Can We Predict Churn Rate?: The Lifeline of Startups and Their Customers

Capstone Project: Sprint0

Hal Morishima

The Problem area

- Startups, particularly app-based businesses operating on subscription models like music streaming services - the source of our dataset for this project - often face customer churn as one of their most pressing challenges. The ability to predict customer churn is paramount considering that retaining existing customers usually comes at a lower cost than acquiring new ones.

- The early identification of churn risk customers provides startups with the opportunity to proactively engage with potentially departing customers, thus improving retention. This key metric is especially critical for startups seeking funding, as investors often scrutinize churn rates to evaluate the startup's viability and growth potential.

The User

- The primary users of a churn prediction model would be the customer service and marketing departments of a company. These teams could use the model's insights to design targeted interventions for customers predicted to churn, improving overall customer satisfaction and reducing churn rates.

The Big Idea

- The main idea is to apply machine learning techniques to predict KKBox's customer churn. By using available customer data such as user demographics, subscription type, and usage behavior, the model can identify patterns associated with churn. This way, the model could predict which users are most likely to churn in the near future.

The Impact

- Predicting churn can have a significant impact on a company's bottom line. By identifying at-risk customers early, companies can take proactive measures to improve customer retention.

- This could involve offering incentives, personalizing the user experience, or addressing any issues the customer might be facing. In the long term, a reduction in churn rates could lead to increased customer loyalty, higher customer lifetime value, and ultimately, improved profitability.

The Data

- We're using a dataset from Kaggle's WSDM - KKBox's Churn Prediction Challenge. It includes about 9GB datuser details, subscription info, and user logs showing how users interact with the platform.

- We aim to use this data to guess if a user will leave the service after their subscription ends. A special part of this dataset is the churn label, which is based on KKBox's own rules.