**ANÁLISIS SOBRE LA PREDICCIÓN DEL CLIENTE PARA DIFERIR LAS CUOTAS DE SU CRÉDITO BANCARIO**

**AUTORES:**

**CYNDY ELIZABETH PANTOJA TAMAYO**

**DANIELA ESTEFANY ROJAS PORTILLA**

**DOCENTE:**

**ÁLVARO OÑATE BOWEN**

**UNIVERSIDAD POPULAR DEL CESAR**

**FACULTAD DE INGENIERIAS Y TECNOLOGIA**

**INGENIERÍA DE SISTEMAS**

**VALLEDUPAR – CESAR**

**2021 – 1**

**INTRODUCCIÓN**

En esta época de la tecnología es de vital importancia para todas las organizaciones la búsqueda de una buena prestación de servicio para sus clientes y usuarios, de modo que de esto depende la fidelidad del usuario hacia su compañía, por lo que es importante satisfacer las necesidades al cliente totalmente.

Los objetivos de las compañías deben ser claros y estar enfocados hacia las necesidades del cliente, sin embargo, este punto no se encuentra totalmente fijo desviando de un lado las necesidades de los consumidores.

Para que se pueda fortalecer el servicio al cliente dentro de las organizaciones se debe tener en cuenta como primera medida la estructura organizacional y un adecuado sistema y administración que facilite que la entidad cumpla con su misión y las estrategias corporativas y de servicio.

Cuando una entidad se enfoque hacia el servicio a sus clientes obtendrá mejores resultados en la misión logrando un cambio hacia la cultura del servicio.

El concepto de Marketing, varia en su significado de una organización a otra, pero en términos generales se puede simplificar como el sistema de investigar un mercado, ofrecer valor y satisfacer al cliente con un objetivo de lucro, se responsabiliza de estudiar el comportamiento de los mercados y de las necesidades de los consumidores.

**ESTADO DEL ARTE**

El marketing bancario es el conjunto de acciones y técnicas empleadas en el sector bancario para mejorar la comercialización de sus productos o servicios. Los bancos lo aplican en su día a día, dentro de un entorno muy complejo y cambiante, con un cliente cada vez, más exigente.

Según SOTO [1], Una de las problemáticas que se afrontan en el contexto del marketing bancario, es como ofrecer adecuadamente los productos que se poseen, es decir, la realización de una estrategia enfocada en los requerimientos de cada cliente, sin incurrir en el desagradable marketing viral. En este ámbito, se busca generar un modelo, que formule políticas que ayuden en la toma de decisiones, al momento de ofrecer los productos que posee la institución, de manera que esto se realice de forma efectiva, con el objetivo de maximizar los beneficios asociados a las políticas generadas por el método utilizado.

Para entrar en contexto el proceso de esta investigación, utilizo una metodología denominada CRISP-DM, el cual es un modelo de procesos perteneciente al área de minería de datos, que posee varias etapas: utiliza la tecnología suministrada por SQLalchemy (https://www.sqlalchemy.org/) para consultar toda la información y al almacenarla en un único dataframe y una matriz de tenencia de productos, otra de la correlación entre productos y finalmente una base que posea el rating de los productos para cada usuario. El dataframe, está conformado por las columnas “user id” que poseer el identificador de cada cliente, sociado a cada producto y el nombre de dicho producto. Finalmente, el rating o calificación que cada usuario posee de dicho producto en base al modelo RFM.

El modelo RFM es una metodología con la que se puede segmentar clientes según características similares. Este modelo contempla la Recencia, Frecuencia y Valor Monetario para cada cliente a partir del cual determinamos el comportamiento o evolución de compra de estos.

Según Muñoz, Para examinar el comportamiento del cliente se utilizan métodos estadísticos como el método CHAID. Es un método de árboles de decisión, el método utiliza 2 tipos de variables predictivas o independientes, utiliza como principal herramienta el estadístico chi-cuadrado con el cual particiona el conjunto de datos en subconjunto mutuamente exclusivos que describen de mejor manera la variable independiente. La liberalización de los mercados y el aumento de la competencia incrementan las exigencias de los clientes, que desean un mejor servicio, más adaptado a sus necesidades y más personalizado. Por ello, el marketing de la sociedad de la información trata de integrar al cliente en la empresa para crear una relación permanente y beneficiosa entre ambos, que sirva, además, de referencia para la captación de nuevos clientes. Para llevar a cabo esto se necesita disponer de grandes bases de datos interactivas tanto de clientes actuales como potenciales, incluso en grandes mercados de consumo. Esto solo ha sido posible cuando el desarrollo de las TIC ha dotado los profesionales del marketing de potentes sistemas informáticos. [2]

Implementar metodologías para lograr la retención eficiente de los clientes, es una meta diaria que toda compañía debe generarle mayor importancia, es decir establecer el conjunto de procedimientos a realizar desde un marco conceptual hasta el modelamiento mediante técnicas de minería de datos, para su posterior implementación en el Banco.

[3]Cross-Industry Standard Process for Data Mining (CRISP-DM) Actualmente es una de las más utilizada ya que puede llevarse a cabo en cualquier software minero, se compone de 6 etapas.

• Entendimiento del negocio: se concentra en los propósitos y requisitos desde un punto de vista del negocio.

• Entendimiento de los datos: recaudo de datos, posteriormente familiarización con ellos y formulación de hipótesis, a partir de ellos.

• Preparación de datos: consiste en construir la base final, dentro de esta etapa se debe realizar tareas que incluyen tabular, documentar y seleccionar particularidades de igual manera transformar variables.

• Modelado: consiste en la aplicación de diversas técnicas, donde se calibran los lineamientos de forma óptima, sin embargo, es posible que se tenga que modificar la data si una de las técnicas tiene requisitos específicos para poder aplicarse.

• Evaluación: se evalúa el modelo, si cumple con los objetivos propuestos y si tiene en cuenta todos los requerimientos inicialmente planteados.

• Despliegue: el conocimiento adquirido de acuerdo con el modelo realizado debe organizarse y presentarse de forma que la compañía lo pueda usar, es decir se debe implementar como parte de un proceso de toma de decisiones de una empresa

Según los autores Altamiranda, Peña, Ospino, Volpe, y Ortega [4], Actualmente se estima que el suministro de datos del mundo se duplica cada 20 meses. Lo anterior implica un crecimiento excesivo en el volumen de datos que se maneja tanto en la comunidad científica como en los sectores productivos de la economía, que ha sobrepasado la capacidad humana de analizar, resumir y extraer conocimientos a tales cantidades de datos. Lo anterior, hace necesaria una nueva generación de herramientas capaces de automatizar el análisis de los datos almacenados. El conjunto de estas herramientas lo estudia un nuevo campo de investigación llamado minería de datos. La minería de datos se ha convertido en una herramienta estratégica para la toma de decisiones de mercadeo, producción, organización y demás factores en la empresa, que de cierta manera la hacen más competitiva. El presente artículo analiza la manera como la minería de datos, una técnica ampliamente relacionada con la investigación de operaciones, incide en el diseño de estrategias de mercadeo B2B, en empresas de sectores industriales. Esta investigación guarda relación directa Minería de datos como herramienta para el desarrollo de estrategias de mercadeo en sectores productivos, afines a los colombianos.

Por otro lado, el ámbito financiero es uno de los sectores más complejos de nuestra sociedad, pues está sometido a grandes cambios, fluctuaciones y modificaciones que se producen por un extenso número de factores.

Además, es un sector muy competitivo y todavía tradicional, en el que las grandes entidades bancarias y financieras tienen cada vez más necesidad de diferenciarse unas de las otras, por lo que es imprescindible llevar a cabo procesos de innovación, tanto en productos y servicios, como en trato al cliente, atención y comunicación.

**METODOLOGÍA CRISP – DM**

CRISP – DM es un método probado para orientar trabajos de minería de datos, sus siglas en ingles significan Cross-Industry Standard Process for Data Mining.

* Como metodología, incluye descripciones de las fases normales de un proyecto, las tareas necesarias en cada fase y una explicación de las relaciones entre las tareas.
* Como modelo de proceso, CRISP – DM ofrece un resumen del ciclo vital de minería de datos.

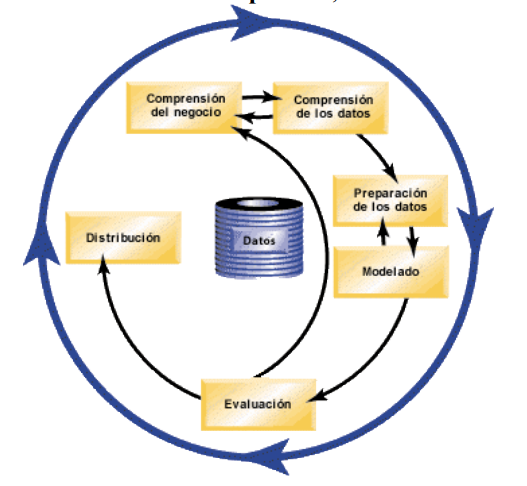


Figura 1. Metodología CRISP – DM

El ciclo vital del modelo contiene seis fases con flechas que indican las dependencias más importantes y frecuentes entre fases. La secuencia de las fases no es estricta. De hecho, la mayoría de los proyectos avanzan y retroceden entre fases si es necesario. El modelo de CRISP-DM es flexible y se pueden personalizar fácilmente.

**IMPLEMENTACIÓN DE LA METODOLOGÍA**

1. **FASE 1: COMPRENSIÓN DEL NEGOCIO**

El objetivo de esta fase es alinear los objetivos del proyecto de data mining con los objetivos del negocio. Tratando así de evitar embarcarnos en un proyecto de minería de datos que no produzca ningún efecto real.

En esta fase deberemos ser capaces de:

Establecer los objetivos de negocio.

Evaluar la situación actual.

Fijar los objetivos a nivel de minería de datos.

Obtener un plan de proyecto.

Como ejemplo ilustrativo de la metodología vamos a considerar el caso con una entidad bancaria que pretende predecir los clientes que aplazaran el pago de las cuotas de su crédito o la adquisición de uno nuevo y lograr a su vez maximizar la cobertura de los plazos fijos.

Para ello se procederá a realizar un estudio de una dataset descargada en los repositorios de datos de https://archive.ics.uci.edu/ml/datasets/Bank+Marketing

El objetivo final es pronosticar que clientes realizaran el plazo de su crédito, reducir los llamados tratando de contactar a la mayor cantidad de clientes propensos de adquirir un plazo fijo, mejorando la precisión diagnóstica.

Una vez establecidos y alineados los objetivos se creará un plan de proyecto que incluya las fases, tareas y actividades necesarias para conseguir los objetivos.

* 1. **ECONOMÍA BANCARIA**

El efecto a largo plazo de la relajación actual de varias regulaciones bancarias, por ejemplo, relacionadas con diferir créditos a largo plazo, es una consideración clave para ofrecer los servicios bancarios. Observamos que la mayoría de los reguladores han seguido el mismo enfoque, con medidas muy similares. En el corto plazo, consideramos que esta relajación es positiva para los bancos, ya que debería permitirles navegar por lo peor de la crisis y al mismo tiempo proporcionarles más flexibilidad para gestionar los efectos inmediatos.

Sin embargo, es demasiado pronto para predecir si algunos de estos cambios se volverán más duraderos. De ser así, un debilitamiento permanente en los objetivos de capital y liquidez de los bancos, o una menor transparencia en el reconocimiento de la deuda incobrable y retrasos en la provisión adecuada de la misma, podría conducir a balances más débiles.

* 1. **Determinación de los objetivos comerciales**
     1. **Compilación de la información de la empresa**

Los datos están relacionados con campañas de marketing directo de una entidad bancaria portuguesa. Las campañas de marketing se basaban en llamadas telefónicas. A menudo, se requería más de un contacto con el mismo cliente para acceder a si el producto (depósito bancario a plazo) sería (o no) suscrito.

* + 1. **Describir el área problemática**

La entidad bancaria portuguesa recopilo una gran cantidad de datos que incluyen perfiles de aquellos clientes que tienen que suscribirse a depósitos a plazo y aquellos que no se suscribieron, por lo tanto, al no conocer dichos perfiles, no saben cuál han tenido más éxito, ya que esto permitiría la capacitación para definir acciones de gran valor como mejora del proceso de telemarketing.

* + 1. **Definición de los objetivos del negocio**

Para el desarrollo de este proyecto estableceremos los siguientes objetivos de negocio:

* Determinar cuáles son los clientes que suscribirá (sí / no) un depósito a plazo (variable y).
* Reducir los llamados tratando de contactar a la mayor cantidad de clientes proclives de adquirir un plazo fijo.
  + 1. **Valoración de la situación**

Para este proyecto se obtuvo la información brindada por la entidad bancaria portuguesa.

* **Datos:** La información recopilada es de todos los clientes que tienen que suscribirse a depósitos a plazo y aquellos que no. Básicamente, han sido dos conjuntos de datos en el primero se presentan 16 atributos, con una cantidad de estancias de 4521 y en el segundo presentan 16 atributos, con una cantidad de estancias de 45211 que son importantes para el desarrollo del proyecto.

Los datos se encuentran disponible en la información suplementaria del siguiente link: https://archive.ics.uci.edu/ml/datasets/Bank+Marketing

* **Factores de riesgo:** El proyecto puede tener falencias por el tiempo de desarrollo, puesto que no hay posibilidad de volver a realizar todo de nuevo una vez realizada la evaluación, entonces debe establecerse los pasos de la mejor manera posible para tener éxito.
  + 1. **Inventario de recursos**
       1. **Recursos del hardware**

Para este proyecto se cuenta con un equipo de cómputo con pocas características, puesto que se realizará un tratamiento de datos con baja necesidad de requerimientos, sin embargo, posee una capacidad de almacenamiento media ya que los datos tienen un volumen un tanto medio/bajo.

* + - 1. **Recurso del software**

El equipo de cómputo tiene un sistema operativo Windows 10 de 64 bits, a su vez para el procesamiento de los datos y el análisis exploratorio se cuenta con software de open source como lo son Python, anaconda, spyder(anaconda).

* + - 1. **Recurso humano**

Para este proyecto, se tiene a disposición dos personas con conocimientos afines a las tecnologías de información y un rendimiento del estudio de la minería de datos.

* + 1. **Requerimientos**
* Se requiere realizar un análisis exploratorio sobre los atributos suministrados en los datos de los clientes.
* Visualizar la información de forma óptima para que sea entendible a la vista de individuos que no poseen entendimiento de minería de datos.
  + 1. **Supuestos**
* Las herramientas para el análisis y la aplicación de los métodos deben funcionar en su totalidad.
* Los datos son otorgados por una fuente confiable, por consiguiente, se implica que la información ahí suministrada es veraz.
* Los datos posean una variable categórica que permitirá la predicción.
  + 1. **Restricciones**
* Las fases de la metodología tienen que cumplirse a cabalidad.
* La evaluación se realizará una vez terminada cada una de las etapas anteriores.
* Los instrumentos de investigación de datos tienen que proporcionar la información veraz.
* Los datos tienen que ser del todo procesados.
  1. **Orígenes de datos y almacenes de conocimientos**

Los datos a estudiar fueron encontrados en un repositorio, estos se encuentran disponible en el siguiente link <https://archive.ics.uci.edu/ml/datasets/Bank+Marketing>

* 1. **Terminología**
* **Dataset:** El data set contiene en sí, un conjunto de datos que han sido volcados desde el proveedor de datos.
* **Marketing:** analiza la gestión comercial de las empresas con el objetivo de captar, retener y fidelizar a los clientes a través de la satisfacción de sus necesidades.
  1. **Determinación de los objetivos de minería de datos**

Este proyecto tiene como finalidad aplicar metodologías y técnicas de minería de datos como lo son modelos predictivos y descriptivos, para lo cual se desarrollaron los siguientes objetivos.

* Construir un conjunto de datos que sea viable para la aplicación de técnicas de minería de datos
* Hacer un estudio exploratorio de los datos por medio de técnicas de visualización de datos
* Desarrollar un modelo predictivo para conocer cuáles son los clientes que suscribirá (sí / no) un depósito a plazo (variable y).

1. **FASE 2: COMPRENSIÓN DE LOS DATOS**

El desarrollo de la segunda fase de la metodología CRISP – DM se centra en encontrar la información a través de los datos, para introducirnos un poco más al problema real en sí, desde los datos presentados y los informes estadísticos que podamos obtener de este, a su vez, nos permitirá conocer la veracidad de los datos y la calidad que estos posean.

# Bibliografía

|  |  |
| --- | --- |
| [1] | E. A. F. SOTO, «Diseño e implementacion para un sistema recomendado de marketin bancario,» SANTIAGO, CHILE, 2019. |
| [2] | R. Muñoz, «MINERIA DE DATOS APLICADA AL MANEJO DE LA RELACION DEL CLIENTE EN UNA ENTIDAD BANCARIA,» QUITO, 2013. |
| [3] | M. P. L. R. D. CLIENTES, «Restrepo, Yeimy,» 2008. |
| [4] | L. Altamiranda, A. M. Peña, M. Ospino, I. Volpe y D. D. Ortega, «MINERÍA DE DATOS COMO HERRAMIENTA PARA EL DESARROLLO DE ESTRATEGIAS DE MERCADEO EN SECTORES PRODUCTIVOS, AFINES A LOS COLOMBIANOS: UNA REVISIÓN DE CASOS,» 2018. |
| [5] | C. Belman, «CLASIFICACIÓN DE CLIENTES DE LA INDUSTRIA BANCARIA POR MÉTODOS ESTADÍSTICOS Y REDES NEURONALES ARTIFICIALES USANDO R (CLASSIFICATION OF CUSTOMERS BELONGING TO BANKING INDUSTRY BY STATISTICAL METHODS AND ARTIFICIAL NEURAL NETWORKS USING R),» 2019. |
| [6] | F. Sánchez, «MARCO TEORICO PARA DIAGNÓSTICO Y PROPUESTA DE MEJORA EN LA RED DE AGENCIAS DE UNA EMPRESA BANCARIA DEL RUBRO DE LAS MICROFINANZAS EN LIMA METROPOLITANA EMPLEANDO HERRAMIENTAS DE MARKETING ANALYTICS Y DATA SCIENCE,» Lima, 2020. |
| [7] | D. López, «LA INFLUENCIA DE LAS TIC EN EL MARKETING BANCARIO.,» 2017. |