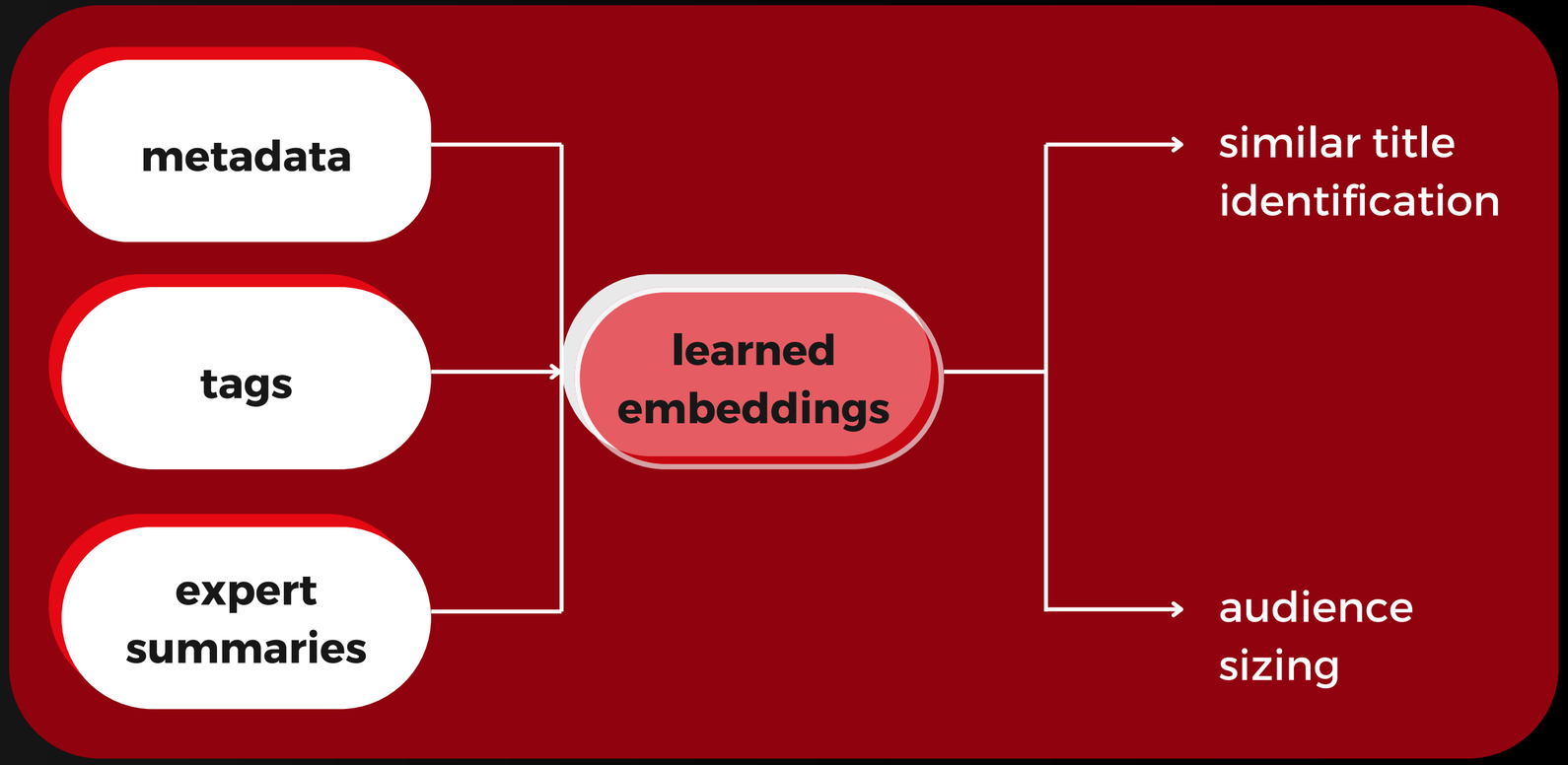
# NETFLIX

## CONTENT CREATION

### LEVERAGING VIEWER FEEDBACK: ENHANCING CONTENT CREATION THROUGH COMPREHENSIVE REVIEW ANALYSIS

MAHRUKH SHAMAS, AASNA SHAH, KRITIKA NAYYAR, MERIEM MEHRI

# Current State and Project Overview




*Netflix. (2021, October 4). "Supporting Content Decision Makers with Machine Learning." Netflix TechBlog.*

Our project aims to enhance targeted content creation for Netflix by integrating user reviews into their existing content analysis pipeline.

We experiment with Production Region's influence on viewer perceptions through NLP analysis of user reviews

Our custom pipeline solution is comprised of a multiclass classifier for predicted rating category and an extractive summarizer for general and role based summaries

# Approach + Method

## Data Collection

- Scraped data from IMDB and Rotten Tomatos

Data: Reviews + Ratings of 10 different popular Netflix shows: 5 North American + 5 East Asian

**Total 4 Datasets:**

1. No region
2. With region
3. East Asian
4. North American

## Pre-Processing

- Tokenization
- Stopwords (normal+custom)
- POS Tagging
- Lemmatization
- OneHot Encoder + Label Encoder
- SMOTE
- Categorized ratings into bins

**Experimented with:**

- n-grams  in CountVectorizer
- n-grams in TF-IDF
- custom Word2Vec model

## Experimentation + Model Testing

**Models Tested:**

1. Multinomial Logistic Regression
2. Linear SVC
3. Multinomial Naive Bayes
4. Random Forest
5. Gradient Boosting
6. Multiclass XGBoost

- Tested various models for each dataset + compared the classification accuracies and other parameters

# Approach + Method

## Model Selection + Experiment Outcome

Model Selected: Multinomial Logistic Regression

- Best classification accuracy across all datasets
- Better class balances across model
- Better F1-score, recall, precision scores

**Experiment Outcome:**

- Regional models not required for rating category prediction

## Summarizer: General

- Tokenizer
- TF-IDF
- Cosine Similarity
- Sumy Summarizer

**Output is transformed into learned embeddings for use in Netflix's models.**

## Summarizer : Role Focus

Complementary extension to summarizer to provide insights to:

- **Content Creation Execs**
- **Localized Assets Marketing Execs**
- **Studio Production Execs**

catering to different aspects of the content creation process by rating category.

# Business Insights + Implications

## 1 Diverse Perspectives and Feedback

**Broadening Understanding:** Expert summaries provide professional insights, while viewer reviews reflect diverse audience reactions, offering a broader understanding of content.

**Real-world Reception:** Viewer summaries reveal the public reception of content, highlighting areas of resonance or improvement and showing differences between expert opinions and viewer perceptions.

## 2 Enhanced Personalization and Recommendation

**Viewer Preferences:** Reviews provide insights into viewer preferences and dislikes, aiding in refining recommendation algorithms for personalized content matching.

**Content Discovery:** Analyzing viewer summaries can uncover less-known content favored by specific audience segments, enhancing content discovery beyond expert highlights.

# Business Insights + Implications

## 3 Quantitative and Qualitative Feedback Loop

**Volume of Feedback:** The abundance of viewer reviews offers a comprehensive dataset for uncovering trends and preferences, surpassing the limited scope of expert summaries.

**Qualitative Insights:** Viewer reviews provide detailed qualitative feedback, offering in-depth insights into the viewing experience beyond what is obtainable from ratings or expert analyses.

## 4 Benefits of Role-Focused Summarizer

**Marketing:** Utilizing summarized viewer reviews can help refine investments into dubbing and subtitles, making marketing campaigns more relevant and effective for different audience segments.

**Studio Production:** Insights from summaries can guide improvements in storytelling and production quality, enhancing viewer satisfaction and engagement.

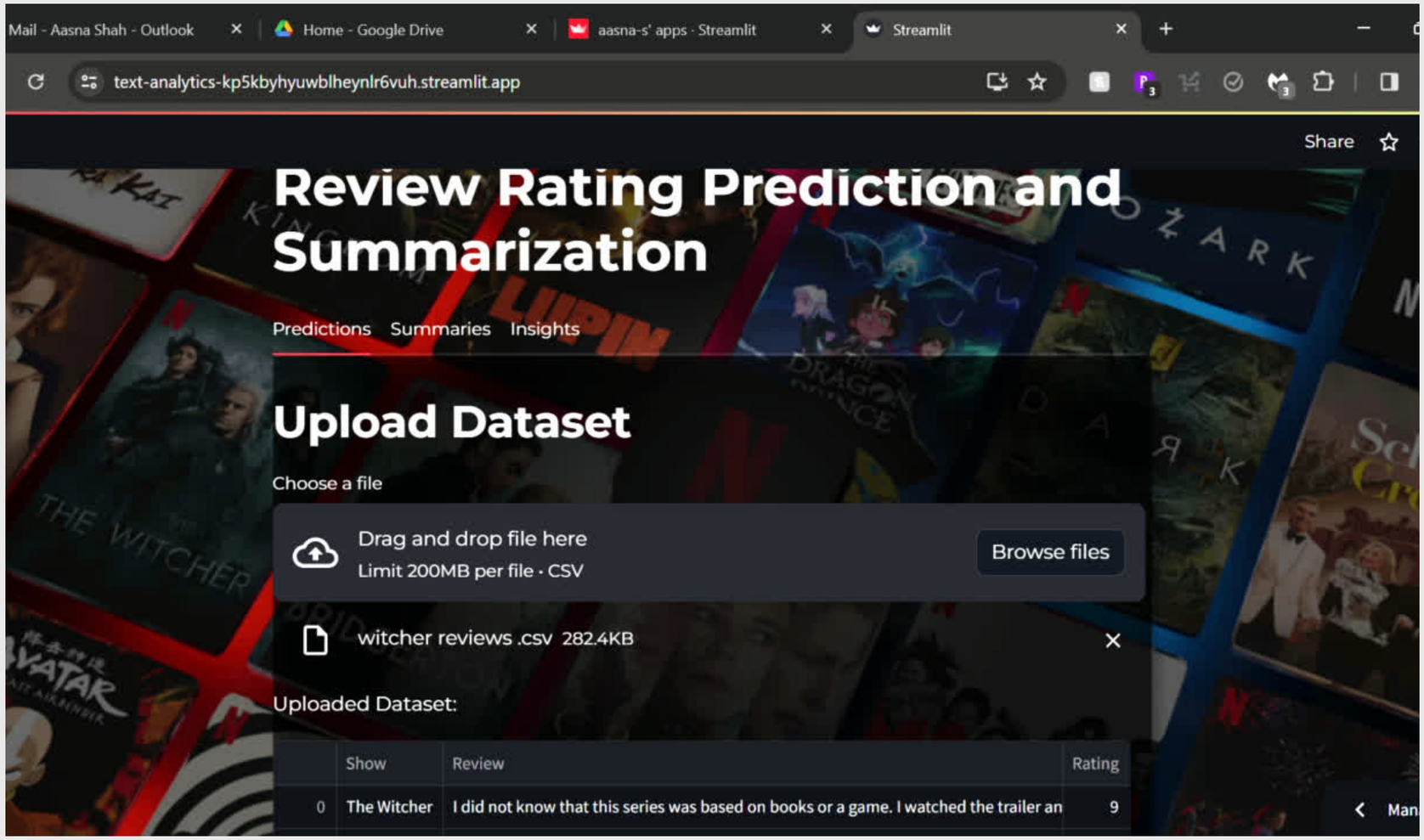
**Content Creation:** Analyzing viewer preferences through summaries can inspire content creators to develop stories and themes that resonate more deeply with their target audience.

# Behind the Scenes:

## Live Code Demo

To visually demonstrate the impact of user reviews on content matching & Netflix's recommendation system.
through network graphs.

**Streamlit Application**: https://text-analytics-kp5kbyhyuwblheynlr6vuh.streamlit.app/



# THANK YOU!