

1. Problema

La empresa TelecomX enfrenta una alta tasa de cancelaciones de clientes (aproximadamente 27%). Este fenómeno representa una amenaza significativa para la estabilidad económica y la rentabilidad a largo plazo. El propósito de este análisis es identificar patrones y factores asociados con la pérdida de clientes, para apoyar el desarrollo de modelos predictivos y estrategias de retención.

2. Metodología

Se utilizó Python con las bibliotecas pandas, numpy, seaborn y plotly. El proceso se estructuró bajo el enfoque ETL:

- Extracción de datos desde una API pública (JSON).
- Transformación de datos: limpieza, detección y tratamiento de valores nulos, conversión de tipos de datos.
- Carga y análisis exploratorio con visualizaciones descriptivas.

3. Principales Hallazgos

- Los contratos mensuales presentan mayor propensión a cancelaciones.
- Clientes con cargos mensuales altos (mayores a \$80) tienden a cancelar con más frecuencia.
- La permanencia (tenure) menor a 12 meses se asocia con mayor churn.
- El método de pago electrónico está correlacionado con altas tasas de cancelación.
- Género y presencia de dependientes no tienen impacto significativo.

4. Recomendaciones para Modelado Predictivo

- Convertir variables categóricas mediante codificación dummy.
- Aplicar algoritmos como Regresión Logística, Random Forest y XGBoost.

- Utilizar técnicas de balanceo como SMOTE o undersampling.
- Explorar importancia de variables con herramientas como SHAP.

5. Recomendaciones Estratégicas para el Negocio

- Implementar programas de fidelización enfocados en usuarios con contratos mensuales.
- Ofrecer descuentos u ofertas para migrar a contratos anuales.
- Activar alertas tempranas para clientes con bajo tenure y alto cargo mensual.