**Nombre de la investigación:** Análisis predictivo de la deserción de clientes en el sector de telecomunicaciones utilizando técnicas de aprendizaje automático

**Base de datos a utilizar:** La base de datos “Telco Customer Churn” fue descargada de la plataforma Kaggle. Este dataset corresponde a una empresa de telecomunicaciones y contiene información detallada sobre los clientes y sus servicios contratados. Está compuesta por variables que permiten realizar un análisis integral del comportamiento del cliente y su propensión a abandonar el servicio (churn). Dentro de este dataset, se encuentra la siguiente información:

|  |  |  |
| --- | --- | --- |
| **Variable** | **Descripción** | **Tipo de datos** |
| Customer ID | Identificador único de cliente | Categórico |
| Gender | Género del cliente (masculino / femenino) | Categórico |
| Senior Citizen | Indica si el cliente es adulto mayor (sí = 1, no = 0) | Numérico |
| Partner | Indica si el cliente tiene pareja (sí / no) | Categórico |
| Dependants | Indica si el cliente tiene dependientes (sí / no) | Categórico |
| Tenure | Cantidad de meses que el cliente ha permanecido con la compañía | Numérico |
| Phone Service | Indica si el cliente tiene servicio telefónico (sí / no) | Categórico |
| Multiple Lines | Indica si el cliente tiene múltiples líneas telefónicas (sí / no / sin servicio de internet) | Categórico |
| Internet Service | Proveedor de servicio de internet del cliente (DSL / Fiber optic / no) | Categórico |
| Online Security | Indica si el cliente contrata servicio de seguridad en línea (sí / no / sin servicio de internet) | Categórico |
| Online Backup | Indica si el cliente contrata servicio de respaldo en línea (sí / no / sin servicio de internet) | Categórico |
| Device Protection | Indica si el cliente tiene protección de dispositivos (sí / no / sin servicio de internet) | Categórico |
| Tech Support | Indica si el cliente tiene soporte técnico (sí / no / sin servicio de internet) | Categórico |
| Streaming TV | Indica si el cliente tiene streaming de TV (sí / no / sin servicio de internet) | Categórico |
| Streaming Movies | Indica si el cliente tiene streaming de películas (sí / no / sin servicio de internet) | Categórico |
| Contract | Tipo de contrato del cliente (mensual / anual / bianual) | Categórico |
| Paperless Billing | Indica si el cliente utiliza facturación sin papel (sí / no) | Categórico |
| Payment Method | Método de pago del cliente (cheque, transferencia bancaria, tarjeta de crédito) | Categórico |
| Monthly Charges | Monto cobrado al cliente de manera mensual | Numérico |
| Total Charges | Monto total cobrado al cliente | Numérico |
| Churn | Indica si el cliente abandonó el servicio (sí / no) | Categórico |

**Motivo de la selección de los datos:** La selección del dataset de deserción de clientes en telecomunicaciones responde a la relevancia práctica del fenómeno de “churn” en la industria de servicios. Este conjunto de datos contiene información que permite analizar los factores que influyen en la decisión de abandonar el servicio. Además, el dataset está bien estructurado, es de acceso público y ha sido ampliamente utilizado en la literatura académica y en proyectos de ciencia de datos, facilitando la comparación de resultados y el acceso a metodologías validadas.

**Introducción del contexto:** En el competitivo mercado de las telecomunicaciones, la retención de clientes es un objetivo estratégico fundamental. La alta oferta de servicios y la facilidad de migración entre proveedores incrementan la tasa de deserción, lo que afecta directamente la rentabilidad y sostenibilidad de las empresas del sector. Identificar oportunamente a los clientes con mayor propensión a abandonar el servicio permite a las compañías diseñar campañas y políticas personalizadas de fidelización, reduciendo costos asociados a la adquisición de nuevos usuarios y fortaleciendo la relación con su base actual de clientes. El propósito de esta investigación es anticipar el comportamiento de deserción y aportar valor a la toma de decisiones.

**Línea de tiempo de investigación:**

|  |  |
| --- | --- |
| **Semana** | **Tarea** |
| 16 de junio | Selección de la base de datos y análisis de literatura |
| 23 de junio | Análisis exploratorio inicial de los datos, revisión y limpieza de datos |
| 30 de junio | Preprocesamiento y análisis descriptivo |
| 7 de julio | Evaluación del modelo predictivo más adecuado |
| 14 de julio | Redacción del reporte y preparación de la presentación |
| 21 de julio | Entrega final del proyecto |