

Modelo Predictivo para la Fidelización de Clientes

Marini, Ian Denis / Sosa, Cristhian Sebastian / Sanchez Vottero Simon Azul

Innovación en Gestión de Datos - TSCDIA - 2024

Fecha de entrega: 06/09/24

Resumen del Proyecto

Descripción:

Este proyecto se adentra en el desafío de predecir el churn o abandono de clientes. A través del análisis de datos históricos sobre el comportamiento y las interacciones de los usuarios con los servicios de una empresa, se busca construir un modelo predictivo que identifique a aquellos clientes con mayor probabilidad de cancelar o dejar de utilizar los servicios permitiendo a la empresa tomar medidas proactivas para retener a estos clientes.

Objetivos:

El objetivo principal de este proyecto es desarrollar un modelo de clasificación que permita predecir con precisión la probabilidad de churn de cada cliente. Esto brindará a la empresa información valiosa para implementar estrategias de retención personalizadas y efectivas, mejorando la satisfacción del cliente y reduciendo las pérdidas asociadas al abandono. Además, el proyecto busca identificar las variables o características más relevantes que influyen en la decisión de los clientes de abandonar los servicios, lo que permitirá a la empresa comprender mejor las necesidades y expectativas de sus usuarios y adaptar sus estrategias en consecuencia.

Metodología a utilizar:

Se empleó un Diagrama de Gantt con el propósito de organizar las prioridades de las tareas y su asignación a días laborables específicos. Tales actividades se coordinaron a través de reuniones recurrentes (cada dos días) con el fin de monitorear el progreso de las mismas.

[illegible]

Relevancia del proyecto:

Mediante el análisis de datos históricos y la aplicación de técnicas de aprendizaje automático, este proyecto permite:

Identificar clientes en riesgo

Comprender las causas del abandono

Optimizar la inversión en retención

Mejorar la satisfacción del cliente

Aumentar la rentabilidad

Descripción de la Base de Datos

Descripción General:

Esta base de datos relacional está diseñada para almacenar y gestionar información sobre los clientes, sus llamadas telefónicas y planes a los que están suscritos.

Su objetivo principal es el análisis del comportamiento de los clientes (HISTÓRICO), permitiendo identificar patrones de uso, evaluar la escalabilidad de los planes y predecir posibles abandonos (CHURN).

1. Tabla Area

Descripción: Almacena información sobre las áreas geográficas, como el código de área y el estado asociado.

2. Tabla Planes

Descripción: Contiene información sobre los diferentes planes ofrecidos, como si incluyen llamadas internacionales o correo de voz, y la antigüedad requerida para acceder a ellos.

3. Tabla Clientes

Descripción: Almacena información detallada sobre los clientes, incluyendo sus datos personales, información de contacto, el plan al que están suscritos y su historial de abandono (churn).

4. Tabla Llamadas

Descripción: Registra el detalle de las llamadas realizadas por los clientes, incluyendo la duración y los cargos por diferentes tipos de llamadas (día, tarde, noche, internacionales) y el número de llamadas al servicio al cliente.

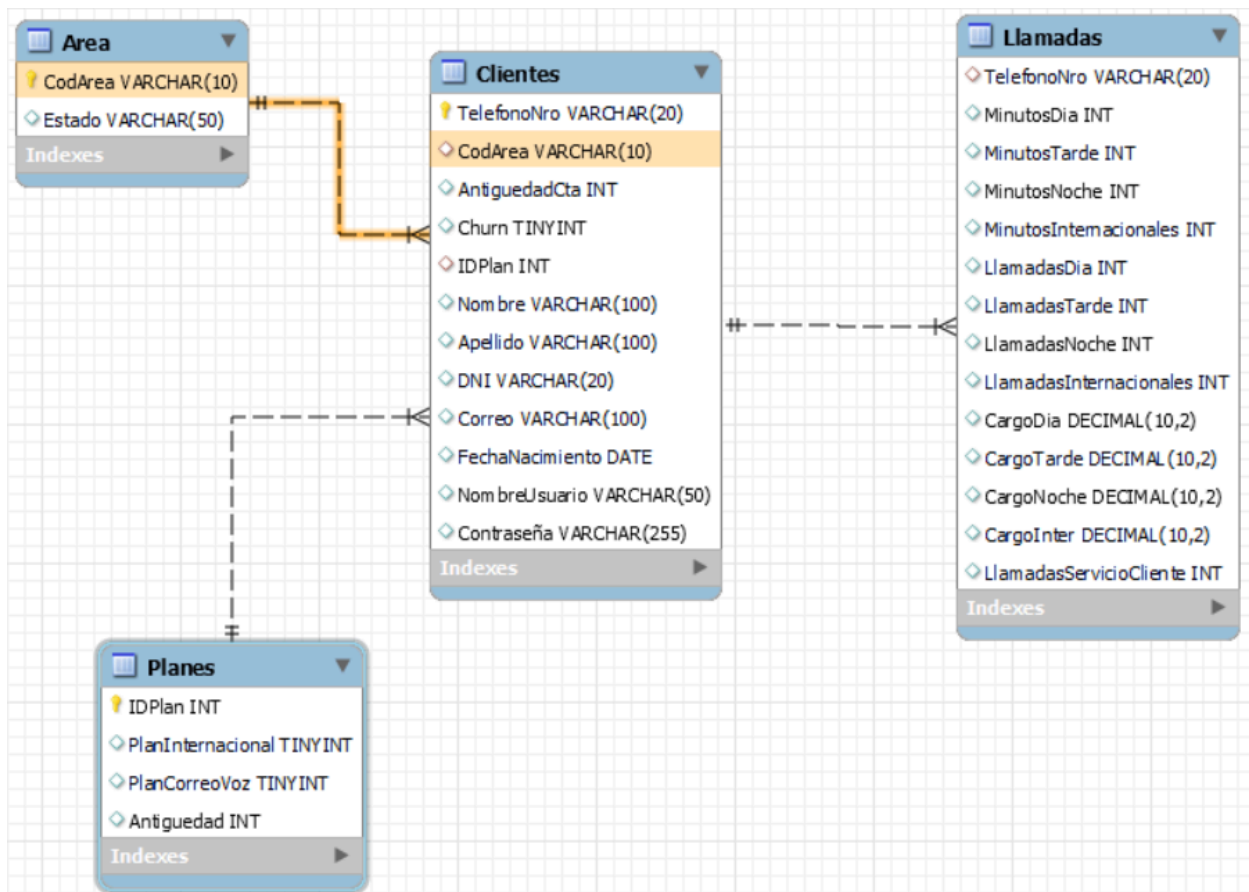
Relaciones:

Uno a Muchos (1:N):

Área - Clientes: Un área puede tener muchos clientes.

Planes - Clientes: Un plan puede ser contratado por muchos clientes.

Clientes - Llamadas: Un cliente puede realizar muchas llamadas.



Referencias:

Discovery hacía formato APA:

<https://normas-apa.org/formato/>

Requerimiento de distintas consultas hacia el Gestor MySQL

<https://dev.mysql.com/doc/#:~:text=Accelerate%20query%20performance%20with%20HeatWave%20MySQL.%20Query%20data>

Conclusión

Para llevar esta base de datos a la Tercera Forma Normal (3FN), primero debemos identificar las PK y las FK presentes en el set de datos. Luego, crearemos nuevas tablas para eliminar estas dependencias y establecer las relaciones adecuadas entre las mismas.

En atención a las peticiones efectuadas en el Módulo de Programación, se añaden campos a la tabla de Clientes donde se almacenará el historial de los mismos.

La base de datos está en 3FN, lo que minimiza la redundancia de datos y facilita su mantenimiento. La separación de la información de área geográfica en una tabla aparte permite actualizar o agregar nuevas áreas sin afectar los datos de los clientes.

La base de datos permite realizar análisis sobre el comportamiento de los clientes, como la duración de las llamadas, los cargos, el uso de planes internacionales y de correo de voz, y la tasa de abandono (churn).

Se analiza la posibilidad de incorporar una tabla de tiempo para almacenar información sobre fechas y horas, lo que facilita el análisis de tendencias a lo largo del tiempo.