

Video Streaming Churn Project | Feature Engineering (Milestone 4)

Executive Summary Report

Prepared by: PhD Aleksandar Osmanli


Project Overview


I'm currently developing a data analytics project aimed at increasing overall growth by preventing monthly user churn on video streaming service. For the purposes of this project, churn quantifies the number of users who have cancelled the monthly subscription of the video streaming service. Feature engineering offer possibility to enhance the ML models and achieve better results by combining the most predictive basic features into new features, which can add predictive power to the modeling. **This report offers details and key insights from Milestone 4, which impact the future development of the overall project.**

Correlation of Features with Churn


Feature	Correlation
AccountAgeAverageViewingDuration	-0,211278
AccountAgeViewingHoursPerWeek	-0,201495
AccountAgeContentDownloadsPerMonth	-0,199352
AccountAge	-0,197736
AccountAge_Per_MonthlyCharges	-0,194331
ViewingHoursPerWeekAverageViewingDuration	-0,172806
AverageViewingDurationContentDownloadsPerMonth	-0,171814
TotalChargesAverageViewingDuration	-0,160681
ViewingHoursPerWeekContentDownloadsPerMonth	-0,160450
AverageViewingDuration_Per_MonthlyCharges	-0,157554
TotalChargesViewingHoursPerWeek	-0,152027
TotalChargesContentDownloadsPerMonth	-0,151333
AccountAge_Per_UserRating	-0,151181
AccountAgeUserRating	-0,146917
AverageViewingDuration	-0,146897
ContentDownloadsPerMonth_Per_MonthlyCharges	-0,145559
ViewingHoursPerWeek_Per_MonthlyCharges	-0,143736
ContentDownloadsPerMonth	-0,129752
ViewingHoursPerWeek	-0,128645
MonthlyCharges_Per_AverageViewingDuration	0,124072
MonthlyChargesSupportTicketsPerMonth	0,120682
TotalCharges	-0,120529
SupportTicketsPerMonth_Per_AverageViewingDuration	0,118895
AccountAge_Per_SupportTicketsPerMonth	-0,115397
AccountAgeWatchlistSize	-0,115164
AverageViewingDuration_Per_UserRating	-0,114473
MonthlyCharges_Per_ViewingHoursPerWeek	0,110414
MonthlyCharges_Per_AccountAge	0,110037
SupportTicketsPerMonth_Per_AccountAge	0,106711
SupportTicketsPerMonth_Per_ViewingHoursPerWeek	0,106020
ContentDownloadsPerMonth_Per_UserRating	-0,104855
AverageViewingDurationUserRating	-0,103852
ViewingHoursPerWeek_Per_UserRating	-0,102198
TotalCharges_Per_UserRating	-0,100723
MonthlyCharges	0,100473

Milestone 4 – Feature Engineering

 **Target Goal:** Create new interactive features and check their additional predictive power on user subscription churn.

 **Methods:**

- Assessment of churn risk among the most predictive basic features
- Creation of interactive features
- Evaluation of their correlation with churn

 **Impact:** By deep engineering, the new features can reveal important relationships and predict the target feature

Next steps

→ Taking into account the benefits of tree-based ML models, evaluate different types of models to assess the best roc_auc score on the validation dataset.