# Customer Churn Prediction & Driver Analysis

---

## Introduction

For any subscription-based company, understanding \*\*why customers leave (churn)\*\* is paramount to future revenue and growth. This project addresses that critical challenge by developing a robust predictive model for customer churn within a telecom context. My goal was to not only accurately forecast churn but also to identify the \*\*key factors driving customer retention and attrition\*\*, enabling the development of targeted intervention strategies.

The analysis utilizes the "Customer Churn Prediction: Analysis," a synthetic dataset available on Kaggle.

---

## Methodology

The project workflow encompassed several key stages:

1. \*\*Data Preparation:\*\* Initial steps involved addressing missing values, converting data types, and ensuring overall data cleanliness.

2. \*\*Feature Engineering:\*\* This crucial phase involved transforming existing data into more insightful features. This included:

\* \*\*Grouping continuous variables:\*\* Categorizing customer age (e.g., into '25-39', '40-54' etc.) and tenure (e.g., 'under 1 year', '1-3 years') into distinct groups.

\* \*\*Creating composite features:\*\* Developing variables like `Services\_Count` (total core services subscribed to) and an \*\*interaction term\*\* (e.g., to understand how customer tenure specifically impacts those on a month-to-month contract).

3. \*\*Model Development:\*\* A \*\*Logistic Regression\*\* model was built for churn prediction, chosen for its strong interpretability.

4. \*\*Feature Importance Analysis:\*\* Post-modeling, the model's coefficients were analyzed to quantitatively identify the most influential drivers of churn and retention.

---

## Outcome

The developed Logistic Regression model proved to be highly effective, achieving \*\*98.5% overall accuracy\*\*. Its performance in identifying churners was exceptional, with \*\*100% precision\*\* (meaning nearly all predicted churners actually churned) and \*\*98% recall\*\* (correctly identifying almost all actual churners).

Analysis of the model's feature importance revealed the following dominant factors:

### Key Features Driving Customer Churn:

\* \*\*Month-to-Month Contracts:\*\* Customers on these contracts are highly prone to churn. This risk intensifies the longer they remain on this flexible plan, indicating a compounding effect of contract type and tenure.

\* \*\*High Monthly Charges:\*\* A strong indicator of churn, suggesting potential price sensitivity or a perceived lack of value for the services received.

\* \*\*Lack of Tech Support:\*\* Customers without technical support services are significantly more likely to churn, underscoring the importance of robust customer assistance.

\* \*\*New Customers (Under 1 Year Tenure):\*\* This group exhibits a higher inherent churn risk, highlighting the critical need for proactive engagement during the initial customer lifecycle phase.

### Key Features Driving Customer Retention:

\* \*\*Tech Support Subscriptions:\*\* Customers actively subscribing to technical support demonstrate significantly higher retention rates.

\* \*\*Longer-Term Contracts (One-Year/Two-Year):\*\* These contract types provide substantial stability and effectively reduce churn compared to month-to-month options.

\* \*\*Mid-to-Long Term Tenure:\*\* Customers who have stayed for 1-3 years or over 3 years exhibit strong loyalty, becoming less likely to churn over time.

---

## Additional Finding: The Nuance of Internet Service

Beyond the primary churn drivers, an intriguing pattern emerged concerning \*\*Internet Service Type\*\*.

While customers who \*only\* subscribe to our phone service (i.e., no internet service through us) show a higher propensity to stay, customers \*with\* any form of our internet service (including both Fiber Optic and DSL) exhibit a minor increased churn risk. This counter-intuitive finding is significant: it suggests customers might be finding more reliable or cheaper internet options elsewhere, making this a critical area for competitive analysis and re-evaluation of our internet service's value proposition.

---

## Recommendations

Based on these insights, I would advise the telecom company to:

\* \*\*Target Month-to-Month Customers:\*\* Implement proactive retention campaigns for month-to-month customers, especially as they approach the 6-12 month mark. Highlight long-term contract benefits and exclusive incentives.

\* \*\*Enhance Tech Support & Contract Promotion:\*\* More aggressively promote the value and benefits of Tech Support subscriptions and longer-term contract options (One-Year, Two-Year).

\* \*\*Investigate Internet Service Competitiveness:\*\* Conduct a deeper analysis into customer satisfaction with current internet services and competitor pricing/reliability. If internet service is indeed a churn driver, it likely points to external market pressures or perceived service gaps.

---

## Technologies Used

\* \*\*Python:\*\* Programming Language

\* \*\*Pandas:\*\* Data manipulation and analysis

\* \*\*NumPy:\*\* Numerical computing

\* \*\*Scikit-learn:\*\* Machine learning model development (Logistic Regression, preprocessing)

\* \*\*Matplotlib:\*\* Data visualization

\* \*\*Seaborn:\*\* Enhanced data visualization