Viewer Retention in OTT Platforms: Diagnosing Engagement Patterns

Case Study

Capstone Project for Winter Consulting 2025

Consulting & Analytics Club, IIT Guwahati

RAJVI BUJAD

Decoding Viewer Drop-Off: A Data-Driven Approach

Problem Diagnosis & Analysis Framework

Objective

Identify the key drivers of viewer drop-off across 33,000+ episodes on our streaming platform.

Methodology

Utilizing correlation analysis, advanced clustering techniques, and predictive modeling for robust insights.

Key Findings

- Strong correlation (-0.96) between average watch percentage and drop-off probability.
- Pacing score and hook strength are the top predictors of viewer retention.

Our primary challenge: Addressing high drop-off rates in specific genres like Drama and Mystery, and content with high cognitive load.

Unpacking Viewer Behavior: Data-Driven Insights

Genre Impact on Retention

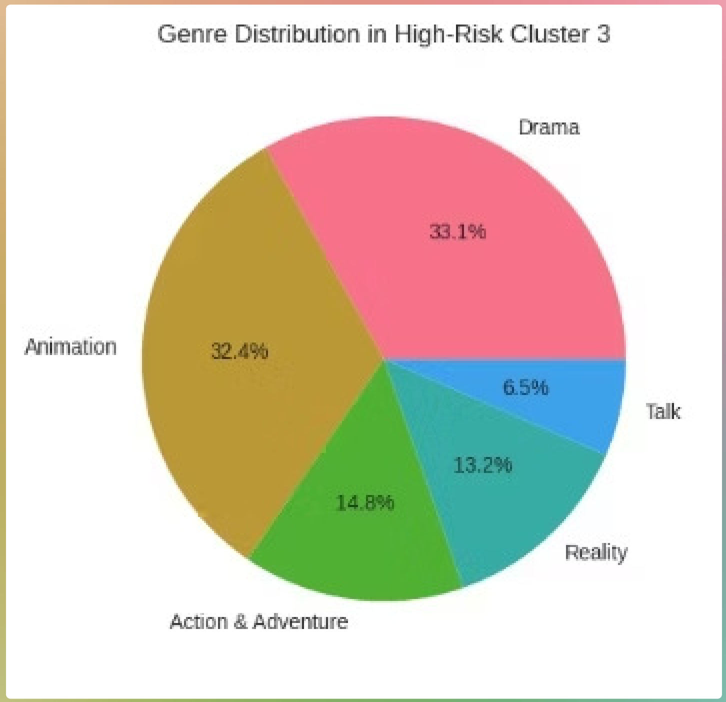
- **High-Risk Genres:** Drama, Mystery, Western show significantly higher drop-off rates.
- **Low-Risk Genres:** Comedy, News, and Kids' content consistently maintain better viewer retention.

Critical Content Attributes

- **Dialogue Density:** High dialogue density correlates with a 30% increase in viewer drop-off.
- **Cognitive Load:** Content with a cognitive load score of 9 sees a substantial 79.6% drop-off.
- **Pacing:** A medium pacing score (4–6) is identified as optimal for viewer retention.

Early Behavioral Signals

- **Engagement Cues:** Frequent pausing and rewinding often serve as early indicators of potential drop-off.
- **First Episode Vulnerability:** The inaugural episode of a series experiences a 3.6% higher drop-off rate, highlighting its critical role.



Strategic Segmentation: Identifying Key Viewer Clusters

1

Cluster 0 (Best)

Characterized by high hook strength and medium pacing, resulting in a 0% drop-off rate.

2

Cluster 1 & 2 (Balanced)

These clusters exhibit balanced content attributes, leading to a respectable ~65% average watch percentage.

3

Cluster 3 (Worst)

Defined by low hook strength, slow pacing, and high dialogue density, contributing to a significant 55.4% drop-off rate.

Actionable Segment Focus

Our strategy will concentrate on intervening in Cluster 3 episodes while replicating the successful attributes of Cluster 0 in high-risk content.

The ML Solution: Targeted Interventions for Retention

Predictive Model Highlights

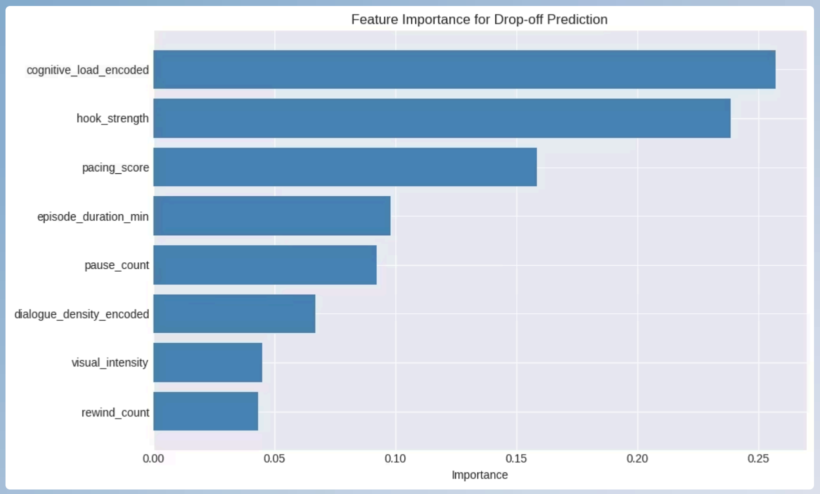
- **Top Features:** Cognitive load, hook strength, and pacing score are critical.
- **Accuracy:** Our model demonstrates 85% accuracy in predicting viewer drop-off.

k-NN Recommender System

This system intelligently matches low-performing episodes with high-performing peers to suggest improvements. For example, "Stranger Things – Episode 1" could be paired with episodes demonstrating better pacing to guide content adjustments.

Key Intervention Levers

- **Pacing Adjustment:** Target an optimal pacing score of 4–6.
- **Hook Enhancement:** Strengthen narrative hooks within the critical first 5 minutes of an episode.
- **Cognitive Load Reduction:** Simplify complex narratives in high-risk genres to ease viewer comprehension.



Prioritization Matrix: Maximizing Impact

This matrix helps us strategically deploy resources for maximum impact on viewer retention.

Quick Wins

- Reduce dialogue density in Drama genres.
- Shorten episodes exceeding 50 minutes.
- Implement mid-episode hooks to re-engage viewers.
- A/B test various pacing scores for optimal results.

Strategic Bets

- Develop genre-specific content templates based on successful attributes.
- Implement cluster-based recommendations for personalized content improvements.
- Personalize content delivery by viewing time (day/night).
- Integrate early alerts for high pause/rewind counts.

High Impact, Low Effort

Focus on optimizing hook strength within the crucial first episodes to quickly capture viewer attention.

High Impact, High Effort

Undertake a comprehensive redesign of content identified with high cognitive load in high-risk genres.

Proactive Measures: Risk Mitigation Strategies

Content Risks

Addressing inherent content challenges to maintain viewer engagement:

- **High Cognitive Load:** Implement strategies to simplify complex narratives and reduce mental effort for viewers.
- **Extreme Pacing:** Moderate pacing to the identified optimal 4–6 range, avoiding both overly slow and excessively fast narratives.

Viewer Risks

Tailoring content delivery to diverse viewing habits:

- **Mobile Viewers:** Optimize content for shorter segments and mobile-friendly consumption patterns.
- **Binge Platforms:** Ensure seamless episode-to-episode flow to facilitate continuous viewing and minimize drop-off between episodes.

Monitoring Framework

Establishing robust systems for continuous performance tracking:

- **Real-time Drop-off Scoring:** Implement a dynamic scoring system to monitor drop-off probability in real-time.
- **Episode "Health Score":** Develop a composite score combining pacing, hook strength, and average watch percentage to provide a holistic view of content performance.

Conclusion & Future Roadmap: Driving Retention Forward

01

Summary of Findings

Viewer drop-off is significantly influenced by pacing, hook strength, cognitive load, and dialogue density. Our clustering analysis reveals actionable segments for targeted interventions.

02

Immediate Next Steps (Business-Focused)

1. Launch Targeted Content Pilot Program
2. Operationalize the Content Improvement Engine
3. Establish a Content Health Monitoring Dashboard

03

Ambitious Goal

Achieve a substantial 15–20% reduction in overall viewer drop-off within the next 6 months, enhancing platform engagement and content value.

Thank You

By embracing a data-driven approach, we are committed to continuously improving viewer engagement and fostering a more compelling content experience.

For further questions or to discuss next steps:

Email: rajvibujad2906.com