Streaming Services Case Study

Content

1. Assumptions
2. Analysis Framework of Churn Issue
3. Approach to Churn Prediction Model

Assumptions

- All relevant data are available.

- Available data are complete and, in the format required for all necessary analysis and modelling.

- Product and Customer Support teams refer to the relevant teams that has the knowledge and understanding of customer behaviour for the respective markets, and that these teams exists.

- Limited or minimal computational restrictions on Churn Prediction Model.

Analysis Framework of Churn Issue

Define | Hypothesize | Analyze | Recommend & Perform

Churn as an input to Business Decisions

Leveraging market knowledge of customer behaviour, and harnessing insights from data.

Process	Description	Example
Churn Definition	• Identifying the appropriate Churn definition and Churn metric tracked influences scope of analysis.	• Churn refers to customers who cancels or downgrade plans voluntarily.
Hypothesis Generation	• Gather list of hypothesis that might explain the churn Customer Behaviour Workshops • Conduct workshops with Product or Customer Support teams* by markets to develop understanding of customer behaviour Customer Segmentation Analysis • Identify churn data across different customer segments	• Teams hypothesize that (i) weekly pricing plan might be more receptive in Philippines due to weekly salary disbursement within the country (ii) bundling products to increase revenue and reduce churn. • Deploy K-means clustering for segmentation in each market. Within each cluster, perform usage and interaction analysis.
Data Validation & Analysis	• Validate hypothesis with available data Correlation Study • Identify correlation between customer actions and churn (e.g. Customers that Wishlist content has lesser churn) Feature Analysis • Determine which features are the most/least valued by customers. Might lead to feature improvement discoveries Lifecycle Analysis • Identify which point of lifecycle will customers tend to churn	• Based on list of hypothesis, perform correlation study to accept or reject hypothesis. • Depending on the type of data used, correlation coefficient used could either be Pearson or Spearman correlation coefficients. Check for multicollinearity and interpret correlation cautiously.
Recommendations & Performance	• Prioritize recommendations based on ROI and implementation duration • Define success metrics to measure performance	• Based on findings, introducing weekly pricing plans in Philippines with auto-renewal option has higher ROI and lower implementation duration. • Metrics to measure includes uptake in weekly plan, churn rate, and churn rate by different plans.

*Or equivalent teams with the understanding of customer behaviour

Approach to Churn Prediction Model

Aim | Engineer | Evaluate & Select | Interpret & Improve

Churn Prediction Model

Constant monitoring, refining, and experimenting to stay ahead of Churn.

Process*	Description	Example	Factors to Consider
Aim of Model	• Use same Churn definition from previous framework. • Determine aim of Churn Prediction to choose the appropriate underlying model.	• Model to predict if customer will churn or not. Model not intended to provide estimated time to churn.	• Consistent churn definition allows appropriate strategies to be deployed.
Feature Engineering	• Identify relevant features that influences churn. • Create new features if necessary.	• Perform Exploratory Data Analysis to identify potential features that might influence churn. • Start with a common model (e.g. Logistic Regression) with different combination of features. Each combination is considered 1 model.	• Consider using different techniques such as Principal Component Analysis, Decision Tree, Random Forest to further enhance selection of features.
Model Evaluation & Selection	• Split data into training and testing sets. • Evaluate model performance. • Select best performing model.	• 80% of data for training and 20% for testing. • Use metrics like Accuracy, Precision, and Recall to evaluate and compare each model's effectiveness. Metrics can also be prioritized. If "being correct" (e.g. correctly identifying churn) is more important, Precision metric will be prioritized.	• Beware of overfitting, and multicollinearity. • If different machine learning models are deployed, k-fold cross-validation can be used to better evaluate model performance. • Different models might be implemented across markets due to different influence of features.
Model Interpretation and Improvement	• Understand coefficient outputs to propose strategy to reduce churn and increase stickiness. • Deploy models and monitor performance of model.	• If the odds of churn reduces by 0.1, for every use of the app, the team can consider sending suggested content notifications to customers who use the app lesser than the average customer.	• Given the iterative nature of modelling process, constant monitoring and periodic review of performance can be conducted. • Model interpretability should be considered to ensure that the team understands the working mechanics to provide recommendations.

*Processes should be iterated frequently

End