

EdRetain: A Smart Churn Prediction and Predictive Retention Analytics System for Premium Users in EdTech Platforms

Identify the Problem Area

The problem area of this project is customer retention and segmentation within subscription-based EdTech platforms, with a particular focus on premium users. From a marketing perspective, the interest lies in understanding and managing the ongoing relationship with existing paying learners: keeping them engaged, maintaining their premium status, and strategically segmenting them based on their likelihood to stay, downgrade, or cancel.

Conduct Preliminary Research

Recent trends in the EdTech sector reveal a high churn rate among premium subscribers, driven by declining motivation, limited personalization, and the abundance of free alternatives. Studies and market analyses show that most learners abandon courses early or fail to perceive the long-term value of premium features. Competition is intense, and while companies invest heavily in customer acquisition, they often neglect the retention phase. Addressing this issue through predictive analytics offers an opportunity to identify at-risk customers before they churn, enabling timely and personalized marketing interventions that improve both learner outcomes and brand loyalty.

Define a Specific Problem

The specific, measurable problem is to determine which premium users are most likely to downgrade or churn and what behavioral indicators signal this risk early enough for marketers to act. The project aims to uncover patterns in engagement and spending that precede cancellation, helping marketing teams predict potential downgrades and design strategies to retain those customers.

Propose a Solution with Methodology

To solve this problem, the project proposes the development of a retention prediction model that continuously monitors learner activity and spending patterns to detect early warning

signs of disengagement. The model will assign a churn-risk score to each premium user and identify the key behavioral features that most strongly predict churn. These insights will inform personalized marketing campaigns designed to re-engage learners and maintain their subscription status.

Data Collection

The solution will rely on data that reflects premium customers' engagement, learning progress, and subscription behavior. In particular, the analysis will require information about how users interact with learning content, how active they are on the platform over time, how their subscriptions evolve, and how they respond to different marketing or communication efforts.

Analytical Techniques

To address the problem of identifying premium EdTech users at high risk of downgrading or churning, the project will use a combination of predictive, behavioral, and segmentation-focused analytical methods from a marketing perspective:

- **RFM Analysis** to create an initial behavioral view of premium learners based on recency and frequency of activity and their overall value, providing a simple way to distinguish more and less engaged users.
- **Customer Segmentation** using clustering techniques such as **K-Means** to group learners into meaningful segments (e.g., highly engaged, at-risk, low-value) based on RFM scores and other engagement indicators, so that different retention strategies can be designed for each group.
- **Churn Prediction Modeling** with classification algorithms like **Logistic Regression** or tree-based models to estimate the probability that a premium user in each segment will downgrade or cancel within a given period.

- **Survival Analysis** to model how long premium customers are likely to remain subscribed before downgrading, adding a time dimension to churn risk and helping to identify critical windows for intervention.
- **Customer Lifetime Value (CLV)** estimation that combines information on revenue, churn probability, and expected subscription duration to quantify the long-term financial value of each premium user and to prioritize retention efforts where they have the greatest impact.

Implementation Plan

- **Real-Time Churn Scoring** by integrating the prediction model into the marketing platform to continuously assess risk levels of premium users
- **Personalized Retention Campaigns** targeting high-risk segments with customized content, special offers, and motivational messaging
- **Monitoring Dashboards** for marketing teams to track churn risks and campaign performance metrics in real-time
- **Targeted Interventions** executed based on the risk scores and segment classifications generated by the model
- **Continuous Feedback Loop** through ongoing data collection to refine models and improve campaign strategies

Expected Outcomes

- Increased customer loyalty and engagement.
- Higher average subscription duration, and overall customer lifetime value
- Higher recurring revenue
- More efficient, better-targeted marketing campaigns
- Strengthened brand reputation and competitive advantage in the EdTech market.

Evaluation Metrics

- Increase in subscription duration and customer lifetime value (CLV)
- Retention rate improvement
- Reduction in churn rate over time
- ROI of retention campaigns
- Campaign open and response rates