Letterboxd

Movie Review Analysis for Letterboxd

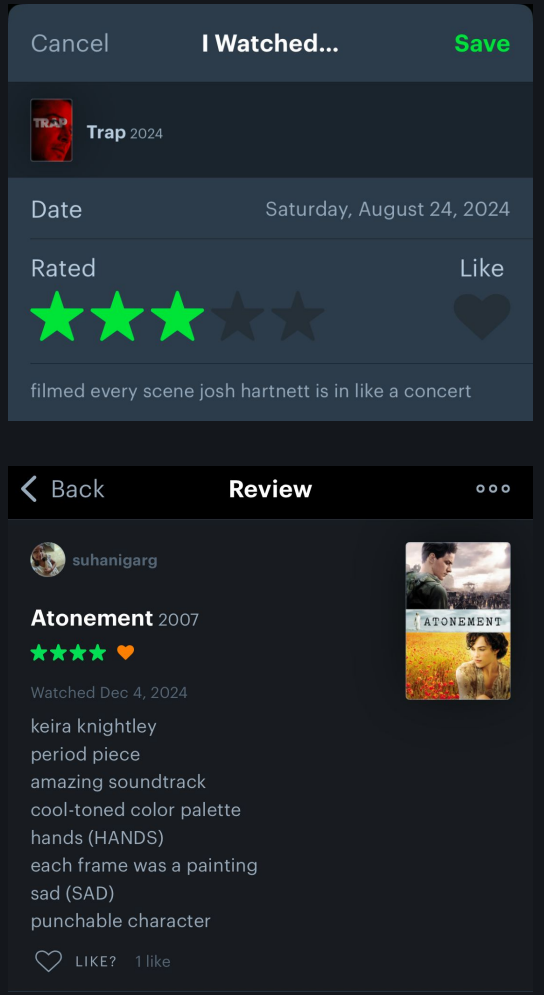
Adithya Neelakantan (SUID: 663182945)
Aditi Mangesh Jadhav (SUID: 544349325)
Pratik Satpute (SUID: 566041095)
Shruti Vasave (SUID: 427009403)
Jeevan Tumkur Venkatesh (SUID: 740874557)

Letterboxd

What is Letterboxd?

A social media application to let you share your love for film and blog your movie experience. Letterboxd has tools to let you –

- Keep track of every film you've watched.
- Write and share reviews, and follow friends and interact with others movie reviews.
- Rate each film on a five-star scale.
- Keep a diary of your film watching.
- Compile, share lists and recommendations of films on any genre or topic and keep a watchlist of films to see.



Letterboxd

Internet discourse about Letterboxd

Letterboxd

Two level approach + Reasoning

- Sentiment Analysis
- Sarcasm Check
- Validation of the Sentiments to distinguish between Actual Positive, Positive but ironic (neutral or negative), Actual Neutral, Actual Negative, Negative but snarky (negative, neutral or positive).

Letterboxd

Types of analysis

- Sentiment Analysis – Is the review a positive, negative, or neutral one.
- Engagement Analysis – How are reviews interacted with, by movie and genre.
- Time Analysis – Is there a trend with a certain genre of movie over time? More engagement by time?
- Genre Analysis – Are certain genres “in” right now, did a year be a great moments for a genre of movies?
- Sarcasm Analysis – Are the movie reviews snarky, contain irony? Is it an earnest review or just made for fun.

Letterboxd

Data and Preprocessing

– Tokenization:
Splits text into individual words or tokens using TweetTokenizer.

– Stop Words Removal:
Filters out common words (e.g., the, and, is) to retain meaningful words.

– Lemmatization:
Reduces words to their base form using WordNetLemmatizer (e.g., movies → movie).

– Joining Tokens:
Combines the processed words back into a string for further analysis.

Letterboxd

Our Model

Sentiment Analysis – Lexicon-based using NLTK's VADER sentiment analyzer.

VADER Scoring – Calculates scores for negative, neutral, and positive sentiments for each review.

(1 for positive, 0 for neutral, -1 for negative).
Determines sentiment based on the highest score among pos, neu, and neg.

Sarcasm Analysis – Extend sentiment analysis by detecting sarcasm in movie reviews to enhance the understanding of user sentiments and engagement.

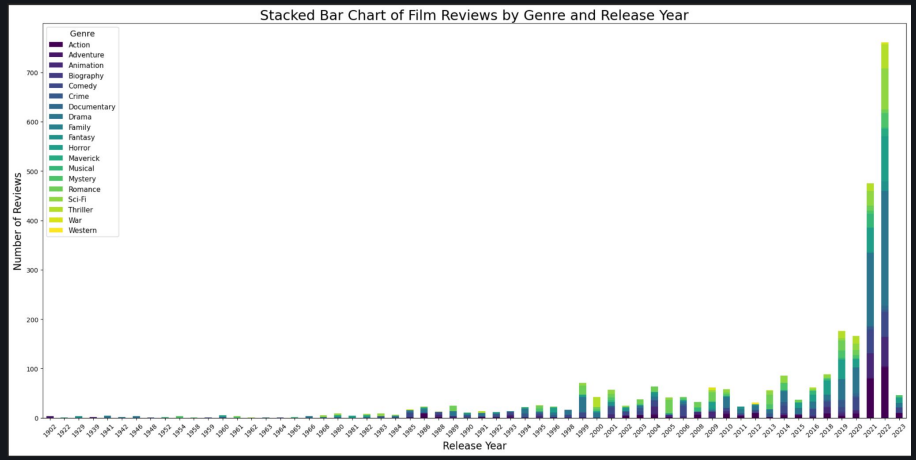
Used a pre-trained Transformers-based sarcasm detection model for accurate classification. Softmax to compute probabilities for sarcastic (1) and non-sarcastic (0) labels.

Model-Name	F1	Precision	Recall	Accuracy
helinivan/english-sarcasm-detector	**92.38**	92.75	92.38	92.42

Letterboxd

Results and Evaluation

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

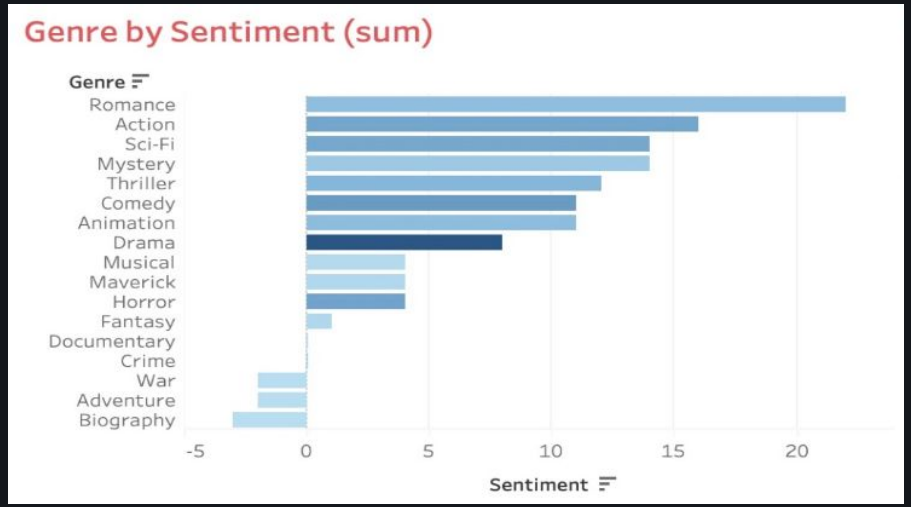


Sentiment Analysis

Engagement and Time Analysis

Letterboxd

Results and Evaluation



Genre Analysis

	movie_name	Release Year	Reviewer name	Clean_Review_date	Clean_Review	Clean_Comment Count	Like count	genre	Processed Review	Sarcasm_Detection
0	Clue	1985	Branson Reese	1996-10-16	My dad got in so much trouble for showing me t...	6	2,286 likes	Comedy	dad got much trouble showing kid started sayin...	{'is_sarcastic': 'False', 'confidence': 0.9499...
1	Beetlejuice	1988	Branson Reese	1999-10-21	Thank GOD Tim Burton made this movie in 1988 a...	12	3,304 likes	Comedy	thank god tim burton made movie 1988 2008 . im...	{'is_sarcastic': 'False', 'confidence': 0.8518...
2	Being John Malkovich	1999	Than Tibbetts	2010-10-04	Malkovich. Malkovich Malkovich Malkovich, Malk...	6	4,300 likes	Comedy	malkovich . malkovich malkovich malkovich , ma...	{'is_sarcastic': 'True', 'confidence': 0.57780...
3	The Muppets	2011	Jeff	2012-03-06	It's fine if you don't like this movie, but it...	31	NaN	Comedy	fine like movie , probably mean angry , hate-f...	{'is_sarcastic': 'False', 'confidence': 0.9903...
4	Mysterious Skin	2004	Cole	2012-03-11	This movie is beautiful, captivating, fascinat...	4	6 23 likes	Drama	movie beautiful , captivating , fascinating , ...	{'is_sarcastic': 'False', 'confidence': 0.9944...
...	...	...	...	...	...	...	...	...	...	...
2832	Drive	2011	k??rsten	NaN	Yes, I just saw it for the first timeYes, I lo...	9	2,160 likes	Action	yes , saw first timeyes , loved everything ity...	{'is_sarcastic': 'False', 'confidence': 0.9778...
2833	Fight Club	1999	hunt??r	NaN	if I was next to brad, I would have dropped th...	19	NaN	Drama	next brad , would dropped soap	{'is_sarcastic': 'False', 'confidence': 0.9857...
2834	The Bling Ring	2013	k??rsten	NaN	not a single good shot or outfit in this entir...	30	NaN	Crime	single good shot outfit entire thing	{'is_sarcastic': 'False', 'confidence': 0.7072...
2835	A Serbian Film	2010	DirkH	NaN	OH MY GOD, LOOK AT HOW CONTROVERSIAL I AM!!!!!!	65	NaN	Horror	oh god , look controversial ! ! !	{'is_sarcastic': 'False', 'confidence': 0.9678...
2836	CODA	2021	James (Schaffrillas)	NaN	Crazy how such a cliche and predictable movie ...	32	NaN	Drama	crazy cliche predictable movie could soooooooo ...	{'is_sarcastic': 'False', 'confidence': 0.8940...

Sarcasm Analysis

Conclusion and Future Scope

To analyze Movie reviews to get a few stats about genres, engagement, sentiments, and sarcasm.

Employed Python's NLTK Vader sentiment analyzer for lexicon-based Sentiment scoring (-1, 0, 1), and a Sarcasm Analysis feature.

Investigates user interactions with the reviews over time to identify patterns in activity and preferences.

Explores genre-specific trends, which could help act as a precursor for a recommender system, based on a user's interactions.