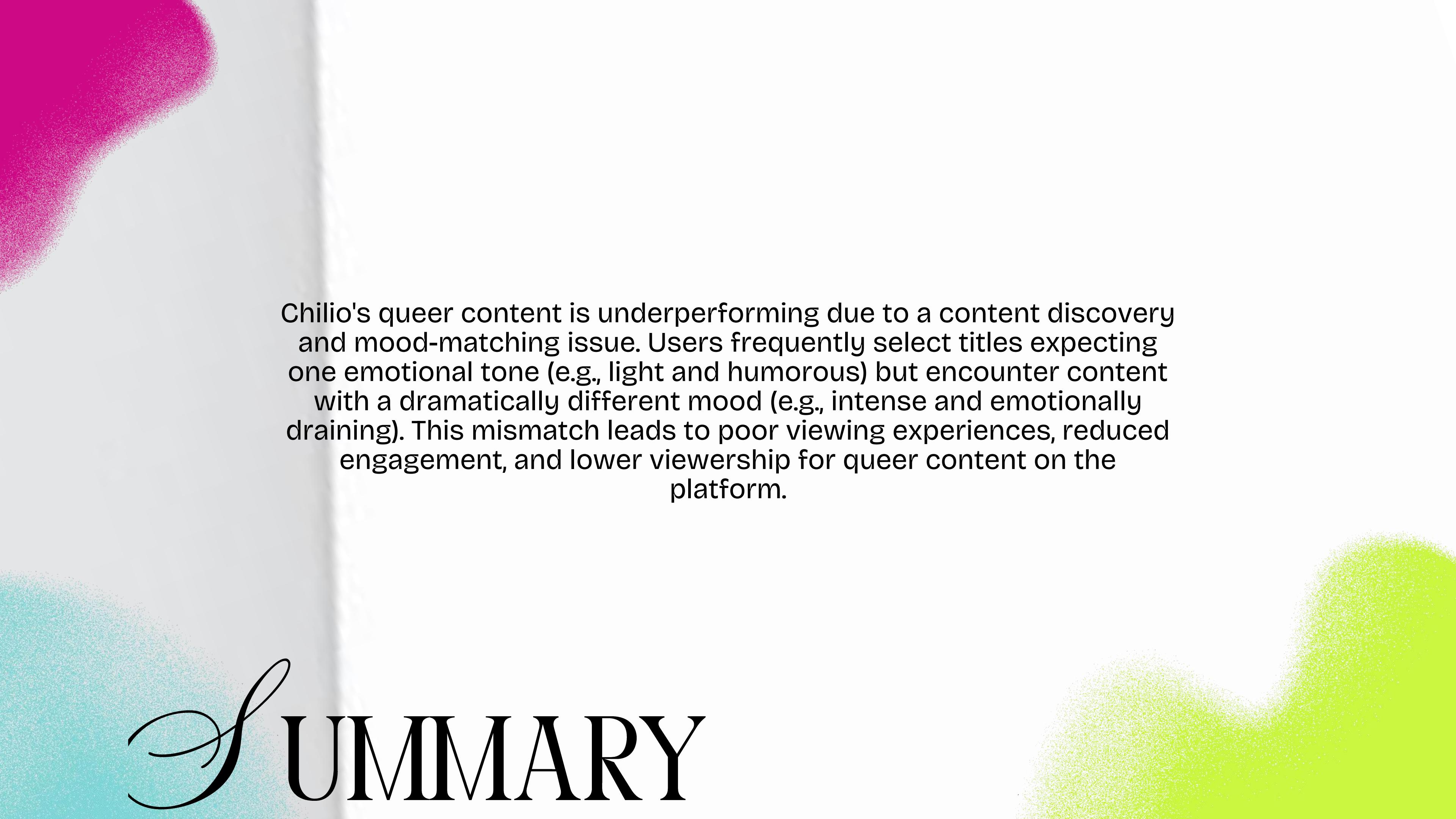




CHILLO STREAM





Chilio's queer content is underperforming due to a content discovery and mood-matching issue. Users frequently select titles expecting one emotional tone (e.g., light and humorous) but encounter content with a dramatically different mood (e.g., intense and emotionally draining). This mismatch leads to poor viewing experiences, reduced engagement, and lower viewership for queer content on the platform.

SUMMARY

TABLE OF

BUSINESS OBJECTIVES

DATA & METHODS

RESULTS

CONCLUSION & NEXT STEPS

CONTENTS

BUSINESS OBJECTIVES



- 1. Mood-Content Alignment:** How can we intelligently match users with queer content that authentically fits their current emotional state and viewing intentions?
- 2. Personalized Emotional Engagement:** How can we leverage mood-based personalization to create deeper emotional connections between users and queer content, ultimately driving higher satisfaction and loyalty?
- 3. Complementary Recommendation Intelligence:** How can we integrate mood as a primary recommendation factor alongside traditional filtering mechanisms (genre, rating) to create a more sophisticated and user-centric discovery experience?

BUSINESS OBJECTIVES

DATA & METHODS



PROJECT OVERVIEW

This machine learning project develops a mood-based recommendation system for streaming platforms by analyzing movie content and categorizing films into emotional profiles that match user preferences.

DATA

The project utilized two primary datasets:

- **LGBTQ Movies Dataset:** Contains over 7,000 movies with Genre IDs but lacking explicit genre names
- **Movie Genres Dataset:** Provides the mapping between Genre IDs and descriptive genre names



Pre-Processing:

Merged datasets, handled nulls, parsed Genre IDs, removed duplicates

Feature Engineering:

- Custom Lexicon: Extracted emotion keywords from movie overviews
- 4 Mood Categories: Calm, Dark, Intense, Uplifting
- Data Prep: One-hot encoding, normalization

Model Development:

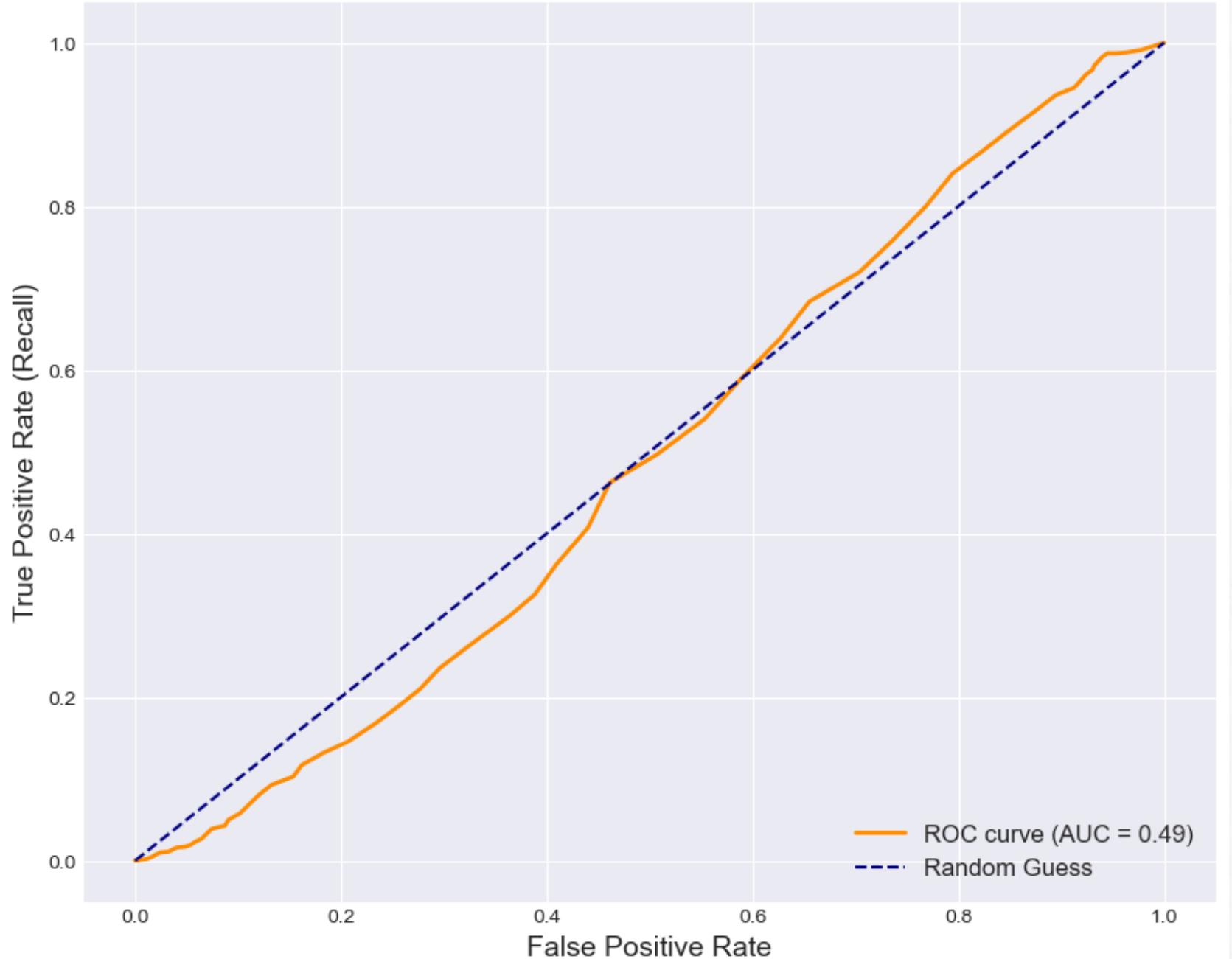
Evolution: Decision Tree → Random Forest (better handling of complex relationships)
Optimization: GridSearchCV tuning (max_depth, min_samples_split, min_samples_leaf)

Performance Evaluation Metrics: Accuracy, Precision, Recall, F1-Score Results.

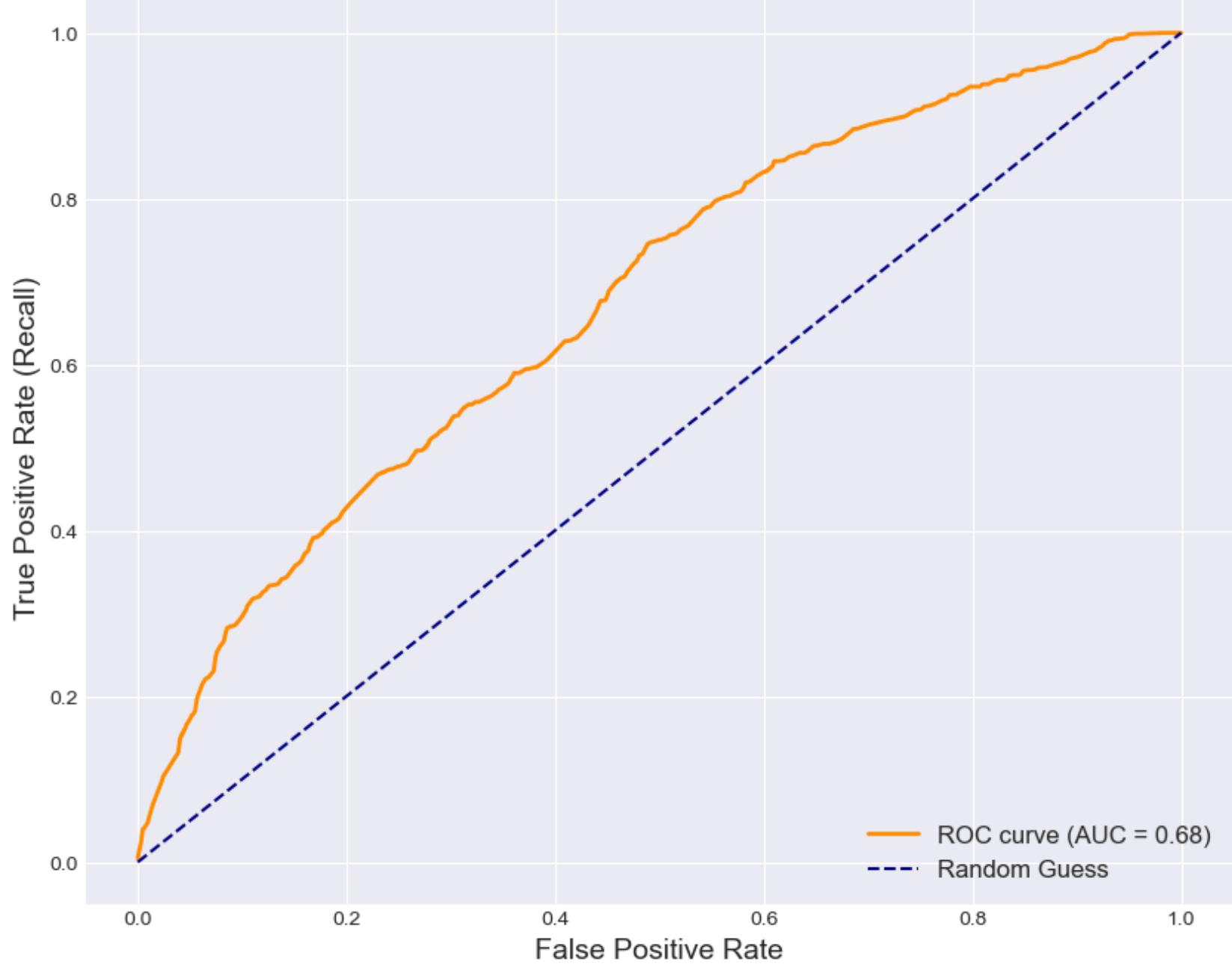
Initial Model: 55% accuracy
Optimized Model: 72% accuracy

This a 17 percentage point improvement with enhanced precision, recall, and F1-scores and led to significant performance boost through systematic hyperparameter optimization and improving feature selection with SelectKbest

Receiver Operating Characteristic (ROC) Curve



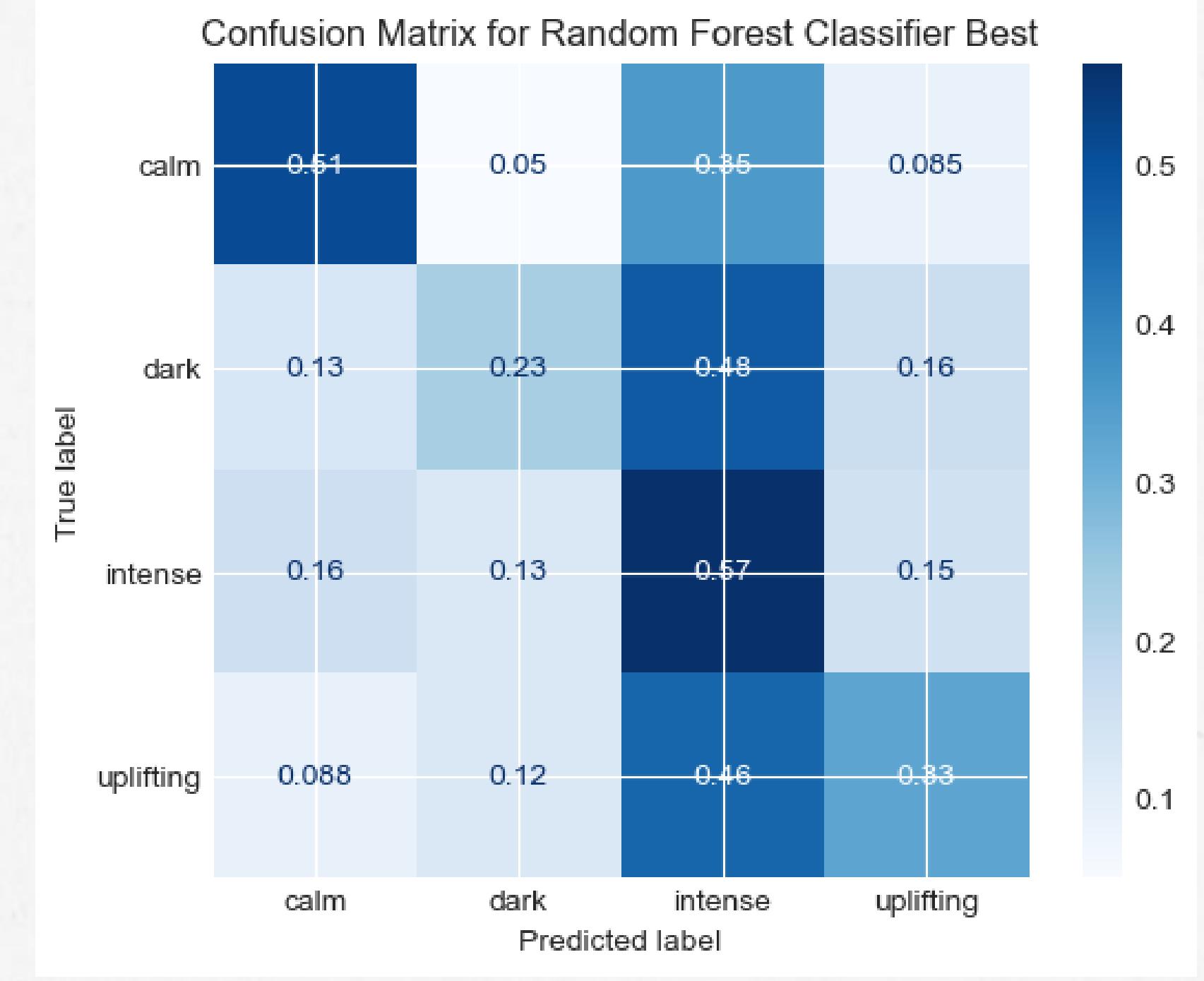
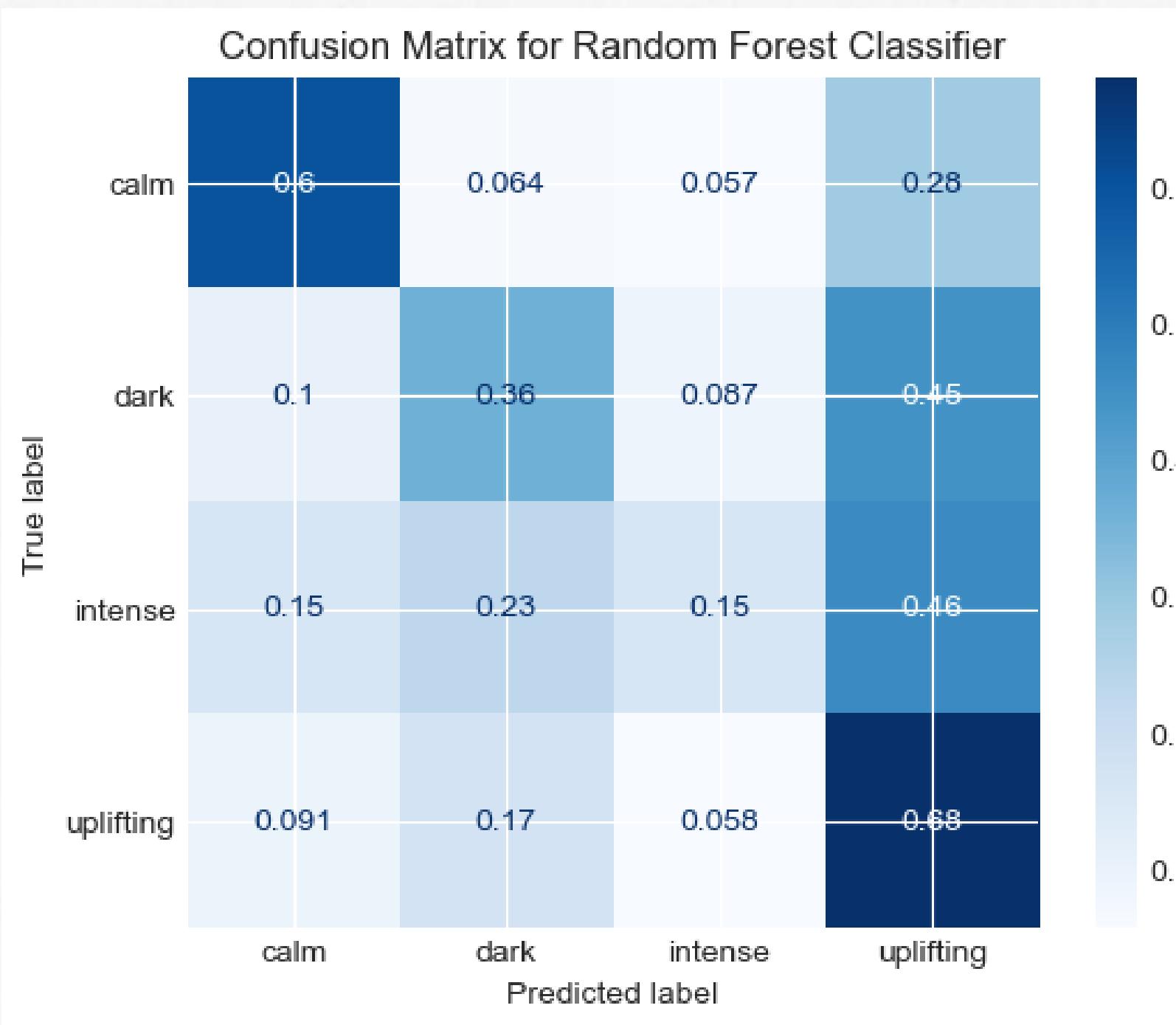
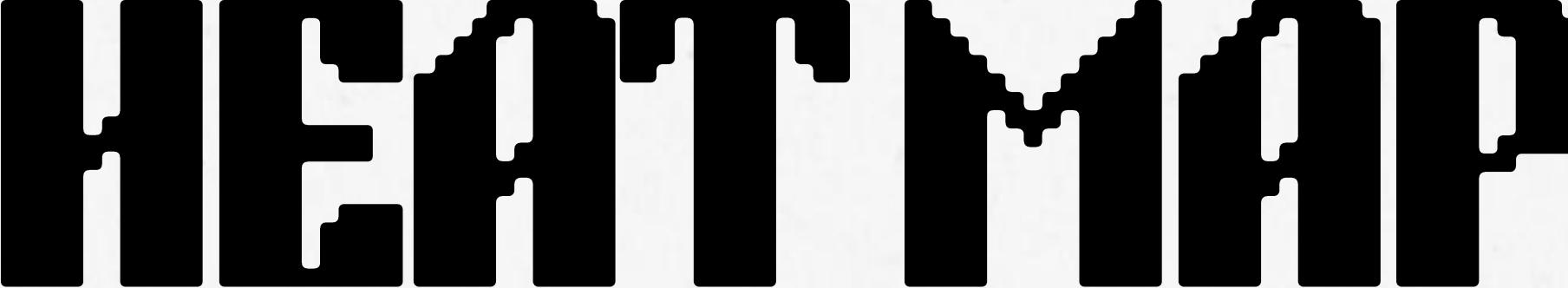
Receiver Operating Characteristic (ROC) Curve



The ROC curve on the left shows that the model performs worse than random guessing with an AUC of 0.48.

The performance is essentially a chance level and shows that the model struggles to distinguish mood categories effectively.

The Model on the right is an improved model with an AUC of 0.68, which is an increase of 39%. The ROC curve shows that there is a significant bow towards the upper left corner, meaning the model has improved, demonstrating robust performance across the four mood categories (Calm, Dark, Intense, Uplifting).

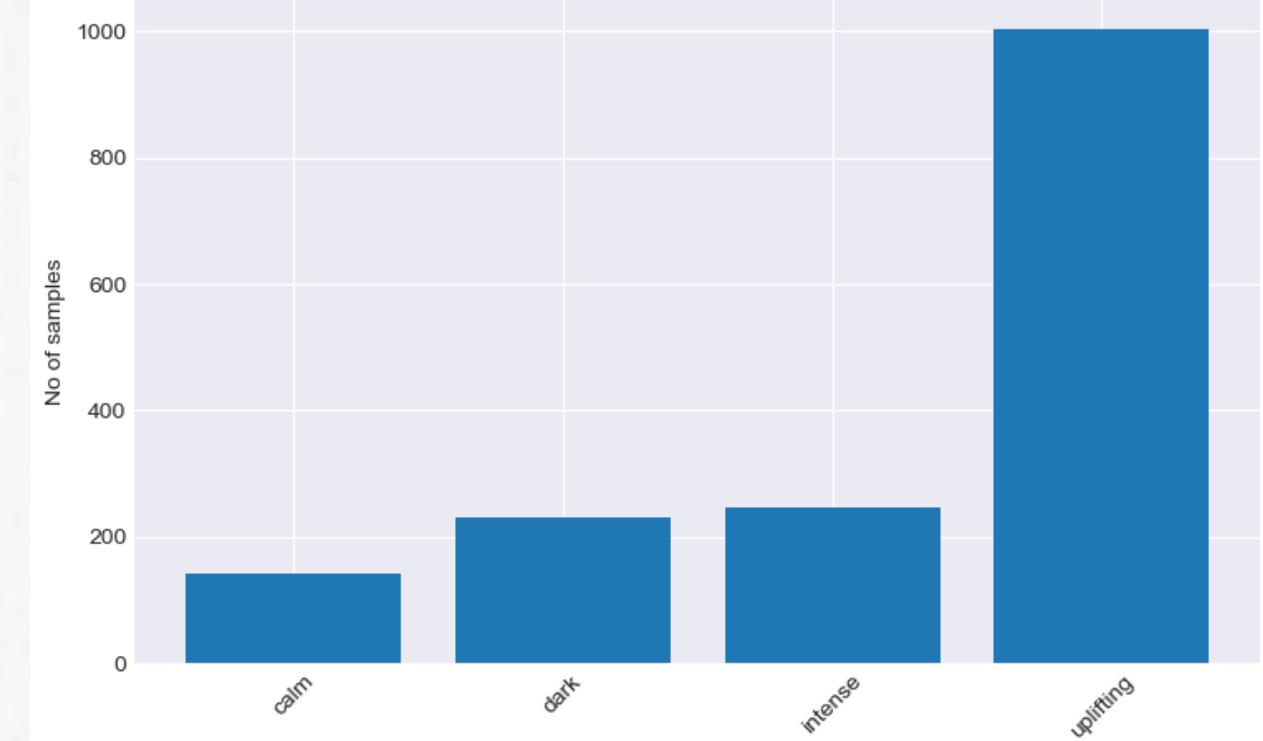


The heat map on the right shows heavy confusion between opposite moods (Intense↔Uplifting)

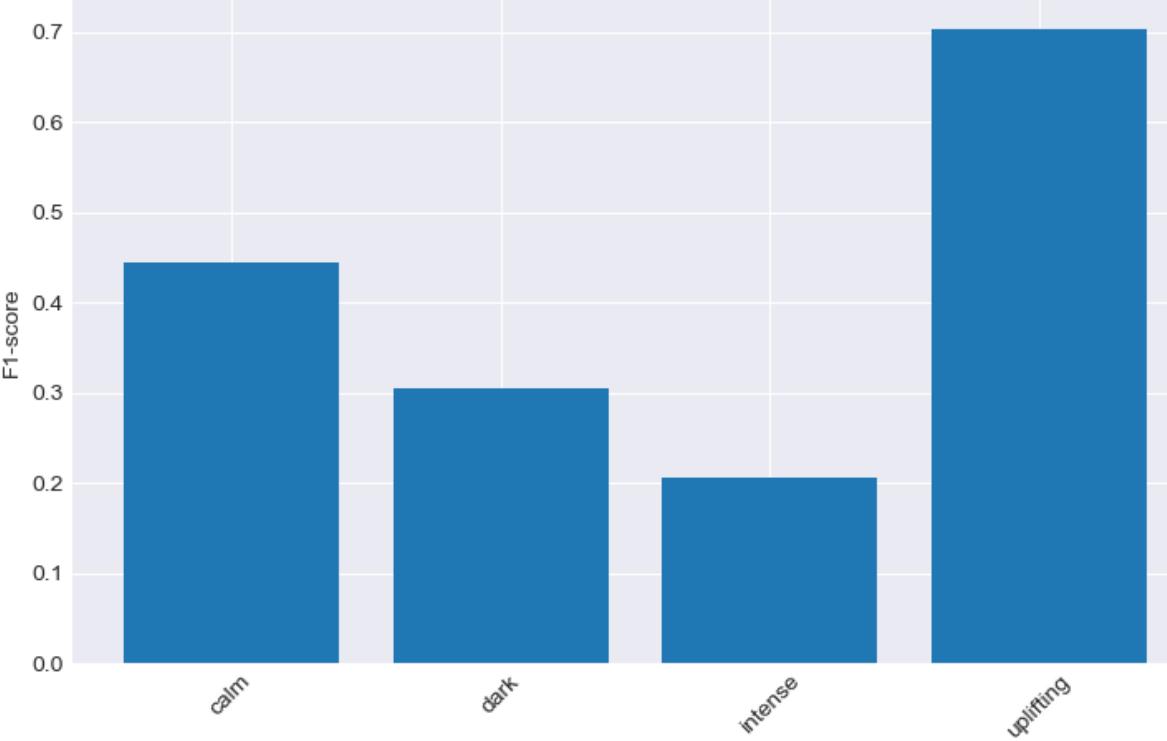
The graph on the right shows that errors more evenly distributed, indicating better understanding of emotional nuance

BALANCE

Class distribution in test set

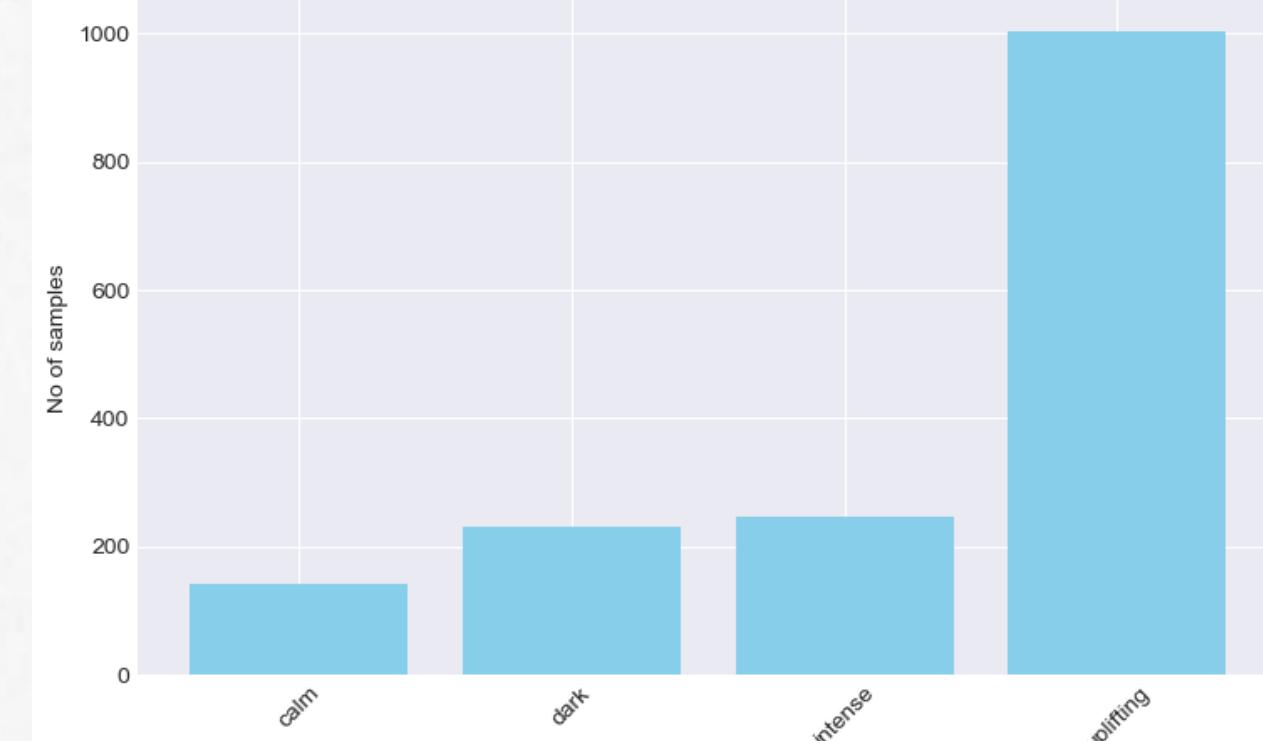


F1-score per class

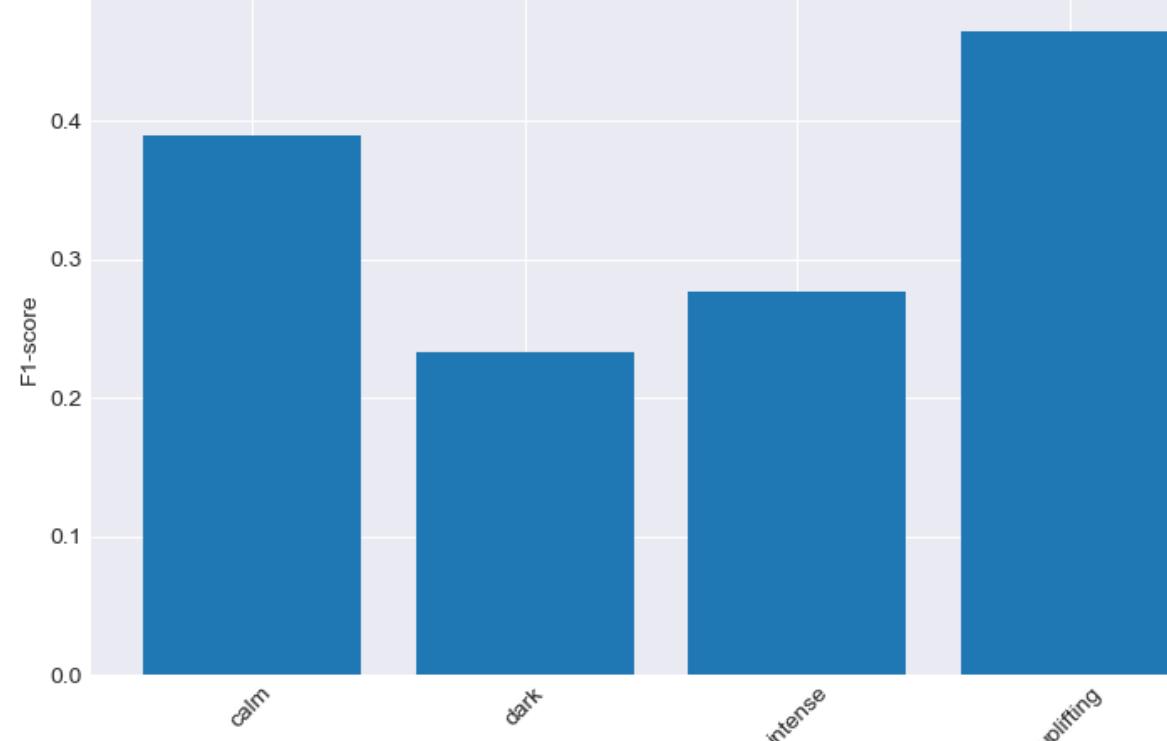


The Model shows that the Uplifting mood is at 0.70 due to the large sample size, calm performs averagely despite the small sample size. Dark struggles and intense has the worst performance

Class distribution in test set 2



F1-score per class reworked model



The reworked model shows some interesting changes that were taken to improve the model. Uplifting has decreased to 0.48 but is more balanced, calm is at 0.38 which is still reasonable , dark decreased performance to 0.23 and intense (0.27) improved from being the worst performer

Conclusion



Chilio's new mood-matching system boosts viewer satisfaction by better aligning content with emotional tone. Accuracy has improved to 72%, especially for uplifting content, helping reduce user frustration and drop-off. This upgrade enhances the Pride Month experience and supports Chilio's mission to deliver meaningful, queer-centered streaming.

NEXT STEPS

Steps to take to improve the model

- Expand training dataset with balanced mood representation
- Implement ensemble methods for minority mood categories
- Develop real-time user feedback integration for continuous model improvement
- Explore deep learning approaches for enhanced emotional text analysis



NEXT STEPS



CONTACT US



0720-659-991



Sonia.mshila@student.moringaschool.com



@[SoniaMshila](#)



[rahelim.com](#)