Problem Statement Worksheet (Hypothesis Formation)

What opportunities exist for KKBox to report a positive percent change in revenue by the end of the current quarter through subscriber retention, attracting new subscribers, or increasing prices?

H

1 Context

KKBox is a music streaming service popular in South East Asia with over 10 million users. It functions on a subscription-based business model, with the majority of subscriptions lasting 30 days. An account is marked as churn if there are no new transactions within 30 days after a subscription has expired. KKBox would like to be able to predict which subscribers are likely to renew within a month of their membership ending and which ones will churn.

2 Criteria for success

This project will be deemed successful if there is a statistically significant decrease in churn one month from the implementation of the solution. If there is no decrease in churn, or if the decrease is statistically insignificant, additional tuning of model hyperparameters and retesting is recommended before overhauling the solution completely.

3 Scope of solution space

By using customers' demographic information, listening history, and transaction history, we will train a classification algorithm to predict if a particular user will renew their subscription or churn.

4 Constraints within solution space

Our data is limited to users' demographic information, listening history, and transaction history. Other factors that may affect a customer's decision to no longer renew their membership, such as a change in income or transferring to a competitor service, cannot be considered by our model as we are not privy to this information.

5 Stakeholders to provide key insight

The Marketing department at KKBox will be a key partner and stakeholder in this project. Predicting which users are at risk of churn can only have an impact on revenue if those predictions are matched with quality advertisements that are able to convince those users to renew their subscriptions.

6 Key data sources

We will use the following CSV files provided by KKBox:

- **train_v2.csv,**
- **members_v3.csv,**
- **transactions_v2.csv, and**
- **user_logs_v2.csv.**

H D E I P