

# Informe Estratégico: Análisis Predictivo de la Cancelación de Clientes

**Preparado por:** Maria Fernanda Rodriguez - Equipo Machine Learning

## 1. Resumen Ejecutivo

El objetivo de este proyecto fue identificar los factores clave que influyen en la cancelación de clientes y desarrollar modelos de Machine Learning para predecir qué clientes tienen una alta probabilidad de cancelación de servicios. El análisis se centró en la evaluación y optimización de tres modelos: Regresión Logística, Random Forest y XGBoost.

El modelo Random Forest, fue seleccionado como la mejor herramienta predictiva y recomendada para implementar una estrategia de retención proactiva y eficiente porque tiene un buen equilibrio entre la capacidad de identificar a los clientes con probabilidad de cancelar (**Recall de 0.719**) y la fiabilidad de sus predicciones (**Precision de 0.555**), por lo tanto es más recomendado para implementar una estrategia de retención proactiva y eficiente.

## 2. Factores Clave que Impulsan la Cancelación

El análisis de la importancia de las variables en los modelos ha identificado los siguientes factores como los más influyentes en la cancelación:

- **Tipo de Contrato:**  
Este es el factor de riesgo más crítico. Los clientes con contratos mensuales (mes a mes) son el segmento de mayor riesgo, ya que no tienen un compromiso a largo plazo.
- **Antigüedad del Cliente**  
Los clientes con menos meses de servicio son los más propensos a cancelar.
- **Servicio de Internet: Fibra Óptica**  
El servicio de internet de fibra óptica está fuertemente correlacionado con una alta probabilidad de cancelación, tal vez pueda haber insatisfacción con la calidad del servicio o alto costo comparado por el mismo servicio con la competencia.
- **Pagos y Facturación:**  
El cargo mensual alto y el uso de Cheque Electrónico también se asocian con un mayor riesgo de cancelación.

### 3. Metodología de Modelado y Resultados

Se implementó un pipeline completo de Machine Learning, incluyendo el manejo del desbalance de clases con SMOTE, para garantizar un rendimiento óptimo.

Resultados obtenidos de la evaluación de los modelos:

Modelo	Accuracy (Exactitud)	Precision	Recall (Sensibilidad)	F1-Score	ROC-AUC
<b>Regresión Logística</b>	0.744	0.512	0.805	0.626	0.843
<b>Random Forest</b>	0.772	0.555	0.719	0.626	0.841
<b>XGBoost</b>	0.770	0.569	0.563	0.566	0.814

El modelo de **Random Forest** fue seleccionado como el más adecuado para este problema debido a su excelente equilibrio entre Precisión y Sensibilidad, adecuado para una estrategia de retención eficiente y efectiva.

### 4. Estrategias de Retención

De acuerdo a los resultados obtenidos, se proponen las siguientes estrategias de retención y planes de acción que permiten enfocar los recursos donde más se necesitan:

- Programa de fidelización para clientes de alto riesgo:**
  - **Plan de acción:** Utilizar el modelo de **Random Forest** para identificar a los clientes con contratos mes a mes en sus primeros 3-6 meses de servicio.
  - **Estrategia:** Ofrecer incentivos personalizados como descuentos o un upgrade de plan gratuito a cambio de un compromiso a largo plazo.
- Mejora del servicio de Fibra Óptica:**
  - **Plan de acción:** Realizar una investigación a fondo para identificar y resolver las causas de la insatisfacción en los clientes de fibra óptica.
  - **Estrategia:** Reforzar el equipo de soporte técnico y crear un sistema de monitoreo proactivo para detectar y solucionar problemas antes de que el cliente se queje.
- Incentivos para métodos de pago más estables:**
  - **Plan de acción:** A los clientes que pagan con **Cheque Electrónico**, ofrecer un pequeño beneficio para que se cambien a un método de pago automático.