**Introducción**

En el siguiente informe, consiste en el análisis de la variación de cancelación del servicio en la compañía Telecom X.

Para ello se realizará un análisis basado en predicciones para determinar los factores que mas influyen en la cancelación del servicio en la empresa Telecom X.

Dicho análisis será realizado con la herramienta de panda y librería de machine learning del lenguaje de programación Python.

**Datos y preparación**

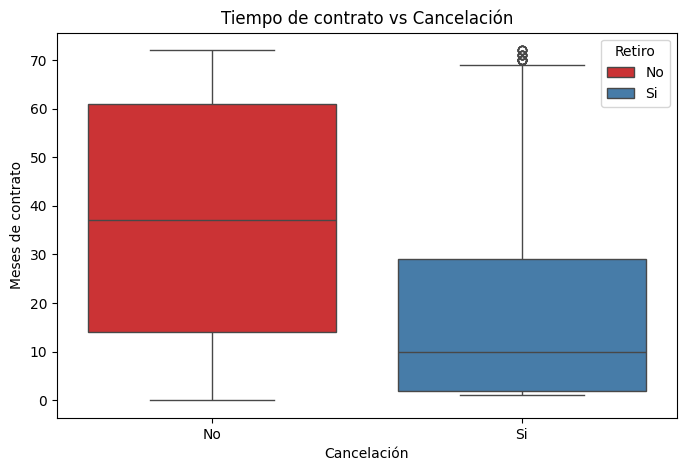
Inicialmente se realizó un bosquejo de los datos trabajados, donde se identificó la variable objetivo (Churn), sin embargo, se observó que dicha variable no era binaria, lo cual iba a causar inconvenientes al momento de realizar la aplicación de los modelos de machine learning, debido a esto, el paso siguiente fue cambiar esta variable a binaria (0 = No, 1 = Sí), y a la ves eliminar las variables únicas, las cuales no eran relevantes para el análisis.

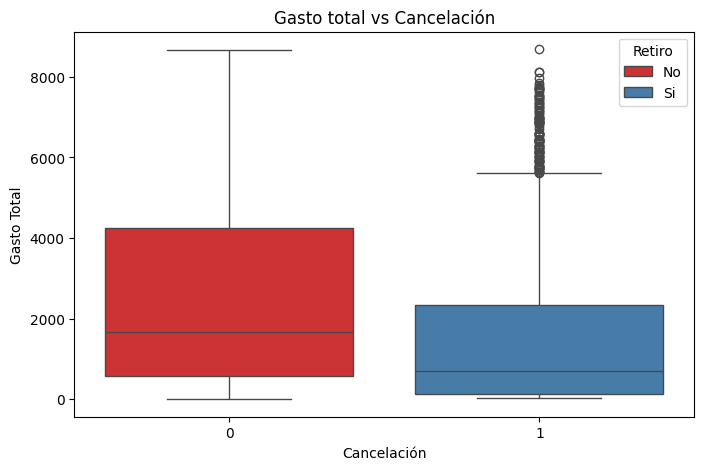
**División:** Train 70 / Test 30.

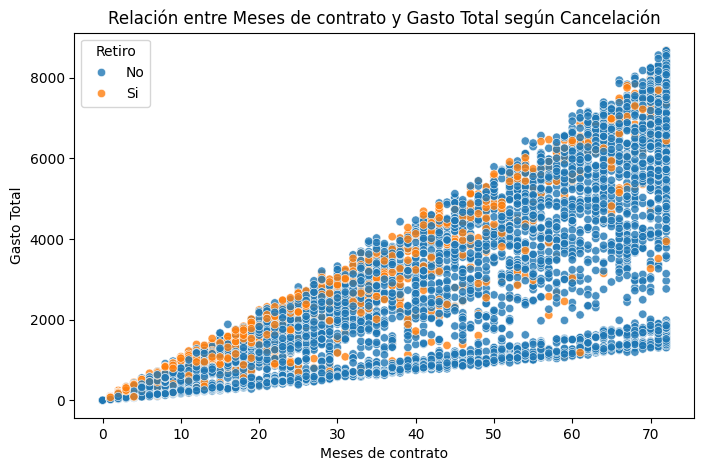
**Balanceo:** SMOTE

**Análisis de datos entre variables influyentes con las cancelaciones.**

Se realizaron varias comparaciones referentes a los tipos de contratos y meses de contrato en relación con la variable de cancelaciones (Churn), tal cual como se muestras en las siguientes imágenes:

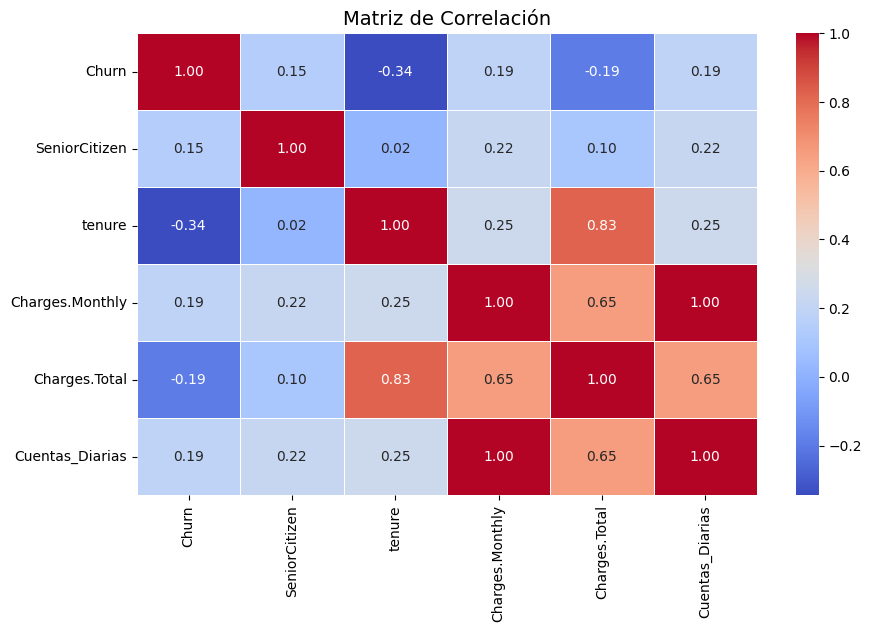






**Correlación y selección de variables**

En el estudio de las correcciones de las variables numérica, se obtuvo el siguiente grafico:



En el grafico se observa que la variable tenure tiene una negativa moderada (-0.34), lo cual indica que entre mayor es el tiempo de permanecía los clientes tienden menos a cancelar, otra de las variables que influyen en la cancelación es la de gastos mensuales, ya que si la factura es más alta los usuarios tienen a cancelar.

**Modelos predictivos**

Para este paso fue necesario la trasformación de los datos con el One\_Hot con el fin de dejar nuestras variables entendibles para los algoritmos de machine learning.

Inicialmente se escogió para nuestra predicción los modelos KNN y Árbol de decisión.

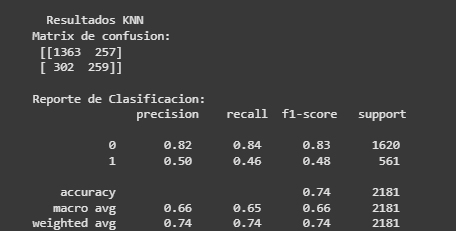
Sin embargo, se evaluaron los modelos:

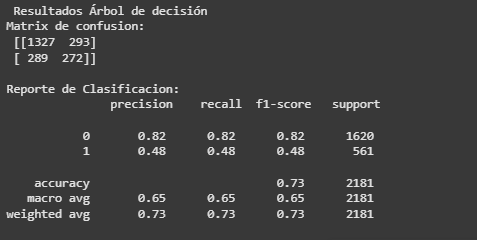
Modelo Baseline: dando un resultado del 74.28 % de personas no deciden cancelar el servicio.

Modelo de árbol de decision: 73.31 % de personas no deciden cancelar el servicio.

Modelo KNN: 74.37 % de personas no deciden cancelar el servicio.

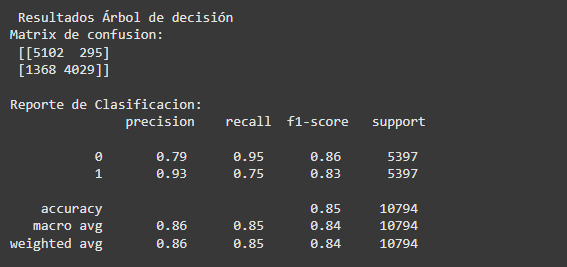
Cuando se realizaron la matrix de confusión con los datos desbalanceado los modeles me arrojaron lo siguiente:

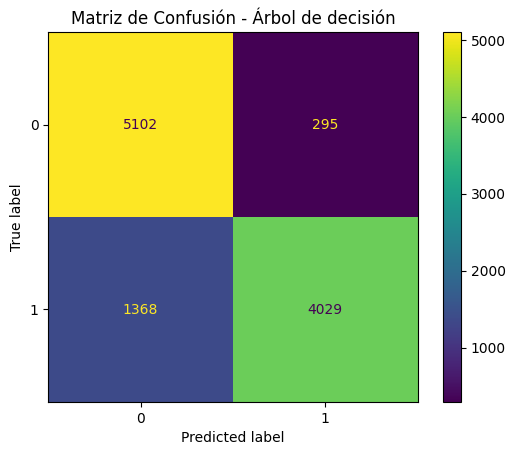


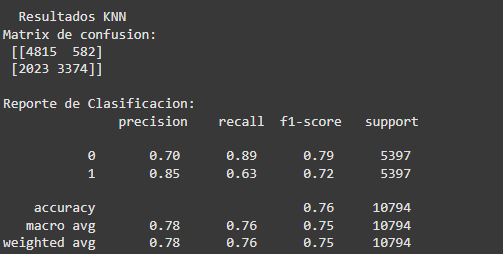


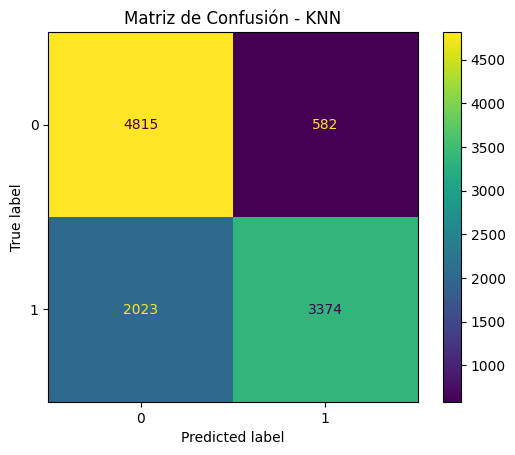
Donde se observa que cuando la variable de cancelación es 0 (No), los modelos tienden a comportarse correctamente, identificando en el recall a la mayoría de cliente, sin embargo, están sesgado para cuando la variable toma valor de 1 (Si), pues es muy poco la identificación de los clientes.

Debido a lo anterior se aplicó la técnica de desbalanceo SMOTE, obteniendo los siguientes resultados:









Teniendo en cuenta los resultados de estos dos modelos, se observa que el arbol de decisión tiene un mayor rendimiento al problema de la cancelación de servicios.

Sin embargo, también se probaron los modelos regresión logística, Random Forest y Modelo SVM, obteniendo las variables de importancia, al realizar la comparación se obtuvo lo siguiente:

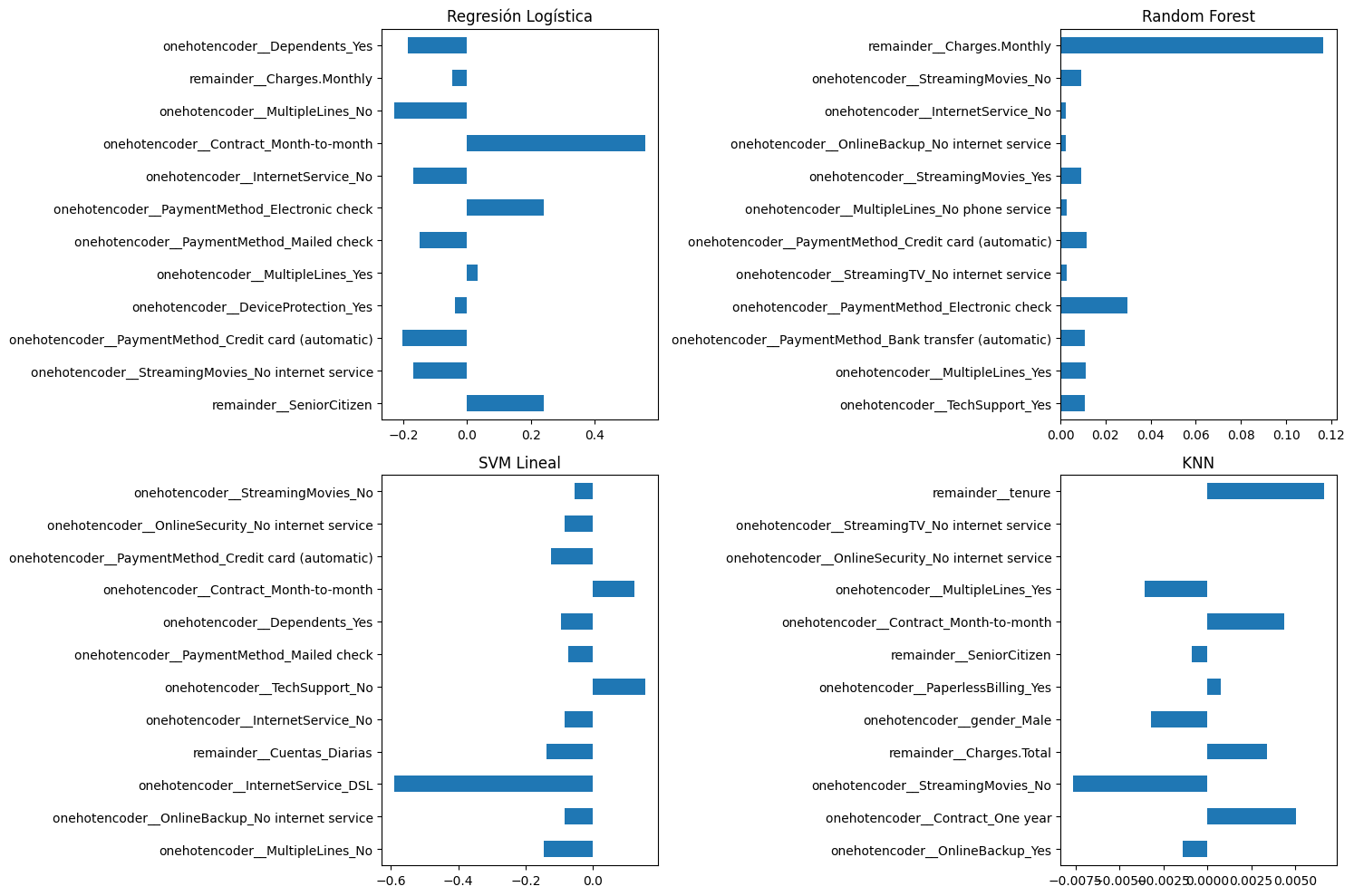
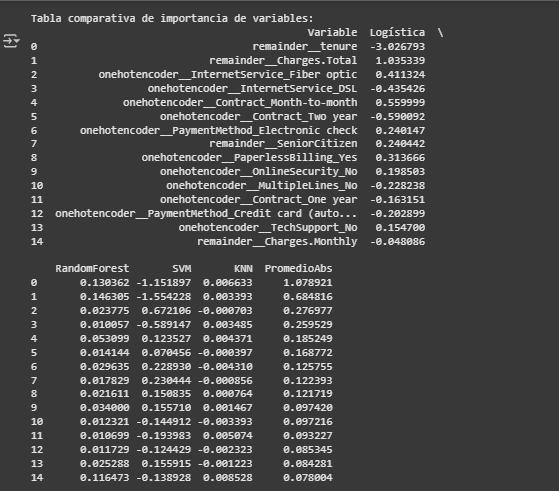


Tabla de comparación de variables importantes



**Conclusión**

Teniendo en cuenta la información suministrada por los modelos, se observa lo siguiente:

1. Entre menos sea la duración del contrato de un cliente, tiende a cancelar mas rápido.
2. Clientes con facturación más alta tiende a cancelar.
3. Personas mayores tienden a cancelar con más facilidad.

Teniendo en cuenta lo anterior, se recomienda lo siguiente:

1. Incentivar a los clientes con planes promocionales si se pasan de un plan mes a mes a uno de 12 meses, ofreciendo incentivos en mayor velocidad de internet entre otros.
2. Crear programas de incentivos por diferidos, y cargar promociones que entre mas servicios tenga activo el cliente menos el valor a cancelar.
3. Crear programas de acompañamiento para los mayores de edad, con el fin que puedan tener una experiencia inigualable en su servicio.