VERTICE EMPRESARIAL S.A.C

Lima 15 de Junio del 2020

Señores

Universidad César Vallejo

Presente

Asunto: Aprobación de la implementación de desarrollo de tesis

De mi mayor consideración:

Tengo el agrado de dirigirme a UD., saludarlo y a la vez hacer de su conocimiento que el alumno Miguel Angel David Castañeda Rojas con Código de estudiante N°6700237990 logro la implementación con éxito de la investigación cuyo título **“MACHINE LEARNING PARA LA GESTION DE VENTAS EN LA EMPRESA VERTICE EMPRESARIAL S.A.C.”** el día 15 de Junio del 2020, en las instalaciones de la empresa ubicado en Jirón Camaná Lima, utilizando los recursos de la empresa.

Sin otro particular, agradeciendo la atención brindada me despido de usted.

Cordialmente.

Representante Legal

Antonio Llanos Cardenas
Gerente General
Dirección: Jr. Camaná N°380
Oficina 510-Cercado de Lima
T (511) 700-8500 - Anexo 500



Anexo Ficha de Registro Eficiencia Pretest

Ficha de Registro	
Investigador	Castañeda Rojas Miguel Angel David
Empresa	Vértice Empresarial S.A.C
Investigación	Machine Learning Para la Gestión de Ventas en la Empresa Vértice Empresarial S.A.C
VARIABLES	Gestión de Ventas
Dimensión	Búsqueda De Cliente
Indicador	Formula
Eficiencia	Predicción Llamadas / Llamadas Reales

Nº	Fecha	Predicción Llamadas	Llamadas reales	Eficiencia
1	01/09/2019	2	6	0,33
2	02/09/2019	3	6	0,50
3	03/09/2019	3	6	0,50
4	04/09/2019	3	5	0,60
5	05/09/2019	4	5	0,80
6	06/09/2019	4	5	0,80
7	07/09/2019	3	5	0,60
8	08/09/2019	5	12	0,42
9	09/09/2019	4	8	0,50
10	10/09/2019	5	12	0,42
11	11/09/2019	6	13	0,46
12	12/09/2019	5	8	0,63
13	13/09/2019	5	9	0,56
14	14/09/2019	3	7	0,43
15	15/09/2019	3	8	0,38
16	16/09/2019	3	4	0,75
17	17/09/2019	3	5	0,60
18	18/09/2019	4	7	0,57
19	19/09/2019	3	5	0,60
20	20/09/2019	3	6	0,50
21	21/09/2019	3	5	0,60
22	22/09/2019	7	11	0,64
23	23/09/2019	5	10	0,50
24	24/09/2019	5	13	0,38
25	25/09/2019	3	5	0,60
26	26/09/2019	4	8	0,50
27	27/09/2019	6	13	0,46
28	28/09/2019	6	9	0,67



Anexo Ficha de Registro Eficiencia PostTest

Ficha de Registro	
Investigador	Castañeda Rojas Miguel Angel David
Empresa	Vértice Empresarial S.A.C
Investigación	Machine Learning Para la Gestión de Ventas en la Empresa Vértice Empresarial S.A.C
VARIABLES	Gestión de Ventas
Dimensión	Búsqueda De Cliente
Indicador	Formula
Eficiencia	Predicción Llamadas / Llamadas Reales

Nº	Fecha	Predicción Llamadas	Llamadas reales	Eficiencia
1	01/02/2020	7	6	0,86
2	02/02/2020	11	9	0,82
3	03/02/2020	11	9	0,82
4	04/02/2020	10	9	0,90
5	05/02/2020	10	8	0,80
6	06/02/2020	10	8	0,80
7	07/02/2020	10	8	0,80
8	08/02/2020	8	6	0,75
9	09/02/2020	10	8	0,80
10	10/02/2020	10	8	0,80
11	11/02/2020	13	11	0,85
12	12/02/2020	7	6	0,86
13	13/02/2020	7	6	0,86
14	14/02/2020	7	5	0,71
15	15/02/2020	12	9	0,75
16	16/02/2020	12	8	0,67
17	17/02/2020	10	9	0,90
18	18/02/2020	9	8	0,89
19	19/02/2020	9	6	0,67
20	20/02/2020	7	7	1,00
21	21/02/2020	12	9	0,75
22	22/02/2020	10	9	0,90
23	23/02/2020	11	10	0,91
24	24/02/2020	8	8	1,00
25	25/02/2020	7	6	0,86
26	26/02/2020	9	8	0,89
27	27/02/2020	7	5	0,71
28	28/02/2020	11	7	0,64



Anexo Ficha de Registro Productividad Pretest

Ficha de Registro	
Investigador	Castañeda Rojas Miguel Angel David
Empresa	Vértice Empresarial S.A.C
Investigación	Machine Learning Para la Gestión de Ventas en la Empresa Vértice Empresarial S.A.C
Variables	Gestión de Ventas
Dimensión	Búsqueda De Cliente
Indicador	Formula
Productividad	Horas Trabajadas /Llamadas Reales

Nº	Fecha	Horas Trabajadas	llamadas reales	Productividad
1	01/09/2019	3	7	0,43
2	02/09/2019	3	7	0,43
3	03/09/2019	2	8	0,25
4	04/09/2019	3	5	0,60
5	05/09/2019	6	13	0,46
6	06/09/2019	3	8	0,38
7	07/09/2019	3	6	0,50
8	08/09/2019	3	6	0,50
9	09/09/2019	5	8	0,63
10	10/09/2019	2	6	0,33
11	11/09/2019	6	8	0,75
12	12/09/2019	5	9	0,56
13	13/09/2019	5	9	0,56
14	14/09/2019	4	8	0,50
15	15/09/2019	4	8	0,50
16	16/09/2019	3	5	0,60
17	17/09/2019	4	7	0,57
18	18/09/2019	4	7	0,57
19	19/09/2019	5	9	0,56
20	20/09/2019	4	9	0,44
21	21/09/2019	4	8	0,50
22	22/09/2019	4	8	0,50
23	23/09/2019	3	8	0,38
24	24/09/2019	4	7	0,57
25	25/09/2019	3	6	0,50
26	26/09/2019	6	9	0,67
27	27/09/2019	5	8	0,63
28	28/09/2019	5	9	0,56



Anexo Ficha de Registro Productividad PostTest

Ficha de Registro	
Investigador	Castañeda Rojas Miguel Angel David
Empresa	Vértice Empresarial S.A.C
Investigación	Machine Learning Para la Gestión de Ventas en la Empresa Vértice Empresarial S.A.C
Variables	Gestión de Ventas
Dimensión	Búsqueda De Cliente
Indicador	Formula
Productividad	Horas Trabajadas /Llamadas Reales

Nº	Fecha	Horas Trabajadas	llamadas reales	Productividad
1	01/02/2020	5	7	0,71
2	02/02/2020	7	8	0,88
3	03/02/2020	8	10	0,80
4	04/02/2020	8	10	0,80
5	05/02/2020	6	8	0,75
6	06/02/2020	7	8	0,88
7	07/02/2020	4	5	0,80
8	08/02/2020	6	8	0,75
9	09/02/2020	8	8	1,00
10	10/02/2020	5	6	0,83
11	11/02/2020	6	7	0,86
12	12/02/2020	6	8	0,75
13	13/02/2020	6	7	0,86
14	14/02/2020	5	6	0,83
15	15/02/2020	8	8	1,00
16	16/02/2020	7	9	0,78
17	17/02/2020	8	10	0,80
18	18/02/2020	6	8	0,75
19	19/02/2020	5	9	0,56
20	20/02/2020	5	8	0,63
21	21/02/2020	7	10	0,70
22	22/02/2020	5	6	0,83
23	23/02/2020	4	5	0,80
24	24/02/2020	8	9	0,89
25	25/02/2020	6	8	0,75
26	26/02/2020	7	8	0,88
27	27/02/2020	5	6	0,83
28	28/02/2020	6	6	1,00



Anexo Ficha de Evaluación de Expertos

TABLA DE EVALUACIÓN DE EXPERTOS

Apellidos y Nombres del experto: **ING. RIVERA CALBOSZANO RENEE**

Título y/o Grado: **ING. DE SISTEMAS**

Doctor.....()	Magister.....(x)	Ingeniero.....()	Otros.....especifique
----------------	------------------	-------------------	-----------------------

Institución donde labora: Universidad Cesar Vallejo Sede Lima Norte

Autores: Castañeda Rojas Miguel Angel David

Fecha: **06-11-19**

TITULO DE TESIS

MACHINE LEARNING PARA LA GESTION DE VENTAS EN LA EMPRESA VERTICE EMPRESARIAL S.A.C

Mediante la tabla de evaluación de expertos, usted tiene la facultad de calificar las metodologías involucradas, mediante una serie de criterios con puntuaciones específicas al final de la tabla.

Evaluar con la siguiente puntuación:

(1) Muy Malo (2) Malo (3) Regular (4) Bueno (5) Muy Bueno

ITEMS	CRITERIOS	METODOLOGÍA			OBSERVACIONES
		CRISP DM	SEMMA	KDD	
1	Comprende el proceso del negocio	4	2	2	
2	Define las Necesidades del cliente	5	3	1	
3	Estudia y comprende los datos	5	4	4	
4	Realiza un análisis de datos y selecciona características.	4	3	3	
5	Logra realizar un modelado con los datos obtenidos.	5	4	1	
6	Evalúa los resultados obtenidos	4	2	2	
7	Realiza un despliegue con los datos obtenidos.	5	1	4	
	TOTAL	32	19	16	

SUGERENCIAS:

.....

.....
Firma del Experto

TABLA DE EVALUACIÓN DE EXPERTOS

Apellidos y Nombres del experto: *Berniejo Terrones Henry Paul*
 Título y/o Grado: *Maestro en Ingeniería de Sistemas*

Doctor.....()	Magister.....(X)	Ingeniero.....()	Otros.....especifique
----------------	------------------	-------------------	-----------------------

Institución donde labora: Universidad Cesar Vallejo Sede Lima Norte

Autores: Castañeda Rojas Miguel Angel David

Fecha: *06/11/19*

TITULO DE TESIS

MACHINE LEARNING PARA LA GESTION DE VENTAS EN LA EMPRESA VERTICE EMPRESARIAL S.A.C

Mediante la tabla de evaluación de expertos, usted tiene la facultad de calificar las metodologías involucradas, mediante una serie de criterios con puntuaciones específicas al final de la tabla.

Evaluar con la siguiente puntuación:

- (1) Muy Malo (2) Malo (3) Regular (4) Bueno (5) Muy Bueno

ITEMS	CRITERIOS	METODOLOGÍA			OBSERVACIONES
		CRISP DM	SEMMA	KDD	
1	Comprende el proceso del negocio	4	3	3	
2	Define las Necesidades del cliente	5	4	3	
3	Estudia y comprende los datos	5	3	4	
4	Realiza un análisis de datos y selecciona características.	5	2	4	
5	Logra realizar un modelado con los datos obtenidos.	4	4	3	
6	Evalúa los resultado obtenidos	4	2	2	
7	Realiza un despliegue con los datos obtenidos.	5	1	4	
	TOTAL	32	19	23	

SUGERENCIAS:

.....

.....
 Firma del Experto

TABLA DE EVALUACIÓN DE EXPERTOS

Apellidos y Nombres del experto: Gálvez Tapia Orleans
 Título y/o Grado: Magister en Ingeniería de sistemas.

Doctor.....()	Magister.....(<input checked="" type="checkbox"/>)	Ingeniero.....()	Otros.....especifique
----------------	--	-------------------	-----------------------

Institución donde labora: Universidad Cesar Vallejo Sede Lima Norte

Autores: Castañeda Rojas Miguel Angel David

Fecha:

TITULO DE TESIS

**MACHINE LEARNING PARA LA GESTION DE VENTAS EN LA EMPRESA
VERTICE EMPRESARIAL S.A.C**

Mediante la tabla de evaluación de expertos, usted tiene la facultad de calificar las metodologías involucradas, mediante una serie de criterios con puntuaciones específicas al final de la tabla.

Evaluar con la siguiente puntuación:

(1) Muy Malo (2) Malo (3) Regular (4) Bueno (5) Muy Bueno

ITEMS	CRITERIOS	METODOLOGÍA			OBSERVACIONES
		CRISP DM	SEMMA	KDD	
1	Comprende el proceso del negocio	5	4	3	
2	Define las Necesidades del cliente	5	4	3	
3	Estudia y comprende los datos	5	4	3	
4	Realiza un análisis de datos y selecciona características.	5	4	3	
5	Logra realizar un modelado con los datos obtenidos.	5	4	3	
6	Evalúa los resultado obtenidos	5	4	3	
7	Realiza un despliegue con los datos obtenidos.	5	4	3	
	TOTAL	35	28	21	

SUGERENCIAS:

.....



 Firma del Experto

Anexo Ficha de Evaluación de Indicadores



UNIVERSIDAD CÉSAR VALLEJO

Evaluación de Experto de Indicadores

Apellidos y Nombres del Experto: Gálvez Tapra Orleans

Fecha:

Título de Proyecto

Machine Learning para la Gestión de Venta de la Empresa Vértice Empresarial S.A.C.

Indicador - Productividad

Mediante la tabla de evaluación de expertos, usted tiene la facultad para medir el indicador, mediante una serie de preguntas con puntuaciones especificada al inicio de la tabla. Asimismo, le exhortamos en la corrección de ítems indicando sus observaciones y/o sugerencias, con la finalidad de mejorar la coherencia de preguntas.

5	4	3	2	1
Muy positiva	Medianamente positiva	Neutra	Medianamente negativa	Muy negativa

ITEMS	PREGUNTAS	VALORACIÓN				
		1	2	3	4	5
1	¿Cumple con el diseño adecuado?					89%
2	¿Tiene relación con el título de la investigación?					91%
3	¿Se mencionan las variables de la investigación?					92%
4	¿Facilitará el logro de los objetivos de investigación?					85%
5	¿Se relaciona con la variable de estudio?					80%
6	¿Facilitará el análisis y procesamiento de datos?					93%
7	¿Son entendible sus alternativas de respuestas?					89%
8	¿Será accesible a la población sujeto de estudio?					90%
9	¿Es claro, preciso y sencillo para que contesten y de esa manera obtener los datos requeridos?					91%
10	¿Los indicadores son los adecuados para la medición?					85%

88.5%

Deuf

TABLA DE EVALUACIÓN DE EXPERTOS

Datos del experto:

1. Apellidos y Nombres: Bermijo Terrones Henry Paul
2. Cargo que sustenta: ~~Docente~~ Docente Tiempo Completo
3. Título y/o grado: Maestro en Ingeniería de Sistemas
4. Fecha: 06/11/19

TESIS:

"MACHINE LEARNING PARA LA GESTION DE VENTAS EN LA EMPRESA VERTICE EMPRESARIAL S.A.C"

Tabla de Evaluación de expertos para el indicador: **PRODUCTIVIDAD**

Mediante la tabla de evaluación de expertos, usted tiene la facultad de calificar los instrumentos utilizados para medir un indicador. Asimismo, le exhortamos en la corrección de los ítems indicando sus observaciones y/o sugerencias, con la finalidad de mejorar la coherencia de las preguntas

Items	Pregunta	Deficiente	Regular	Bueno	Regular	Excelente
		0-20%	21%-50%	51%-70%	71%-80%	81%-100%
1	¿El instrumento de medición cumple con el diseño adecuado?					90%
2	¿El instrumento de recolección de datos tiene relación con el título de la investigación?					95%
3	¿En el instrumento de recolección de datos se mencionan las variables de investigación?					90%
4	¿El instrumento de recolección de datos facilitará el logro de los objetivos de investigación?					92%
5	¿El instrumento de medición será accesible a la población sujeto de estudio?					93%
6	¿El instrumento de medición es claro preciso y sencillo para que contesten y de esa manera obtener los datos requeridos?					94%
Total						923%

FIRMA

TABLA DE EVALUACIÓN DE EXPERTOS

Datos del experto:

1. Apellidos y Nombres: MJ. Rivera Carlosotta Revue
2. Cargo que sustenta: DTC
3. Título y/o grado: INS. DE SISTEMAS
4. Fecha: 06-11-19

TESIS:

"MACHINE LEARNING PARA LA GESTION DE VENTAS EN LA EMPRESA VERTICE EMPRESARIAL S.A.C"

Tabla de Evaluación de expertos para el indicador: **PRODUCTIVIDAD**

Mediante la tabla de evaluación de expertos, usted tiene la facultad de calificar los instrumentos utilizados para medir un indicador. Asimismo, le exhortamos en la corrección de los ítems indicando sus observaciones y/o sugerencias, con la finalidad de mejorar la coherencia de las preguntas

Items	Pregunta	Deficiente 0-20%	Regular 21%-50%	Bueno 51%-70%	Regular 71%-80%	Excelente 81%-100%
1	¿El instrumento de medición cumple con el diseño adecuado?					85%
2	¿El instrumento de recolección de datos tiene relación con el título de la investigación?					89%
3	¿En el instrumento de recolección de datos se mencionan las variables de investigación?					90%
4	¿El instrumento de recolección de datos facilitará el logro de los objetivos de investigación?					92%
5	¿El instrumento de medición será accesible a la población sujeto de estudio?					85%
6	¿El instrumento de medición es claro preciso y sencillo para que contesten y de esa manera obtener los datos requeridos?					90%
Total						88.5%



 FIRMA



UNIVERSIDAD CÉSAR VALLEJO

Evaluación de Experto de Indicadores

Apellidos y Nombres del Experto: Gálvez Tapra Orleans.

Fecha:

Título de Proyecto

Machine Learning para la Gestión de Venta de la Empresa Vértice Empresarial S.A.C.

Indicador - Eficiencia

Mediante la tabla de evaluación de expertos, usted tiene la facultad para medir el indicador, mediante una serie de preguntas con puntuaciones especificada al inicio de la tabla. Asimismo, le exhortamos en la corrección de ítems indicando sus observaciones y/o sugerencias, con la finalidad de mejorar la coherencia de preguntas.

5	4	3	2	1
Muy positiva	Medianamente positiva	Neutra	Medianamente negativa	Muy negativa

ITEMS	PREGUNTAS	VALORACIÓN				
		1	2	3	4	5
1	¿Cumple con el diseño adecuado?					80%
2	¿Tiene relación con el título de la investigación?					85%
3	¿Se mencionan las variables de la investigación?					90%
4	¿Facilitará el logro de los objetivos de investigación?					92%
5	¿Se relaciona con la variable de estudio?					95%
6	¿Facilitará el análisis y procesamiento de datos?					83%
7	¿Son entendible sus alternativas de respuestas?					91%
8	¿Será accesible a la población sujeto de estudio?					90%
9	¿Es claro, preciso y sencillo para que contesten y de esa manera obtener los datos requeridos?					89%
10	¿Los indicadores son los adecuados para la medición?					90%

Dany

TABLA DE EVALUACIÓN DE EXPERTOS

Datos del experto:

1. Apellidos y Nombres: Bermijo Terrones Henry Paul
2. Cargo que sustenta: Docente tiempo completo
3. Título y/o grado: Maestro en Ingeniería de Sistemas
4. Fecha: 06/11/19

TESIS:

"MACHINE LEARNING PARA LA GESTION DE VENTAS EN LA EMPRESA VERTICE EMPRESARIAL S.A.C"

Tabla de Evaluación de expertos para el indicador: **EFICIENCIA**

Mediante la tabla de evaluación de expertos, usted tiene la facultad de calificar los instrumentos utilizados para medir un indicador. Asimismo, le exhortamos en la corrección de los ítems indicando sus observaciones y/o sugerencias, con la finalidad de mejorar la coherencia de las preguntas

Items	Pregunta	Deficiente	Regular	Bueno	Regular	Excelente
		0-20%	21%-50%	51%-70%	71%-80%	81%-100%
1	¿El instrumento de medición cumple con el diseño adecuado?					90%
2	¿El instrumento de recolección de datos tiene relación con el título de la investigación?					90%
3	¿En el instrumento de recolección de datos se mencionan las variables de investigación?					89%
4	¿El instrumento de recolección de datos facilitará el logro de los objetivos de investigación?					92%
5	¿El instrumento de medición será accesible a la población sujeto de estudio?					95%
6	¿El instrumento de medición es claro preciso y sencillo para que contesten y de esa manera obtener los datos requeridos?					90%
Total						91.0%

FIRMA

TABLA DE EVALUACIÓN DE EXPERTOS

Datos del experto:

1. Apellidos y Nombres: Hcy. DÍEBA CALSOSORO REYES
2. Cargo que sustenta: DTC
3. Título y/o grado: ING. DE SISTEMAS
4. Fecha: 06-11-19

TESIS:

"MACHINE LEARNING PARA LA GESTION DE VENTAS EN LA EMPRESA VERTICE EMPRESARIAL S.A.C"

Tabla de Evaluación de expertos para el indicador: **EFICIENCIA**

Mediante la tabla de evaluación de expertos, usted tiene la facultad de calificar los instrumentos utilizados para medir un indicador. Asimismo, le exhortamos en la corrección de los ítems indicando sus observaciones y/o sugerencias, con la finalidad de mejorar la coherencia de las preguntas

Items	Pregunta	Deficiente	Regular	Bueno	Regular	Excelente
		0-20%	21%-50%	51%-70%	71%-80%	81%-100%
1	¿El instrumento de medición cumple con el diseño adecuado?					85%
2	¿El instrumento de recolección de datos tiene relación con el título de la investigación?					90%
3	¿En el instrumento de recolección de datos se mencionan las variables de investigación?					90%
4	¿El instrumento de recolección de datos facilitará el logro de los objetivos de investigación?					97%
5	¿El instrumento de medición será accesible a la población sujeto de estudio?					92%
6	¿El instrumento de medición es claro preciso y sencillo para que contesten y de esa manera obtener los datos requeridos?					91%
Total						90.8%


 FIRMA