

A streaming service provider wants to reduce customer churn (users canceling their subscriptions). The company has historical data about its customers, including demographics, subscription details, usage behavior, and churn status (whether they canceled their subscription). They want to develop a machine learning model to predict which customers are likely to churn, so they can take preventive action

How would you handle imbalanced data if churned customers are fewer than active ones?

- ➔ I would go through the dataset properly and analyse if the given problematic data can be ignored then I will for sure drop it, if not I will try to fix the data and if I fail to fix I will just drop the extra data.

What features are the most important predictors of churn?

- ➔ According to my understanding in case of churn the main factor is ROI that implies features as SubscriptionPlan, tenure, monthlyfee.

How would you explain the model's predictions to a non-technical business team?

- ➔ I would state "This is a special algorithm that takes input as statistic of an individual(feature) it analyses those statistics and predicts whether that individual will churn or not.

What steps would you take to deploy this model into production?

- ➔
 1. Load the dataset
 2. Analyse the data
 3. Clean the data
 4. Pre-process it
 5. Experiment with multiple classification model
 6. Evaluate the model
 7. Hypertune the model
 8. Evaluate the model.