CIS 600 Social Media and Data Mining
Project Report

Movie Review Analysis for Letterboxd Data

Adithya Neelakantan (SUID: 663182945)
Aditi Mangesh Jadhav (SUID: 544349325)
Pratik Satpute (SUID: 566041095)
Shruti Vasave (SUID: 427009403)
Jeevan Tumkur Venkatesh (SUID: 740874557)

Table of Contents

1. Introduction

Letterboxd is a social media blog site for many film enthusiasts around the world to share their movie watching experiences through movie reviews, ratings, and recommendations with friends, other people who they share similar tastes in movies with. The cool thing that sets Letterboxd apart from its competitors is that despite it being a platform to write movie reviews and share recommendations and lists with their friends, these reviews are snarky, ironic, and often not straight forward movie reviews.

The crux of the project was to perform a deeper analysis of these movie reviews, to comprehend the sarcasm, the sentiments behind the reviews (and the snark), do a deep dive on the user behavior and sentiment trends behind these movie reviews to see how they fare every year, how the engagement has been building, what genres audiences tend to prefer and such. We delved into the vast repository of user-generated content on Letterboxd to provide insightful analysis across multiple dimensions. This framework comprised several key components that mainly covered the following areas.

– Sentiment Analysis
– User Engagement Analysis
– Genre Analysis
– Time Analysis
– Sarcasm Analysis

The project investigates different engagement patterns, exploring how user interactions with reviews vary by movie genre, time, and trends. By identifying spikes in engagement and shifts in genre popularity, the analysis sheds light on evolving user preferences and cultural trends. These different analytics collectively help give us a grounded understanding of how users interact with movies and Letterboxd posts, express their opinions about movies, and this in turn could help build recommenders and cater specific genre movies to certain audiences.

The snarky reviews on Letterboxd often make it hard to have an objective analysis system, because of how deeply contextual these reviews might be, in expressing their adoration (or loathing) for a film. These would obstruct a benchmark sentiment analysis model because these reviews seldom use objective words like “good”, or “liked it”. To help further refine the sentiment classification, the sarcasm detection add-on, powered by a pre-trained Transformer-based model, identifies ironic or snarky tones that might obscure true sentiment. By understanding the nuances of sentiment and sarcasm in user movie reviews, the analysis can also contribute to more sophisticated sentiment-based engagement tools, making the Letterboxd experience so much more enriching.

2. Objectives

2.1 Problem Statement

The primary objective of this project is to analyze Letterboxd reviews and user-generated content to uncover insightful patterns and trends in user behavior, sentiment expression, and engagement dynamics. The main goal of this initiative is the task of developing an advanced sentiment analysis framework capable of decoding not only explicit emotions but also the nuanced and deeply contextual sentiments that characterize Letterboxd reviews. Given the platform's unique tone — where irony, and sarcasm are prevalent—traditional sentiment analysis models are insufficient. Therefore, a key focus is on integrating sarcasm detection, powered by Transformer-based models, to accurately interpret the underlying sentiment of reviews that might otherwise mislead due to their indirect expressions. The goal is to bridge the gap between the literal and the intended sentiment to gain a true understanding of user emotions, whether positive, negative, or neutral, even in the absence of overt language markers.

In addition to sentiment analysis, the project attempts to explore user engagement patterns across various dimensions. This involves examining how users interact with content through likes, comments, and logging activity, identifying trends over time, and understanding how engagement fluctuates during specific periods, such as major movie releases or cultural events. By delving into user behavior, we aim to uncover meaningful insights into how interaction levels vary across different types of users, genres, or trends. Furthermore, our project investigates genre-based preferences to identify which movie categories resonate most with audiences, attract higher engagement, and elicit distinct sentiment patterns. A deeper analysis of genre dynamics also reveals how popularity and user sentiment for certain categories evolve over time, shedding light on broader cultural and cinematic trends.

Another critical dimension of this study is a time-based analysis, where we explore temporal patterns in user activity, looking at daily, weekly, and yearly fluctuations in reviews and interactions. By correlating these trends with significant cinematic or cultural happenings, the analysis offers a comprehensive view of how user engagement adapts to changing contexts. Additionally, the project tackles the challenge of sarcasm detection—a particularly complex layer of this work—by systematically identifying and quantifying ironic or snarky tones in reviews. Understanding the prevalence and nature of sarcasm within the community provides a window into the cultural dynamics of the platform and highlights its role in driving engagement.

This analysis aims to generate actionable insights into how users interact with Letterboxd, express their opinions about movies, and engage with the platform's unique social environment. These insights are expected to contribute to the development of smarter recommendation systems, better sentiment-based tools, and more targeted content strategies, enhancing the overall user experience. By leveraging the findings from sentiment, engagement, sarcasm, genre, and temporal analyses, this project aspires to not only deepen our understanding of the Letterboxd community but also to set a foundation for future innovations in sentiment-driven user analytics for similar platforms.

2.2 Analyses and Reasoning

The project employs a comprehensive analytical framework to uncover insights from Letterboxd's rich repository of user-generated content; integrates advanced machine learning techniques, data mining, and contextual understanding of user behavior. The analyses span several interconnected domains to ensure a complete view of the platform's dynamics. We made sure each was grounded in logical reasoning and data-driven methodologies.

Sentiment Analysis with Sarcasm Detection

Traditional sentiment analysis often falters when confronted with the ironic tone prevalent in Letterboxd reviews. To address this, we employed a Transformer model trained on sarcastic tweets, capable of identifying nuanced emotional expressions, includng sarcasm and irony. The sarcasm detection module is key to deciphering hidden sentiments where reviewers might use witty, indirect language to express their opinions. For example, a sarcastic review like 'An absolute masterpiece...of boredom' could mislead a conventional model but would be appropriately flagged by our approach. This enables a more accurate classification of sentiments, creating a foundation for analyzing user preferences, emotional engagement, and cultural expressions in reviews.

User Engagement Analysis

User interactions such as likes, comments and logging are indicators of content resonance and community dynamics. By analyzing these engagement patterns, we identify trends that reveal what type of content draws the most attention. For instance, we explore whether longer, in-depth reviews receive more engagement than snarky one-liners or whether certain genres inspire greater community participation. Reasoning from these patterns, we derive actionable insights for optimizing user retention, identifying high performing c ontent, and understanding how different features of the platform influence engagement.

Genre Analysis

Genre analysis delves into how user preferences vary across movie categories. This is crucial for understanding the diversity of tastes within the Letterboxd community. By correlating genres with sentiment and engagement data, we discern patterns, such as whether horror movies receive more polarizing reviews or whether comedies invite higher engagement due to the witty tone of reviews. These insights are vital for recommending genre-specific content and for aligning platform strategies with user interests. The findings also illuminate how cultural trends and cinematic preferences evolve over time.

Temporal Analysis

By incorporating a temporal lens, the project explores how user engagement and sentiment shift over different time periods. This includes analyzing daily and seasonal activity, as well as engagement spikes linked to major movie releases, award shows, or other cultural events. Temporal analysis helps explain why certain reviews gain traction during specific moments and how these moments influence broader user participation on the platform. Such patterns provide a deeper understanding of user behavior and allow for better planning of platform features, such as time-sensitive recommendations or event-driven promotions.

Understanding Cultural Contexts through Sarcasm
The analysis of sarcasm provides a unique perspective on the cultural nuances of Letterboxd reviews. By quantifying and categorizing sarcasm, we aim to determine how ironic tones influence engagement and the overall sentiment landscape of the platform. For example, identifying sarcastic patterns can help differentiate genuine enthusiasm for a movie from a tongue-in-cheek praise, offering more refined metrics for audience perception. This reasoning aligns with the need to adapt analytical models to the distinct characteristics of platforms like Letterboxd, where humor and irony are central to user interaction.

Integrated Insights and Reasoning
The interconnected nature of these analyses ensures a full understanding of the Letterboxd ecosystem. Sentiment analysis informs genre preferences; engagement data reveals the impact of sarcasm and tone; temporal analysis contextualizes these findings over time. By synthesizing these dimensions, we generate a comprehensive picture of how users interact with movies, express opinions, and engage with the platform. These insights are supported by logical reasoning and validated through data driven experimentation, ensuring that the conclusions drawn are robust, actionable, and aligned with Letterboxd's unique community dynamics.

This approach ensures that the analyses are not only descriptive but also explanatory and predictive, enabling us to understand the why behind user behaviors and to anticipate future trends. The project's reasoning framework, rooted in both qualitative and quantitative insights, lays the groundwork for actionable applications such as improved recommendation systems, enhanced user engagement strategies, and deeper cultural impact studies.

2.3 Significance

This project is highly valuable, both for comprehending the distinct user interactions on Letterboxd and for enhancing analytical models for sentiment-focused social media platforms. By exploring the intricate and complex aspects of Letterboxd reviews, this research tackles a significant issue in natural language processing: understanding sarcasm, irony, and contextual feelings. The frequency of clever, and frequently indirect remarks on the platform makes it an intriguing environment for enhancing sentiment analysis tools and developing models that delve deeper than mere surface interpretation. The incorporation of sarcasm detection improves the precision of sentiment analysis, leading to more advanced tools that could be utilized in various fields where sarcasm and irony play significant roles in user interactions.

From a cultural and film perspective, this project offers important insights into how viewers interpret and interact with films. Through the examination of sentiment and engagement patterns, the research underscores changing user preferences, including transitions in favored genres, the emergence of niche interests, and the cultural events that trigger increases in activity. These results are essential for filmmakers, marketers, and content creators aiming to comprehend audience preferences and develop strategies to meet their needs. The insights derived from data regarding genre popularity and user behavior can enhance the targeting of movie recommendations, promotions, and content curation strategies, providing users with a personalized and enriched experience.

The initiative further adds to the wider research on user involvement in social media contexts. By examining interaction patterns—such as logs, comments, likes, and lists—it reveals the elements that influence user engagement, times of increased activity, and the kinds of content that appeal most to audiences. These insights are crucial for platforms such as Letterboxd, as they can inform improvements to their user interface and recommendation systems, ultimately cultivating a more interactive and involved community.

Also, the temporal analysis provides a deeper understanding of how social events impact user behavior, giving us a lens into the interconnectedness of cultural trends and digital discourse. This sort of statistic can help platforms like Letterboxd anticipate user needs during key moments, such as blockbuster releases or award seasons, ensuring that the platform remains relevant and engaging.

All in all, this project holds significance in advancing sentiment analysis technologies, deepening our understanding of user behavior, and providing actionable insights for both Letterboxd and similar platforms. The findings not only enrich the movie-watching experience for users but also create opportunities for industry stakeholders to better align with audience preferences. By handling the complexities of sarcasm, sentiment, and engagement, the project acts as a foundation stone for future innovations in user-centric social media analytics and sentiment-driven recommendations.

3. Data and Preprocessing

3.1 Dataset

The dataset includes a large number of user-generated film reviews from Letterboxd, a social media site devoted to rating and discussing films. The data covers reviews of a wide range of film genres, from new releases to classic films, and covers around ten years (2014–2023). The title of the movie, the year of release, the reviewer's username, the date of the review, the written content, engagement metrics (mostly likes), and the genre classification are all included in the format of each review entry. Unlike traditional film criticism, these reviews are famous for their conversational, informal tone, which frequently incorporates amusing insights and references to online culture. The degree of involvement varies greatly, with some reviews receiving hundreds of likes and others receiving very little activity.

	movie_name	Release Year	Reviewer name	Clean_Review_date	Clean_Review	Clean_Comment Count	Like count	genre
0	Clue	1985	Branson Reese	1996-10-16	My dad got in so much trouble for showing me t...	6	2,286 likes	Comedy
1	Beetlejuice	1988	Branson Reese	1999-10-21	Thank GOD Tim Burton made this movie in 1988 a...	12	3,304 likes	Comedy
2	Being John Malkovich	1999	Than Tibbetts	2010-10-04	Malkovich. Malkovich Malkovich Malkovich, Malk...	6	4,300 likes	Comedy
3	The Muppets	2011	Jeff	2012-03-06	It's fine if you don't like this movie, but it...	31	NaN	Comedy
4	Mysterious Skin	2004	Cole	2012-03-11	This movie is beautiful, captivating, fascinat...	4	6 23 likes	Drama
...	...	...	...	...	...	...	...	...
2832	Drive	2011	k??rsten	NaN	Yes, I just saw it for the first timeYes, I lo...	9	2,160 likes	Action
2833	Fight Club	1999	hunt??r	NaN	if I was next to brad, I would have dropped th...	19	NaN	Drama
2834	The Bling Ring	2013	k??rsten	NaN	not a single good shot or outfit in this entir...	30	NaN	Crime
2835	A Serbian Film	2010	DirkH	NaN	OH MY GOD, LOOK AT HOW CONTROVERSIAL I AM!!!!!!	65	NaN	Horror
2836	CODA	2021	James (Schaffrillas)	NaN	Crazy how such a cliche and predictable movie ...	32	NaN	Drama

3.2 Data Cleaning

The platform's extensive use of pop culture allusions and modern online lingo suggests that its audience is primarily young and digitally native. This dataset offers important insights into the dynamics of user participation, the evolution of online film conversation, and contemporary film reception patterns. The reviews show how cinema criticism has changed from rigorous analysis to more approachable, community-driven conversations thanks to social media platforms. In the context of movie conversations, this extensive collection is a useful tool for comprehending how modern films are received and how internet users behave.

To guarantee data quality and consistency, our movie review dataset's data cleaning procedure employs a systematic methodology. Our cleaning procedure starts with the raw data from 'letterboxd_data.csv' and covers a number of important areas that need to be corrected and standardized.

Cleaning engagement indicators receive a lot of attention. A custom clean_likes function transforms the 'the Like count' column, which initially contains non-numeric characters and different formats. By removing unnecessary characters and standardizing the information into integers, this function ensures consistency by setting missing data to zero. In a similar manner, any non-numeric entries are handled correctly and replaced with zeros as needed by converting the 'Clean_Comment Count' column to integers.

3.3 Text Preprocessing

Date Standardization
Another essential step in our cleaning procedure is date standardization. A consistent datetime format is applied to the 'Clean_Review_date' column. To maintain temporal continuity in our dataset, any dates that cannot be parsed are first tagged as NaN and subsequently filled using a backfill technique. This method offers a comprehensive timeline for analysis while preserving data integrity. The cleaned dataset is exported to 'cleaned_letterboxd_data.csv' following these changes, with index numbers removed to preserve a clean format. The completed dataset is shown for instant confirmation of the cleaning procedures, and the entire process is intended to be transparent and verifiable.This extensive cleaning procedure guarantees that our dataset is formatted correctly for further analysis, especially for activities that call for precise temporal monitoring or the assessment of engagement metrics. The systematic approach helps maintain data integrity while preparing the information for meaningful analytical insights.

Text Tokenization
The second step in our sentiment analysis project for movie reviews is text preprocessing. We've created a streamlined procedure that turns disorganised review content into data that can be analysed by using a methodical approach to cleaning and standardising raw text data. Our preparation pipeline implements multiple crucial cleaning steps by handling reviews one at a time using a custom preprocess_text function. Tokenization is the first step in the procedure, in which each review is divided into discrete words. Since TweetTokenizer is more adept at handling casual language and social media-style writing, which are prevalent in contemporary movie reviews, we have chosen it over the word_tokenize tool. We remove stop words, common words like "and," "the," and "of" that don't contribute any analytical value, after tokenisation. This stage aids in concentrating our study on the significant stuff that genuinely conveys emotion.

	movie_name	Release Year	Reviewer name	Clean_Review_date	Clean_Review	Clean_Comment Count	Like count	genre
0	Clue	1985	Branson Reese	1996-10-16	My dad got in so much trouble for showing me t...	6	2286	Comedy
1	Beetlejuice	1988	Branson Reese	1999-10-21	Thank GOD Tim Burton made this movie in 1988 a...	12	3304	Comedy
2	Being John Malkovich	1999	Than Tibbetts	2010-10-04	Malkovich. Malkovich Malkovich Malkovich, Malk...	6	4300	Comedy
3	The Muppets	2011	Jeff	2012-03-06	It's fine if you don't like this movie, but it...	31	0	Comedy
4	Mysterious Skin	2004	Cole	2012-03-11	This movie is beautiful, captivating, fascinat...	4	623	Drama
...	...	...	...	...	...	...	...	...

Stemming and Lemmatization

The next important stage in both stemming and lemmatisation is word normalisation. Lemmatisation adopts a more complex strategy by transforming words to their dictionary root form, such as changing "better" to "good," whereas stemming reduces words to their base form by eliminating suffixes (such as changing "watching" to "watch"). By standardising variants of the same word, these procedures improve the accuracy of our analysis.

In order to prepare the processed tokens for sentiment analysis, we finally reassemble them into logical strings. We can process our entire dataset with ease because of this modular method, which also gives us the ability to change individual steps as necessary. The outcome is clear, standardized text data that eliminates noise that can impede our research while preserving the fundamental tone of the original reviews. Because movie reviews frequently use informal language, a variety of punctuation, and numerous writing styles, all of which must be standardised for meaningful analysis,this preprocessing pipeline has shown especially good results.

	movie_name	Release Year	Reviewer name	Clean_Review_date	Clean_Review	Clean_Comment Count	Like count	genre	Processed Review
0	Clue	1985	Branson Reese	1996-10-16	My dad got in so much trouble for showing me t...	6	2286	Comedy	dad got much trouble showing kid started sayin...
1	Beetlejuice	1988	Branson Reese	1999-10-21	Thank GOD Tim Burton made this movie in 1988 a...	12	3304	Comedy	thank god tim burton made movie 1988 2008 . im...
2	Being John Malkovich	1999	Than Tibbetts	2010-10-04	Malkovich. Malkovich Malkovich Malkovich, Malk...	6	4300	Comedy	malkovich . malkovich malkovich malkovich , ma...
3	The Muppets	2011	Jeff	2012-03-06	It's fine if you don't like this movie, but it...	31	0	Comedy	fine like movie , probably mean angry , hate-f...
4	Mysterious Skin	2004	Cole	2012-03-11	This movie is beautiful, captivating, fascinat...	4	623	Drama	movie beautiful , captivating , fascinating , ...
...	...	...	...	...	...	...	...	...	...
2832	Drive	2011	k??rsten	NaN	Yes, I just saw it for the first timeYes, I lo...	9	2160	Action	yes , saw first timeyes , loved everything ity...
2833	Fight Club	1999	hunt??r	NaN	if I was next to brad, I would have dropped th...	19	0	Drama	next brad , would dropped soap
2834	The Bling Ring	2013	k??rsten	NaN	not a single good shot or outfit in this entir...	30	0	Crime	single good shot outfit entire thing
2835	A Serbian Film	2010	DirkH	NaN	OH MY GOD, LOOK AT HOW CONTROVERSIAL I AM!!!!!!	65	0	Horror	oh god , look controversial ! ! !
2836	CODA	2021	James (Schaffrillas)	NaN	Crazy how such a cliche and predictable movie ...	32	0	Drama	crazy cliche predictable movie could sooooooo ...

4. Model Development

This project uses two core components for its analytical framework: a lexicon-based sentiment analysis model using the VADER (*V*alence *A*ware *D*ictionary and s*E*ntiment *R*easoner) Sentiment Analyzer, and a Pre-trained Transformer-based model for sarcasm detection. These components work alongside each other to provide a deeper understanding of engagement by users, capturing both explicit and implicit sentiments in movie reviews.

4.1 Sentiment Analysis Using VADER

The sentiment analysis part of this project employs NLTK's VADER analyzer, a tool that is designed for sentiment categorization in text data, particularly for social media and short text. VADER's suitability stems from its ability to handle nuanced expressions, such as emoticons, slang, colloquial expressions and punctuations, which are prevalent in user-generated content like movie reviews.

The VADER Scoring

VADER uses a pre-built sentiment lexicon corpus that has columns where words are categorized and scored for their intensity in sentiment (like = ~0.2, love = ~0.7, despise = ~ – 0.7 etc.). It's primarily adjectives and other descriptive words like "excellent" or "amazing" that receive positive scores, while descriptors like "terrible" or "horrible" that express negativity are scored negatively. This tool also emphasizes the intensity of expression through repeat letters, expressions, case sensitivity, helping our case vastly, because of how colloquial these reviews are. An "AMAAAAAAAZZZINGGGGGG!!!!!!!!" would be scored higher than an "amazing." Each of these variants are pre-taught to the model through these tags. The contextual elements are also integrated through words like "not bad" or "I didn't hate it" which both through just an analyzer would be negative because of the labels "bad" or "hate", but within context are more neutral-positive leaning.

The VADER sentiment analyzer computes scores across three sentiment dimensions –
– Positive (pos): The intensity of positive sentiment in the review: Compound score > 0.5.
– Neutral (neu): The proportion of review that is deemed sentiment-neutral: Compound score between – 0.5 and 0.5.
– Negative (neg): The intensity of negative sentiment in the review: Compound score < – 0.5.

A compound score is then calculated by normalizing the sum of these scores in a review (entity based) to a range between – 1 (most negative) and +1 (most positive). Each entity aspect of the movie (if separated that way – acting is bad, cinematography is good etc). The VADER compound score is the primary metric used for classification, applied to each review to classify its sentiment based on the compound score. This approach allowed for quick and efficient sentiment analysis of large datasets while maintaining interpretability and robustness intact.

4.2 Sarcasm Detection Using a Transformer Model

There isn't a standard way to tell that the sentiment analyzer actually works, unless we do a two-step pass to make sure we're confident that the reviews we scored are accurate despite them being ironic. This extra step lets us know two things, one – whether the review itself was ironic (sarcastic) or earnest (not sarcastic), and whether the sentiment analyzer was able to catch the sarcasm in the review and account for it in the VADER score. To address this, we incorporated a Pre-trained Transformer-based model. Transformers are state-of-the-art in NLP models and heavily employed due to their ability to understand context and nuanced expressions in text.

The Rationale behind the Sarcasm Add-on Feature
Sarcasm often involves a mismatch between literal text and intended meaning. Take the example of the sentence –
"That's great, another 4 hour war drama seen from the perspective of the pigeon sitting atop a castle."
Here, all a classic NLP model would get is that (1) It's a four hour movie, (2) It's Great (Sentiment ~0.6 POSITIVE), (3) It's a War Drama seen from the perspective of a pigeon? and (4) Set in a castle?
From a literal standpoint, yes, that's correct. But the tone of this implies that the user was bored, and did not like the length of the film. These contextual tones and subtle cues are critical for a sentiment to be accurate.

Transformers, particularly those that are pre-trained on large datasets, are well-suited for this task because they have the ability to encode contextual information through multi-head attention mechanisms. They have infrastructures that can capture relationships between words, even when separated by long distances in the text. Here's where the other analyzer trends could come in handy. In the context of the example used, War movies in general, aren't a crowd favorite, and are often reviewed neutral or negative leaning. Very few movies like Schindler's List are able to stand out, often as outliers. Subsequently, these analytics could be integrated to the model to help the sarcasm detection do better, to be able to tell whether a user is interested in these, or not, depending on other results.

Model Architecture
The sarcasm detection add-on in the model is based on a Transformer architecture, fine-tuned for the specific task of sarcasm detection. The model uses pre-trained weights from a transformer architecture like BERT or RoBERTa, but with mods added, like the SemEval-2018 Task 3 dataset, specifically for the sarcasm features. We compute the overall results in sentiments, for the model, added in with the sarcasm feature, and then pass it through a Softmax layer to give us Output Probabilities for two labels, is_sarcastic (Probability = 1), and a not_sarcastic (Probability= 0).

It uses these probabilities to also check for a confidence score, to see how confident it is that the review our model gave a sentiment score to is ironic or earnest. A confidence score of (> 0.50) implies that it is pretty confident that a review is ironic, and that the sentiment analyzer picks up the irony of the review. A score of (<0.5) implies that the review was an earnest review, and the score is then reflected accordingly. It returns a True/False based on the irony of the review.

Training and Fine-Tuning

Although the model leverages pre-trained weights, fine-tuning was performed to adapt it to the specific characteristics of Letterboxd reviews. The fine-tuning process involved the following features.

– **Dataset Preparation:** Labeling reviews as sarcastic or non-sarcastic using a combination of manual pseudo annotation and heuristic rules (e.g., picking up sarcasm cues like “Yeah, right” or “Totally loved this garbage”).

– **Optimization:** Using the Adam Optimizer and Learning Rate Scheduling to help get faster convergence regions.

– **Validation**: A held-out dataset was used to validate the model’s performance, to make sure generalizability is maintained.

Integrating the two sections

The output of the sarcasm detection model was then used to adjust the VADER-based sentiment classification like so –

– For reviews flagged as sarcastic, the sentiment score was recalibrated by reducing the weight coefficients of the dominant sentiment category.

– This adjustment helped ensure that the sarcastic reviews did not skew the overall sentiment distribution of the review. For example, sarcasm towards a certain entity in the review doesn’t really imply they hated the movie, and that it’s overall a negative score. It just means that they didn’t feel great about an aspect of it, therefore this helps not render the entire review as “Sarcastic”. It helps categorize them into “Actually Positive, Snarky but still Positive, Neutral, Sarcastic but didn’t feel any specific way towards the movie, Ironic and negative, Actually Negative, and such.

By combining VADER’s lexicon-based sentiment analysis with the contextual understanding of a Transformer-based sarcasm detection model, this project achieved an enhanced sentiment analyzing model. The accounting for sarcastic tones improved the reliability and confidence of sentiment scores. It also allowed us to dig deeper on user sentiment beyond a surface-level positive, zero or negative. Both models were computationally efficient and capable of handling large volumes of data, making them suitable for real-time applications especially looking at the value ranges in the engagement in the last few years.

5. Results and Evaluation

Our integrated model has several parameters, which we're using to analyze the movie review data from Letterboxd. There are some elements that are analyzed individually, and there are some that we analyze with respect to each other.
Analyzing Sentiment and Sarcasm in movie reviews help form a good standard, upon which we build with our other metrics.

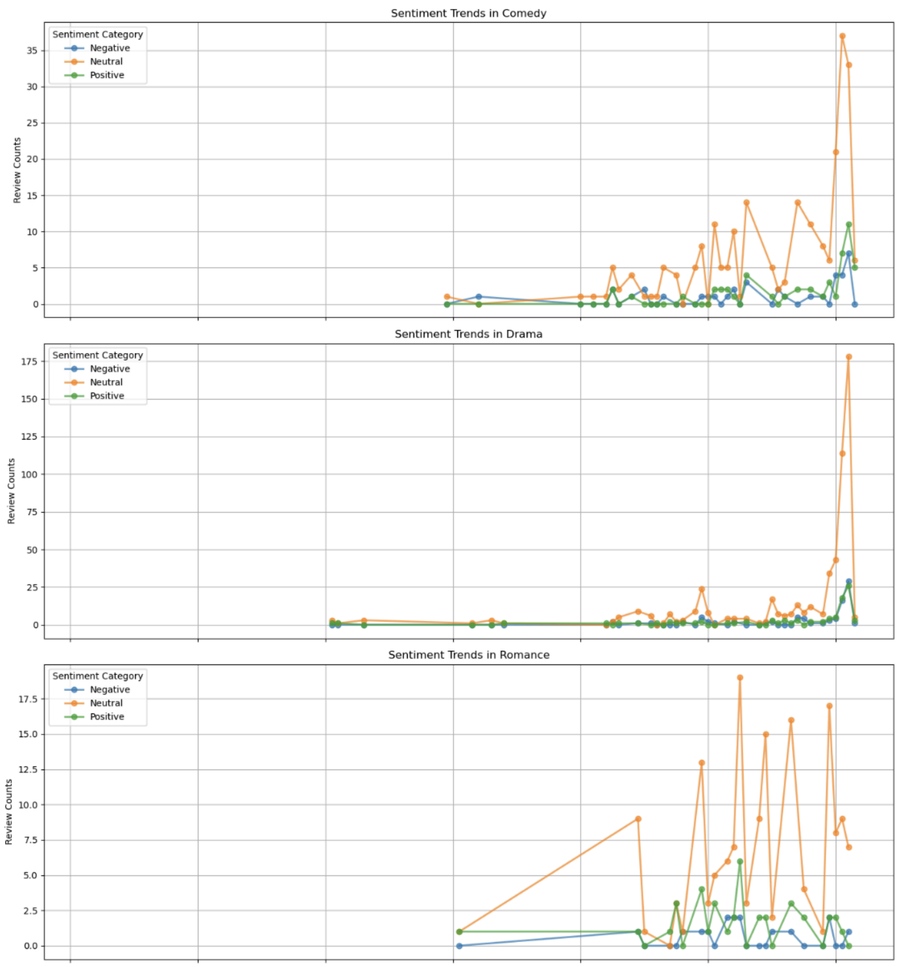
Sentiment Analysis

We used the NLTK VADER Sentiment Analyzer to give us sentiments of movie reviews after performing some processing on the text. Below in the figure is the dataframe, with the Sentiment column added at the end, that shows the overall sentiment of the Reviewer on a certain movie.
For example, a user named Jeff had a Negative Sentiment (Sentiment = – 1) for The Muppets, whilst a reviewer Branson Reese had a neutral review for Beetlejuice (I won't say it three times.) Than Tibbetts' review of Being John Malkovich is a neutral sentiment. The sentiment analysis helps inform us about the overall sentiments of the user on the movie they're reviewing.

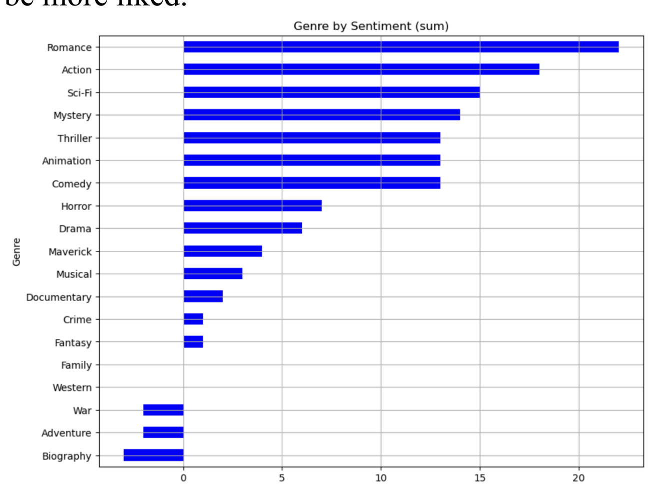
	movie_name	Release Year	Reviewer name	Clean_Review_date	Clean_Review	Clean_Comment Count	Like count	genre	Processed Review	Sentiment
0	Clue	1985	Branson Reese	1996-10-16	My dad got in so much trouble for showing me t...	6	2286	Comedy	dad got much trouble showing kid started sayin...	-1
1	Beetlejuice	1988	Branson Reese	1999-10-21	Thank GOD Tim Burton made this movie in 1988 a...	12	3304	Comedy	thank god tim burton made movie 1988 2008 . im...	0
2	Being John Malkovich	1999	Than Tibbetts	2010-10-04	Malkovich. Malkovich Malkovich Malkovich, Malk...	6	4300	Comedy	malkovich . malkovich malkovich malkovich , ma...	0
3	The Muppets	2011	Jeff	2012-03-06	It's fine if you don't like this movie, but it...	31	0	Comedy	fine like movie , probably mean angry , hate-f...	-1
4	Mysterious Skin	2004	Cole	2012-03-11	This movie is beautiful, captivating, fascinat...	4	623	Drama	movie beautiful , captivating , fascinating , ...	1
...	...	...	...	...	...	...	...	...	...	...
2832	Drive	2011	k??rsten	NaN	Yes, I just saw it for the first timeYes, I lo...	9	2160	Action	yes , saw first timeyes , loved everything ity...	1
2833	Fight Club	1999	hunt??r	NaN	if I was next to brad, I would have dropped th...	19	0	Drama	next brad , would dropped soap	0
2834	The Bling Ring	2013	k??rsten	NaN	not a single good shot or outfit in this entir...	30	0	Crime	single good shot outfit entire thing	1
2835	A Serbian Film	2010	DirkH	NaN	OH MY GOD, LOOK AT HOW CONTROVERSIAL I AM!!!!!!	65	0	Horror	oh god , look controversial ! ! !	-1
2836	CODA	2021	James (Schaffrillas)	NaN	Crazy how such a cliche and predictable movie ...	32	0	Drama	crazy cliche predictable movie could sooooooo ...	1

Genre Analysis

Getting into our results, where we compared different aspects against each other, this first graph shows the different sentiments across specific genres (x axis: release year, y axis: number of reviews), how they usually fare. Across the board, most Letterboxd reviews are usually neutral reviews. But comedies are usually more positively reviewed than they are negative. Interactions with not only recent movies, but also older movies were higher for romance movies, and skewed slightly more positive, than say, drama movies that had a slightly equal distribution of positive to negative sentiments. The code notebook has a larger sect of these trends, covering Horror, Mystery, Fantasy, Crime, War, Musicals, Adventure, Action, Animation etc.

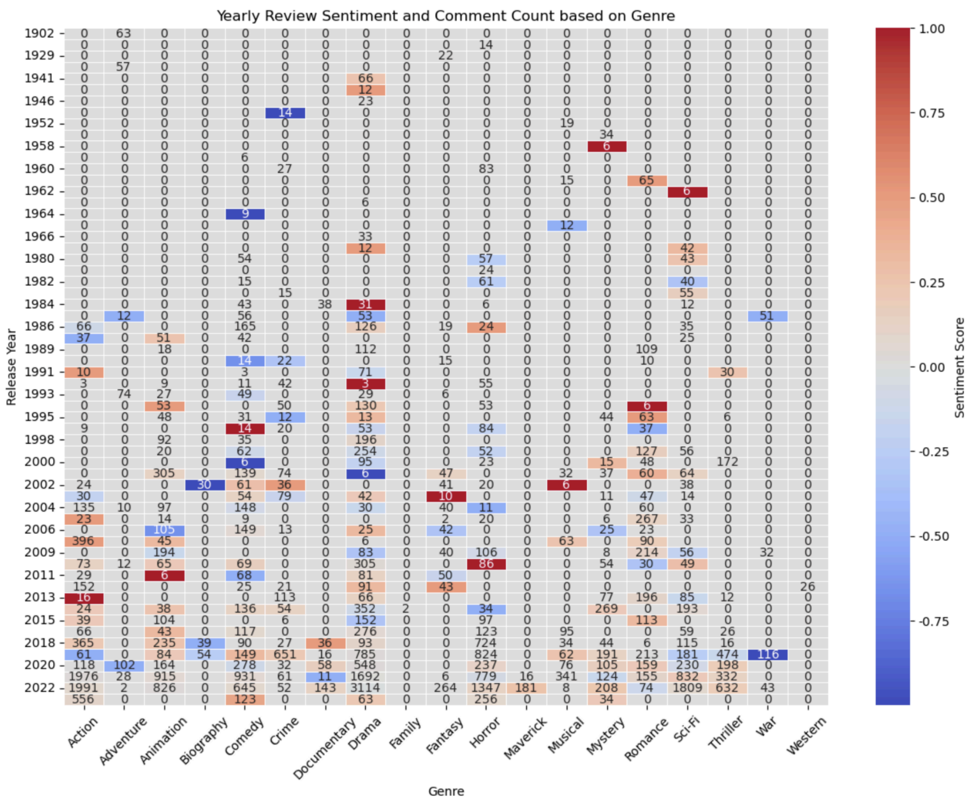


Looking at the overall Genre by Sum of Sentiments graph below, we see an overall pattern in sentiments behind movie reviews across genres, Romance leading the pack. This just means that people like romance movies, that usually have a feel good sensibility attached to it. It doesn't need to go above and beyond for them to like a Romance or an Action movie. But a War or an Adventure Biography might need more crowdpleasing, for the overall sentiments in the reviews and thus the movies to be more liked.



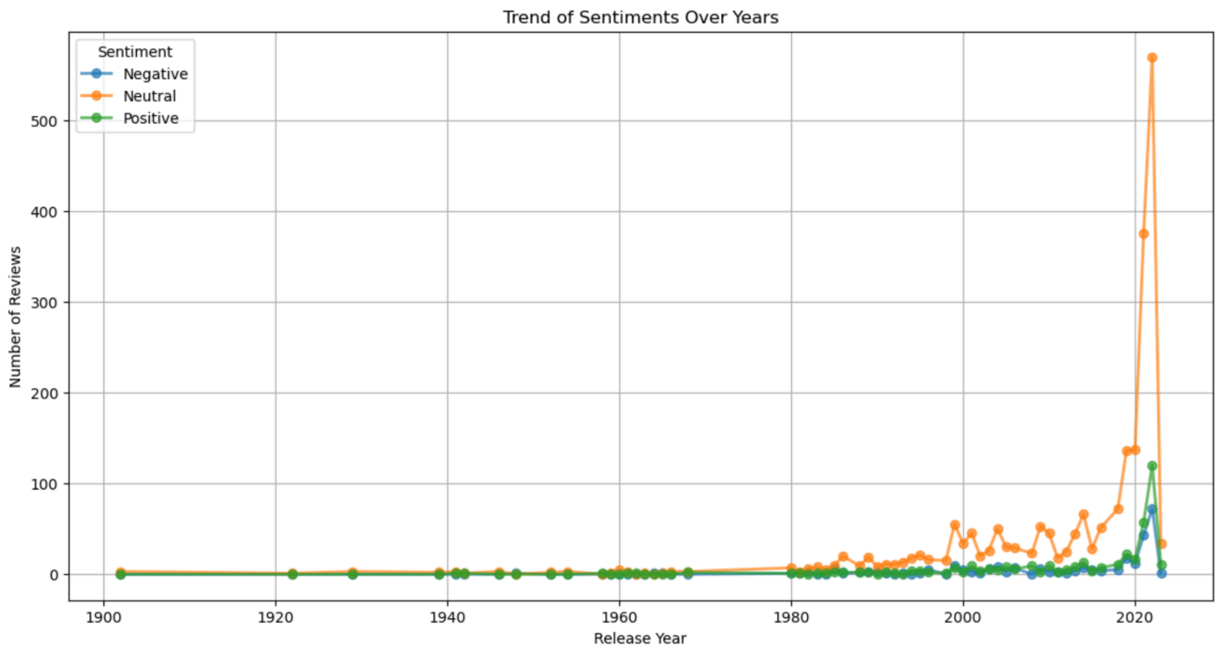
Time Analysis

This does a deeper dive on the release year, and how different movies are reviewed that year, by genre. Drama is the genre that has the highest interactions with older movies. Close followers are Crime, Mystery and Romance. This tracks, because a lot of Hollywood classics are of these genres and this shows that there is still a lot of interest for people to engage with older movies. They're also often reviewed pretty positively, except maybe some genres which may not have aged well with time. We also see the number of reviews yearly increase now, showing that Letterboxd still has an issue with people going back and exploring older classics, as much as they watch recently released movies. This could be an add-on in the future, where the application could not only advertise older classics, and partner with companies that have these movies to watch / stream on their subscription platforms.



User Engagement Analysis

Looking at the graph below, gives us a look on recency, and how engagement has been improving. Great movies have always existed, and some also logged (from the 1900s?), but there was definitely a greater influx of engagement on the internet with movies from the 2000s, with it steadily growing, and movie watchers having stronger sentiments on recently watched movies. The Engagement is directly viewed, with more reviews coming in, and more interactions with these reviews. The trends on the internet usually lean towards neutral with some peaks in positivity, and fewer negative reviews.



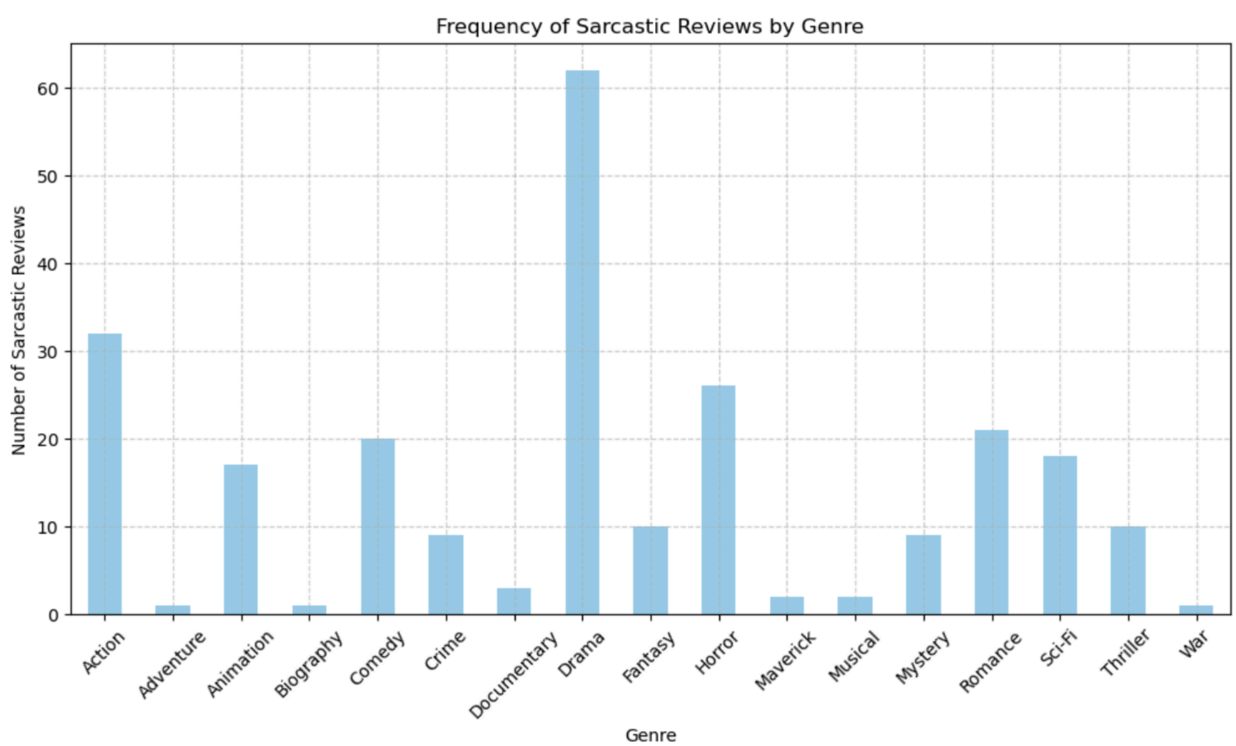
Sarcasm Analysis

We used a Pre-trained Transformer-based model for the Sarcasm Analysis section of the project, to give us a check on whether the movie reviews in question were sarcastic, or earnest. Comparing the same movies we checked individual sentiments for, first, with user Jeff's review of The Muppets. The sentiment returned by just the Sentiment Analysis was negative. This sarcasm has a go at the dataset, and tells us that the review was Not Sarcastic (isSarcastic = False). It's also 0.9903 (99%) confident about this checking. With user Brandon Reese's review of the Beetlejuice (second time), the model says it's Not Sarcastic, and is 85.18% confident of this score. With Than Tibbetts' review of Being John Malkovich, it returns a True (is Sarcastic), and is 57.78% confident of this review being snarky.

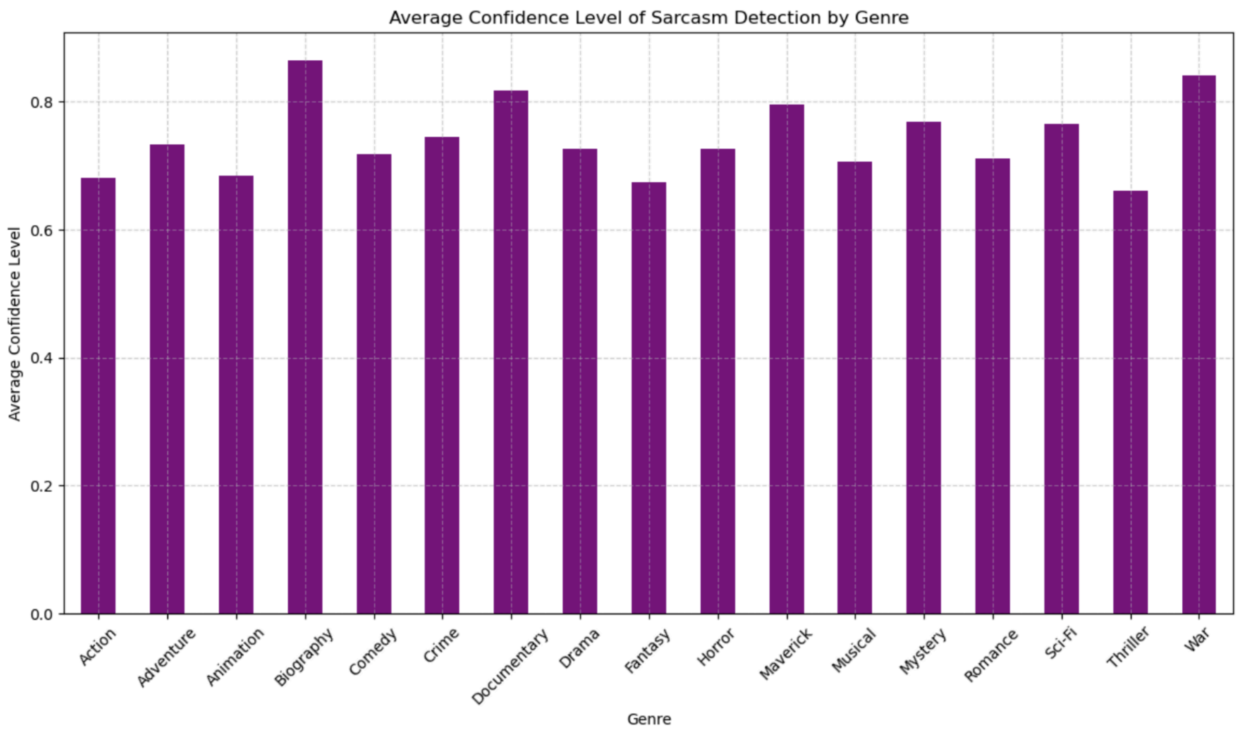
	movie_name	Release Year	Reviewer name	Clean_Review_date	Clean_Review	Clean_Comment Count	Like count	genre	isSarcastic	Confidence
0	Clue	1985	Branson Reese	1996-10-16	My dad got in so much trouble for showing me t...	6	2286	Comedy	False	0.949989
1	Beetlejuice	1988	Branson Reese	1999-10-21	Thank GOD Tim Burton made this movie in 1988 a...	12	3304	Comedy	False	0.851818
2	Being John Malkovich	1999	Than Tibbetts	2010-10-04	Malkovich. Malkovich Malkovich Malkovich, Malk...	6	4300	Comedy	True	0.577804
3	The Muppets	2011	Jeff	2012-03-06	It's fine if you don't like this movie, but it...	31	0	Comedy	False	0.990326
4	Mysterious Skin	2004	Cole	2012-03-11	This movie is beautiful, captivating, fascinat...	4	623	Drama	False	0.994408
...	...	...	...	...	...	...	...	...	...	...
2832	Drive	2011	k??rsten	NaN	Yes, I just saw it for the first timeYes, I lo...	9	2160	Action	False	0.977845
2833	Fight Club	1999	hunt??r	NaN	if I was next to brad, I would have dropped th...	19	0	Drama	False	0.985713
2834	The Bling Ring	2013	k??rsten	NaN	not a single good shot or outfit in this entir...	30	0	Crime	False	0.707220
2835	A Serbian Film	2010	DirkH	NaN	OH MY GOD, LOOK AT HOW CONTROVERSIAL I AM!!!!!!	65	0	Horror	False	0.967895
2836	CODA	2021	James (Schaffrillas)	NaN	Crazy how such a cliche and predictable movie ...	32	0	Drama	False	0.894006

I was curious to see if the last mentioned review actually was snarky. I looked up on Letterboxd Than Tibbett's review for Being John Malkovich and this is what it showed. It is indeed snarky / ironic.

Getting into relative analyses, in the first graph below, we see that the frequency of sarcastic reviews by genres are also very split, meaning that there are some reviews that are definitely more sarcastic leaning in a genre, than others. From the graph, we see that drama movie reviews from our dataframe have shown to be most sarcastic, with a little over sixty sarcastic reviews out of a little under three thousand reviews. Second most are action movies, while War movies tend to be least sarcastic, followed by Adventure and Biography. If you'd notice, this is a pattern. From the relative graph in Genre Analysis, we see these genres be the most negatively reviewed. This could imply that they could be the review sentiments that are most skewed. The negative sentiments from the graphs could mostly be sarcastic ones, probably because the reviewers wanted to juxtapose the very sombre emotional content with some jokes, or snark.



In the graph below, we see the average confidence level of Sarcasm Detection split by Genre. This tells us across the board beyond 70% confidence in the model's detection of sarcasm in movie reviews. There are smaller flourishes, like in Biographies, War (also part of the least sarcastic if you'd compare graphs), which prove that we're confident (Over 90% confident) that they're earnest reviews, while other genres can also be compared, like Horror (around ~26 sarcastic reviews in the corpus), we're confident to about 70% of the sarcasm rates in the reviews.



This way, across reviews, we're able to not only analyze them for their trends, and engagement statuses, but also give a confidence score to reassure other programmers of the ability of the model to identify an earnest review from a snarky one (Markovich Markovich, Markovich? MARKOVICHHHHHH!!) and also have a project that is able to discern a positive review from a negative one.

6. Conclusion and Future Scope

This work opens up a number of exciting directions for future research. While multi-language support would increase the system's usefulness for audiences around the world, the framework might be improved by incorporating more advanced NLP models that can comprehend progressively complicated linguistic patterns. Developing aspect-based sentiment analysis to offer more detailed insights about particular aspects of movies, such acting, directing, or cinematography, also has a lot of promise. The analytical platform could be further enhanced by integrating external data sources, like box office receipts and critic reviews, and implementing real-time analysis capabilities.

In addition to advancing our knowledge of sentiment analysis on social media, this study offers useful advice to platforms looking to improve user experiences. The approaches and results offered here provide useful foundations for further study and advancement in digital content analysis as social media continues to expand as the main platform for cultural debate.

Other Future Prospects

Enhanced Sarcasm Identification

– To increase accuracy and better capture complex behaviors in reviews, we could make use of deep learning models like Transformers (e.g., BERT, RoBERTa) that have been adjusted for sarcasm detection using the latest updates.

Development of Recommender Systems

– We could create a customized movie recommendation system by utilising more engagement data like likes and comments and trends unique to a given genre. For better recommendations, the algorithm might take user preferences, sentiment trends, and interactions into account.

Analysis Across Several Platforms

– In order to analyse user behaviour and sentiment patterns across platforms, we could expand the investigation to other social media blog websites such as Twitter, Reddit, or IMDb. This might offer a more comprehensive perspective on online movie comments.

Temporal Changes in Sentiments

– Potentially analyse how opinions and sarcasm in reviews change over time by conducting longitudinal studies, especially for films with cult followings or during award seasons.

Combining Multimedia Information

– Examine any audio-visual media shared with reviews, such as GIFs, pictures, or memes, to fully understand user humour and expressions, integrate with sentiments and sarcasm checker.

Real-Time Applications and Insights

– Create a dashboard that further can track popular films, genres, and user engagement in real time. A smaller version exists, like Nielson Streaming Records. But studios, distributors, and streaming services may find this helpful in determining how the public will respond to their media.

7. References

A. Joshi, P. Bhattacharyya, and M. Palshikar, "Sarcasm detection: A comparative study," ArXiv, 2021. [Online]. Available: https://ar5iv.labs.arxiv.org/html/2107.02276

M. S. Hossen and N. R. Dev, "Sentiment Analysis of Social Media Reviews: A Case Study," Inspirajournals, 2021. [Online]. Available: https://www.inspirajournals.com

S. Chirgaiya, K. Dashtipour, and P. Shah, "Sentiment Analysis on Movie Reviews: A Comparative Analysis," Elsevier, 2021. [Online]. Available: https://www.sciencedirect.com

Letterboxd, "Letterboxd - Social film discovery," 2024. [Online]. Available: https://letterboxd.com

C.J. Hutto and E. Gilbert, "VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text," Proceedings of ICWSM, 2014. [Online]. Available: https://github.com/cjhutto/vaderSentiment

K. Chakraborty, S. Bhattacharyya, and R. Bag, "A Survey of Sentiment Analysis from Social Media Data," IEEE Transactions, 2020. [Online]. Available: https://ieeexplore.ieee.org

N. Kumar and D. Sharma, "Movie Review Analysis Using Deep Learning Techniques," International Journal of Computer Applications, 2022. [Online]. Available: https://www.ijcaonline.org

Y. Liu et al., "RoBERTa: A Robustly Optimized BERT Pretraining Approach," ArXiv, 2019. [Online]. Available: https://arxiv.org/abs/1907.11692

T. Wolf et al., "Transformers: State-of-the-Art Natural Language Processing," EMNLP, 2020. [Online]. Available: https://aclanthology.org

R. Feldman, "Techniques and Applications for Sentiment Analysis," Communications of the ACM, 2019. [Online]. Available: https://dl.acm.org

D. Jurafsky and J. H. Martin, "Speech and Language Processing," Stanford University, 2023. [Online]. Available: https://web.stanford.edu/~jurafsky/slp3/

M. Zhang and B. Liu, "Social Media Text Classification with Limited Training Data," ACM Digital Library, 2021. [Online]. Available: https://dl.acm.org