Predicting Content Fatigue: A Data-Driven Subscription Retention Strategy

Objective

Move from reactive churn management to proactive risk prediction.

Focus

Uncover hidden consumption patterns that lead to user fatigue and eventual churn.

The Churn Crisis: Fatigue is the Precursor to Loss

Raw transaction data (purchases/logins) often misses the critical signals of user disengagement.

> **Consumption Fatigue** is the slow decline in engagement that precedes cancellation. We must identify at-risk users early.

📝 The Business Impact: It is **5x more expensive** to acquire a new subscriber than to retain an existing one.
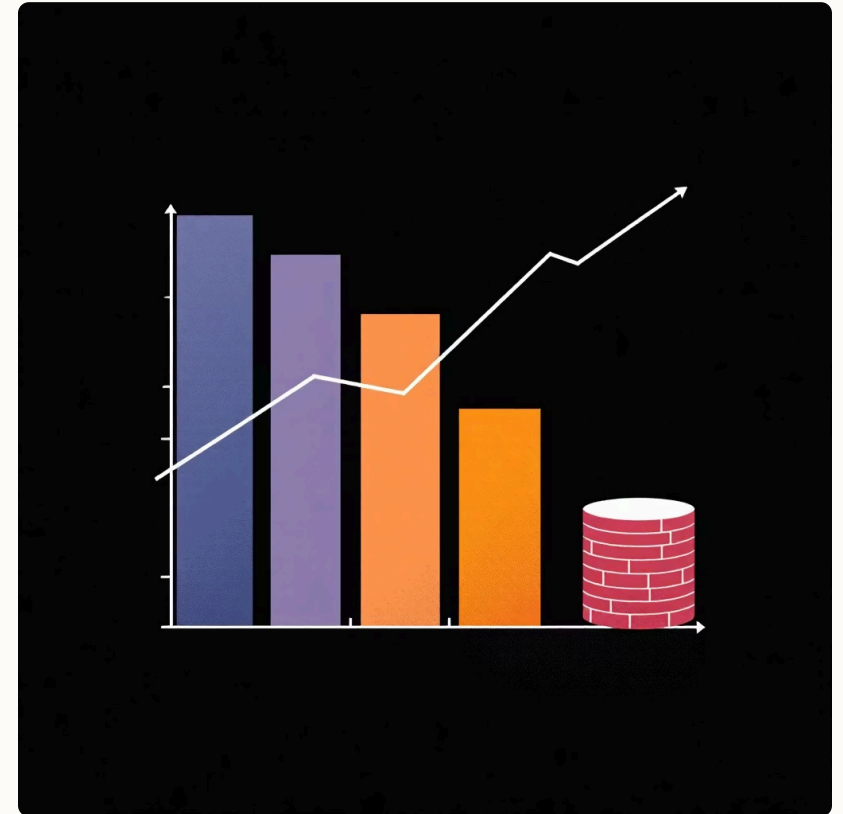
From Raw Data to Actionable Business Intelligence

Our solution pipeline transforms raw user ratings into a predictive model and deploys the results directly to management via a BI dashboard.
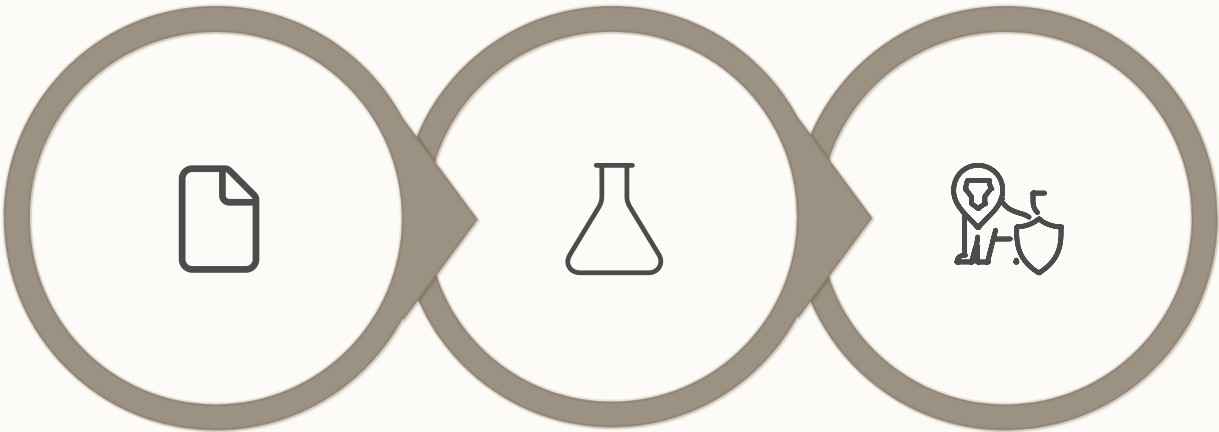
Feature Engineering

Modeling

BI Deployment

Feature Engineering (RapidMiner)

Convert □100,000 ratings into structured, analytical user profiles.

Modeling (Python)

Apply K-Means and Decision Trees to find behavioral groups and predict risk.

BI Deployment (Power BI)

Deliver predictive alerts and segmentation reports to management for action.

Creating the Analytical User Profile (Z-Scores)

We engineered three core features from the raw rating data, all normalized using Z-scores for fair comparison across the user base.

Average Rating (Satisfaction)

Measures how happy and satisfied the user is with the content library overall.

Total Movies Rated (Activity)

Quantifies the user's volume of consumption and engagement with the rating system.

Standard Deviation of Rating (Consistency)

Indicates how stable and predictable the user's taste is over time.

These three metrics form the basis for our K-Means clustering and predictive modeling.

Defining the Customer Universe: The 4 Behavioral Segments

Unsupervised K-Means Clustering (K=4) was used to define distinct behavioral segments based on their Satisfaction, Activity, and Consistency profiles.

Cluster 2: Happy Regular

High Satisfaction, High Consistency. Our benchmark for loyalty and retention.

Cluster 3: Volatile Viewer

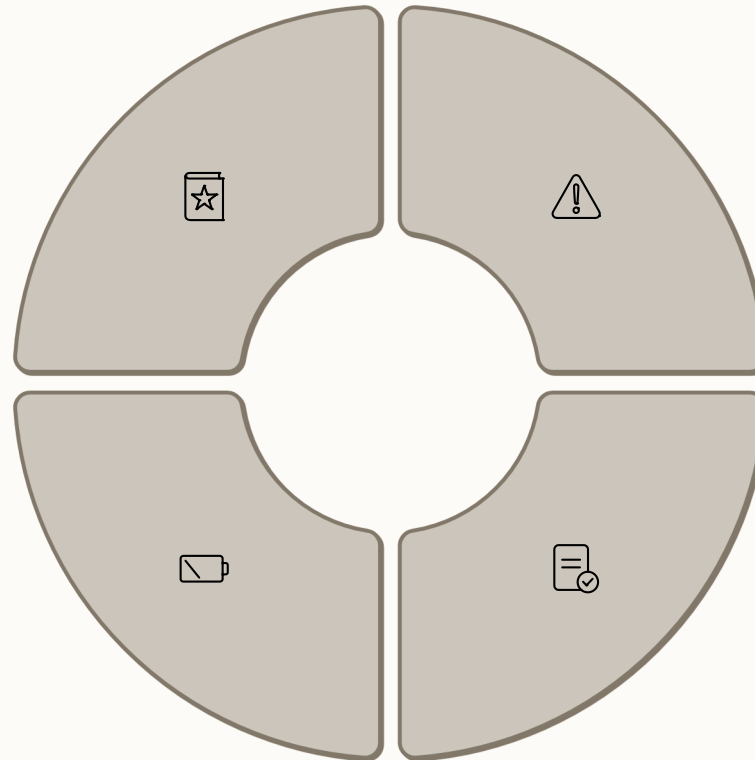Low Satisfaction, High Inconsistency. Our primary high-risk target for churn.

Cluster 4: Dormant Users

Very low activity across all metrics. Already near the point of cancellation.
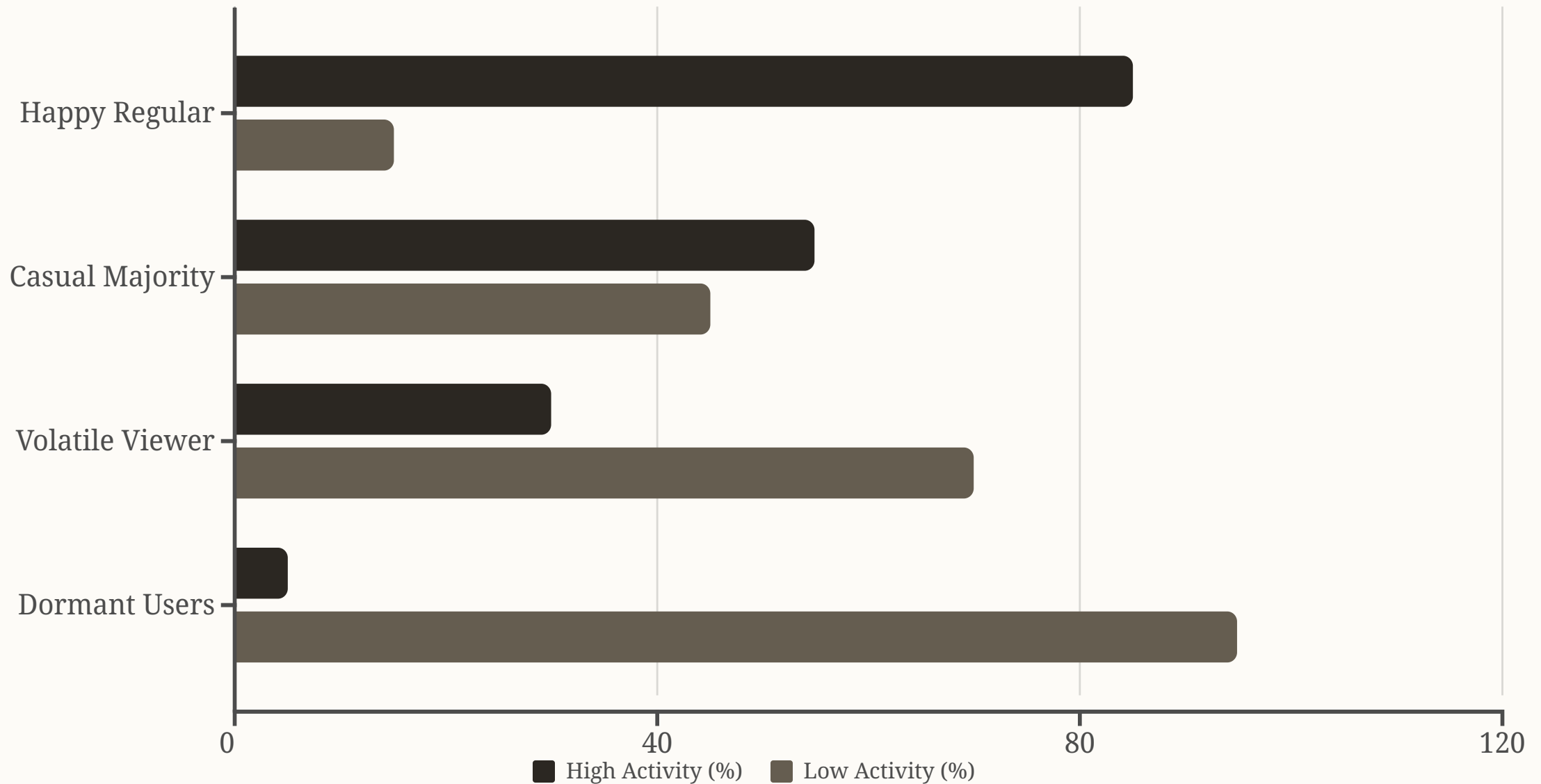
Cluster 1: Casual Majority

Moderate engagement and consistency. A large, stable group.

Segmentation vs. Activity: Identifying Low-Value Segments

By linking the descriptive segments to the predictive outcome (future activity), we gain a clearer picture of where the highest risk lies.

Legend: ■ High Activity (%) ■ Low Activity (%)

Insight: The 'Low Activity' users are disproportionately concentrated within the Volatile Viewer and Dormant clusters, confirming these as our highest priority targets for intervention.

Forecasting Risk: Decision Tree Performance

We used a Decision Tree Classifier to predict whether a user would be 'High Activity' or 'Low Activity' in the subsequent period.
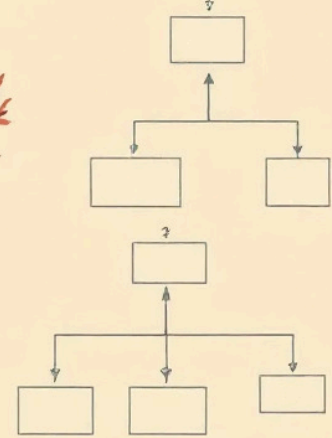
61.20%

Model Accuracy KPI

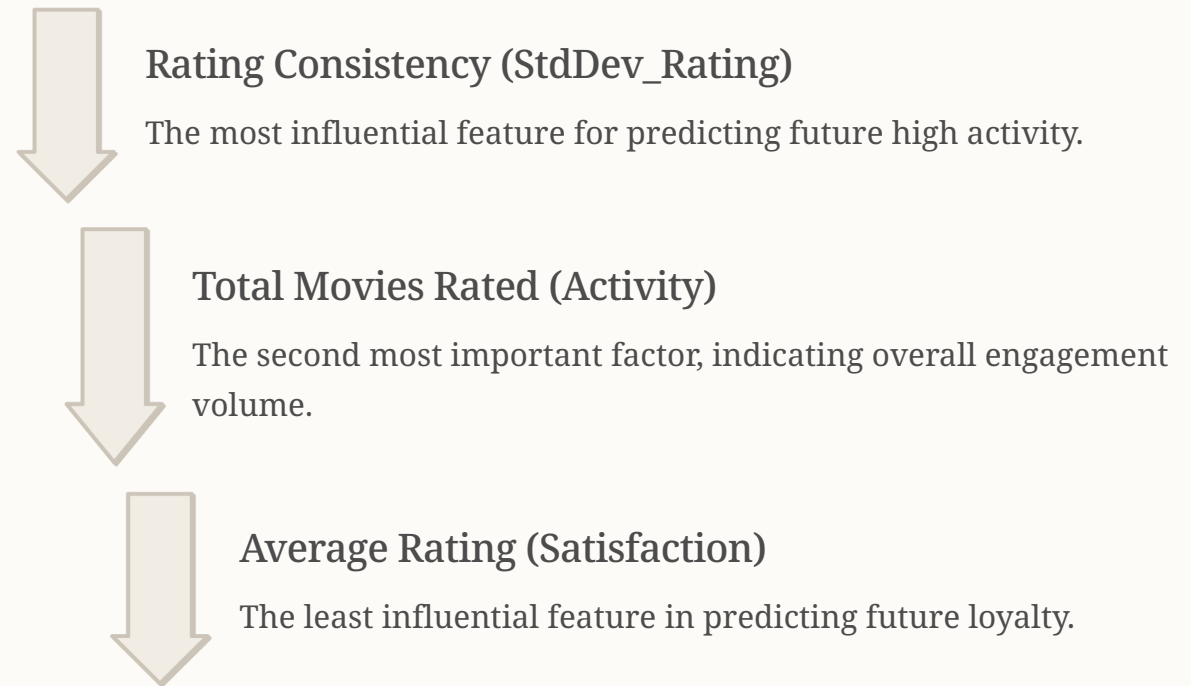The model is significantly better than random guessing at identifying future loyal users.

Critical Error Tracking: We must closely track False Negatives—the number of high-activity users we incorrectly dismiss as low-risk. Minimizing this error is key to maximizing retention efforts.

Predictive Insight: Rating Consistency Drives Loyalty

The Decision Tree analysis revealed a clear hierarchy of influence among our engineered features.

Rating Consistency (StdDev_Rating)

The most influential feature for predicting future high activity.

Total Movies Rated (Activity)

The second most important factor, indicating overall engagement volume.

Average Rating (Satisfaction)

The least influential feature in predicting future loyalty.

Interpretation: Predictable users are loyal users. Chaos in rating patterns leads to fatigue and risk, regardless of how high their average rating is.

Strategic Recommendations for Subscription Retention

Based on the predictive model and segmentation, we propose three actionable strategies targeting the high-risk Volatile Viewer segment.

Proactive Alerting

Immediately flag new users matching the Volatile Viewer profile for personalized curator-picked recommendations.
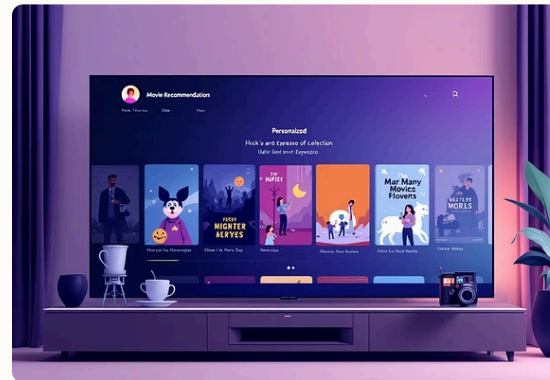
Incentivize Consistency

Run A/B tests rewarding structured, consistent feedback/ratings to shift users toward the 'Happy Regular' profile.

Optimize Acquisition

Focus marketing spend on channels that yield users with 'Happy Regular' characteristics from the start.

Project Value Delivered

61.20%
Predictive Accuracy
Accuracy of the Decision Tree model in forecasting future user activity.

5X
Retention ROI
Cost savings achieved by retaining users versus acquiring new ones.

1
Key Insight
Rating Consistency was identified as the single most influential feature for loyalty.

We successfully transformed raw data into a predictive tool, identifying the Volatile Viewer as the highest priority target. This system provides a framework for moving from **reactive fixes** to **proactive user retention**, maximizing subscriber lifetime value.