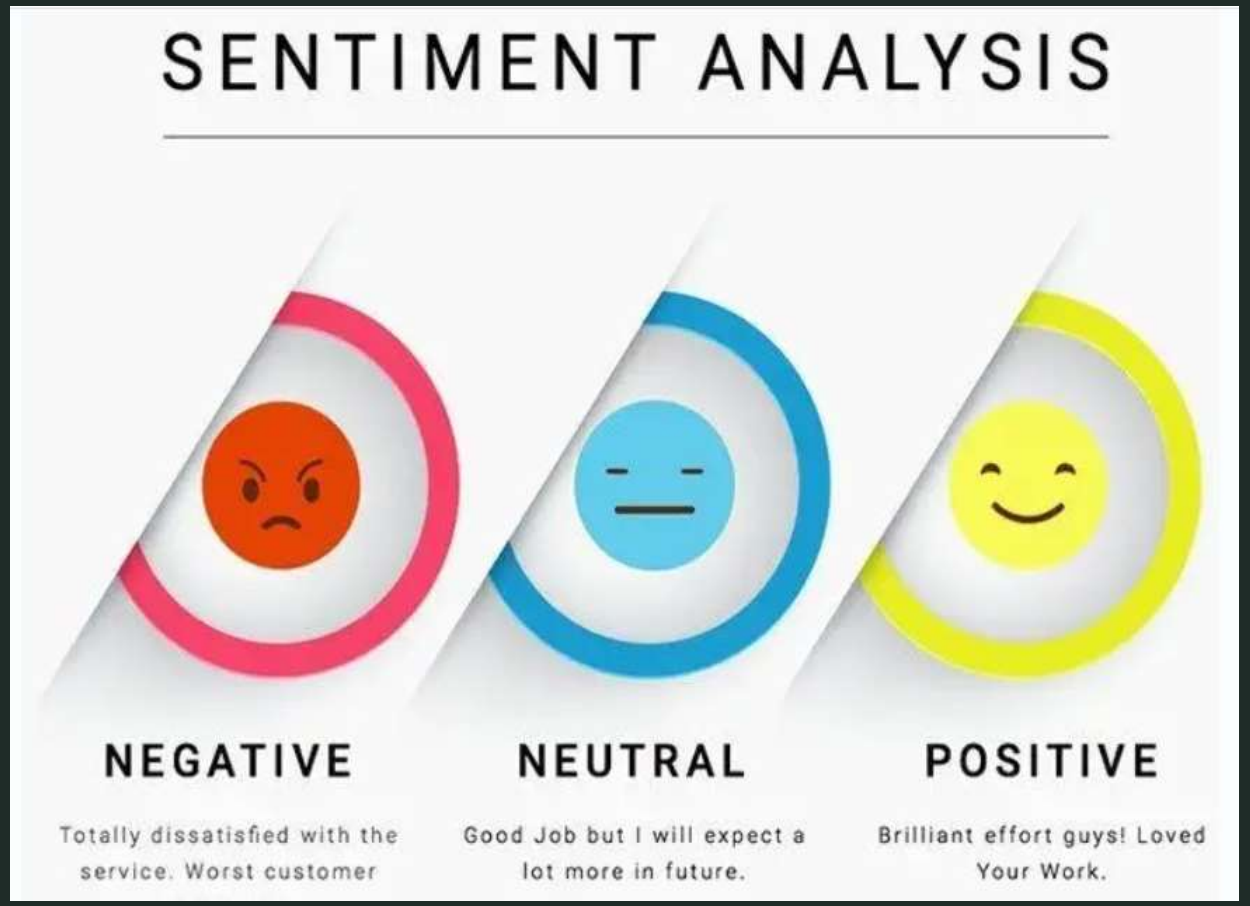
SENTIMENT ANALYSIS OF MOVIE REVIEWS

Submitted by:
Prince Kumar (202410116100149)
Neha Agnihotri(202410116100131)
Samiksha Teotia (202410116100178)
Rahul (202410116100157)

Agenda

1. INTRODUCTION
2. APPLICATION
3. WORK FLOW
4. TECHNIQUES
5. SAMPLE CODE
6. SCREENSHOTS
7. CHALLENGES
8. CONCLUSION

INTRODUCTION



- A method to detect emotions in text
- Also called *opinion mining*
- Commonly used to classify text as:
  - Positive
  - Negative
  - Neutral

ANALYSIS OF MOVIE REVIEWS

- Sentiment analysis is like teaching a computer to read people's comments, reviews, or posts — and understand how they feel about something.
- Sentiment analysis, also known as opinion mining, is a natural language processing (NLP) technique used to determine the emotional tone behind a body of text.
- In the context of movie reviews, sentiment analysis aims to automatically identify whether a review expresses a positive, negative, or neutral opinion about a film.
- By using machine learning algorithms or deep learning models such as Naive Bayes, Support Vector Machines (SVM), or LSTM neural networks, systems can be trained to classify sentiments with high accuracy..

Applications of Sentiment Analysis in Movie Reviews

1.Audience Feedback Analysis
Sentiment analysis helps movie producers and distributors understand how audiences are reacting to a movie.— such as the acting, storyline, direction, or music.

2.Movie Recommendation Systems
Platforms like Netflix, Amazon Prime, and Disney+ use sentiment analysis to improve recommendations. If a user likes movies that get a lot of positive emotional feedback, the system can suggest similar films. This improves personalization and enhances user experience.

3.Box Office Prediction
Before or shortly after a movie's release, early reviews and online discussions can be analyzed to predict how well a movie will perform. If sentiment is mostly positive, it’s likely to attract more viewers, which can indicate a higher box office return.

4.Marketing Strategy Optimization
Studios monitor the audience’s reaction to trailers, posters, interviews, and promotional campaigns.

5.Competitor Analysis
Sentiment analysis allows studios to compare how their movie is performing emotionally compared to competitors. For example, if two action movies release at the same time, comparing audience sentiments helps understand which one resonates more and why.

HOW IT WORKS..

1. **COLLECTION REVIEWS:** The first step is gathering text data — movie reviews in this case. Reviews can be collected from platforms like- google reviews, twitter.

2. **TEXT PREPROCESSING :** Before feeding the data to a model, it must be cleaned and standardized like- remove punctuation ,lowercasing.

3. FEATURE EXTRACTION : Since ML models can't understand text directly, we convert it into numbers.

4. **CLASSIFICATION:** we train a **Machine Learning** or **Deep Learning** model like- naïve bayes , long short-term memory(handles sequence of words).
5. **PREDICT SENTIMENT:** Once trained, the model can predict the sentiment of new reviews:

- "The movie was brilliant!" → **Positive**
- "It was a waste of time." → **Negative**
- "It was okay, not great but not bad either." → **Neutral**

TECHNIQUES

1. **MACHINE LEARNING TECHNIQUES**

- **Naive Bayes-A** simple and fast algorithm based on probability.
- **Support Vector Machine (SVM)-** A powerful classifier that tries to find the best boundary between different classes (e.g., positive vs. negative)

2. **DEEP LEARNING TECHNIQUES**

- **LSTM (Long Short-Term Memory)-**Good at understanding sequences of words (like sentences)

Remembers long-term dependencies in text (e.g., “not good” vs. “good”)

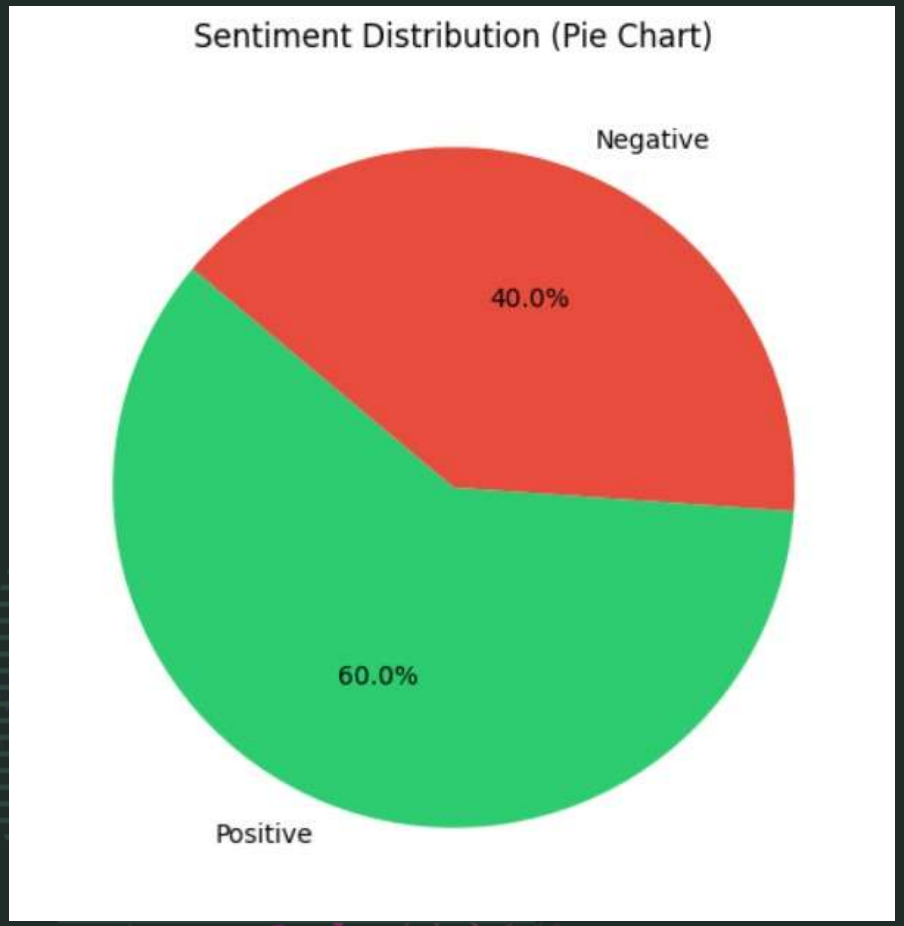
- **CNN (Convolutional Neural Network)-**Fast and efficient for sentence-level classification
- **BERT (Bidirectional Encoder Representations from Transformers) -**Pre-trained model from Google

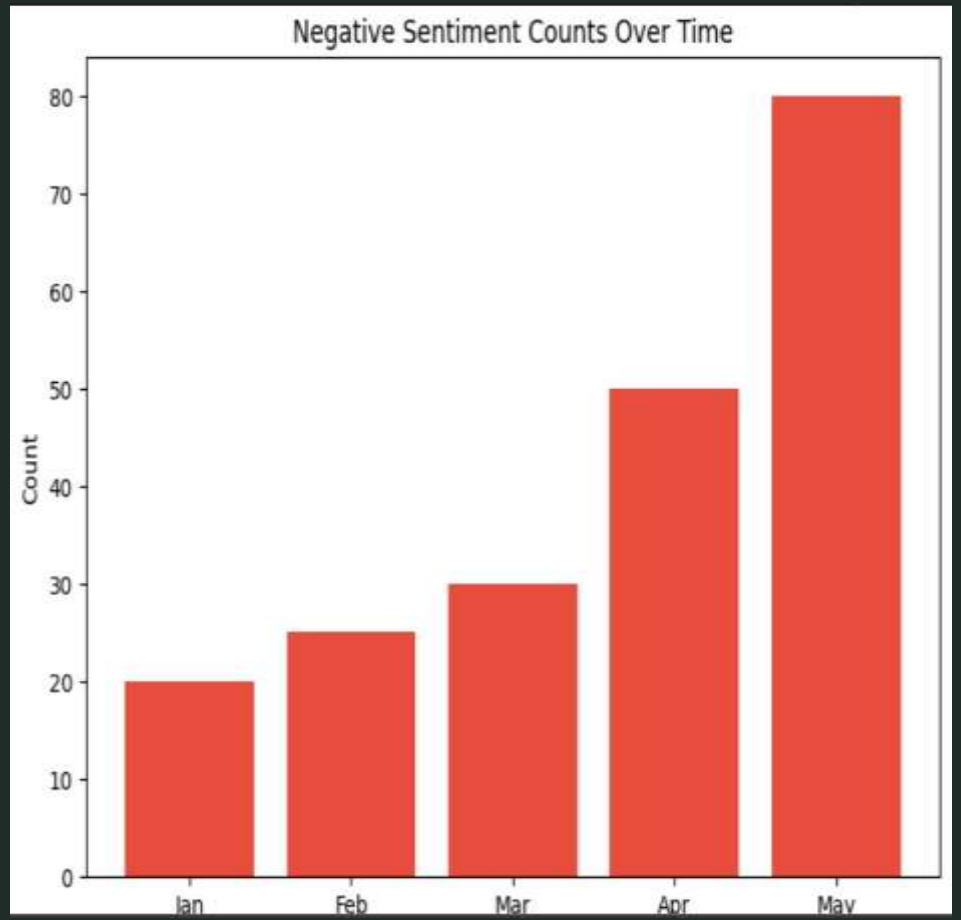
SCREENSHOTS..

```
[nltk_data] Downloading package movie_reviews to /root/nltk_data...
[nltk_data]   Package movie_reviews is already up-to-date!
[nltk_data] Downloading package stopwords to /root/nltk_data...
[nltk_data]   Package stopwords is already up-to-date!
[nltk_data] Downloading package wordnet to /root/nltk_data...
[nltk_data]   Package wordnet is already up-to-date!
============================================================
🎯 Sentiment Analysis on Movie Reviews
============================================================
✅ Model Accuracy: 80.00%

📌 Top Informative Features:
Most Informative Features
   contains(outstanding) = True              pos : neg    =     14.1 : 1.0
         contains(mulan) = True              pos : neg    =      8.3 : 1.0
        contains(seagal) = True              neg : pos    =      7.8 : 1.0
   contains(wonderfully) = True              pos : neg    =      6.6 : 1.0
         contains(damon) = True              pos : neg    =      6.3 : 1.0
        contains(poorly) = True              neg : pos    =      5.7 : 1.0
         contains(awful) = True              neg : pos    =      5.6 : 1.0
        contains(wasted) = True              neg : pos    =      5.6 : 1.0
          contains(lame) = True              neg : pos    =      5.4 : 1.0
    contains(ridiculous) = True              neg : pos    =      4.8 : 1.0

📝 Test on Custom Review
============================================================
👉 Enter your movie review:
```

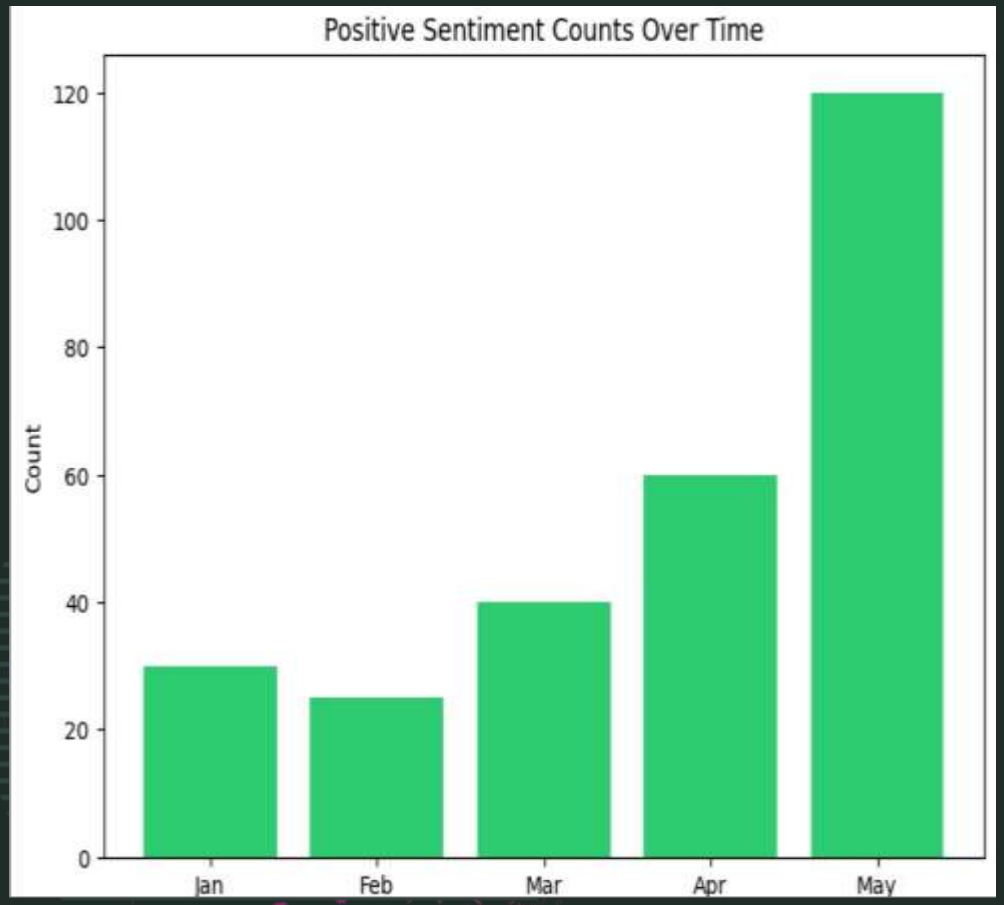
```
[nltk_data]   Package movie_reviews is already up-to-date!
[nltk_data] Downloading package stopwords to /root/nltk_data...
[nltk_data]   Package stopwords is already up-to-date!
[nltk_data] Downloading package wordnet to /root/nltk_data...
[nltk_data]   Package wordnet is already up-to-date!
============================================================
🎯 Sentiment Analysis on Movie Reviews
============================================================
✅ Model Accuracy: 80.00%

📌 Top Informative Features:
Most Informative Features
   contains(outstanding) = True              pos : neg    =     14.1 : 1.0
         contains(mulan) = True              pos : neg    =      8.3 : 1.0
        contains(seagal) = True              neg : pos    =      7.8 : 1.0
   contains(wonderfully) = True              pos : neg    =      6.6 : 1.0
         contains(damon) = True              pos : neg    =      6.3 : 1.0
        contains(poorly) = True              neg : pos    =      5.7 : 1.0
         contains(awful) = True              neg : pos    =      5.6 : 1.0
        contains(wasted) = True              neg : pos    =      5.6 : 1.0
          contains(lame) = True              neg : pos    =      5.4 : 1.0
    contains(ridiculous) = True              neg : pos    =      4.8 : 1.0

📝 Test on Custom Review
============================================================
👉 Enter your movie review: This movie was terrible. The plot was confusing and the characters were dull

🗣️ Predicted Sentiment: Negative 👎
```

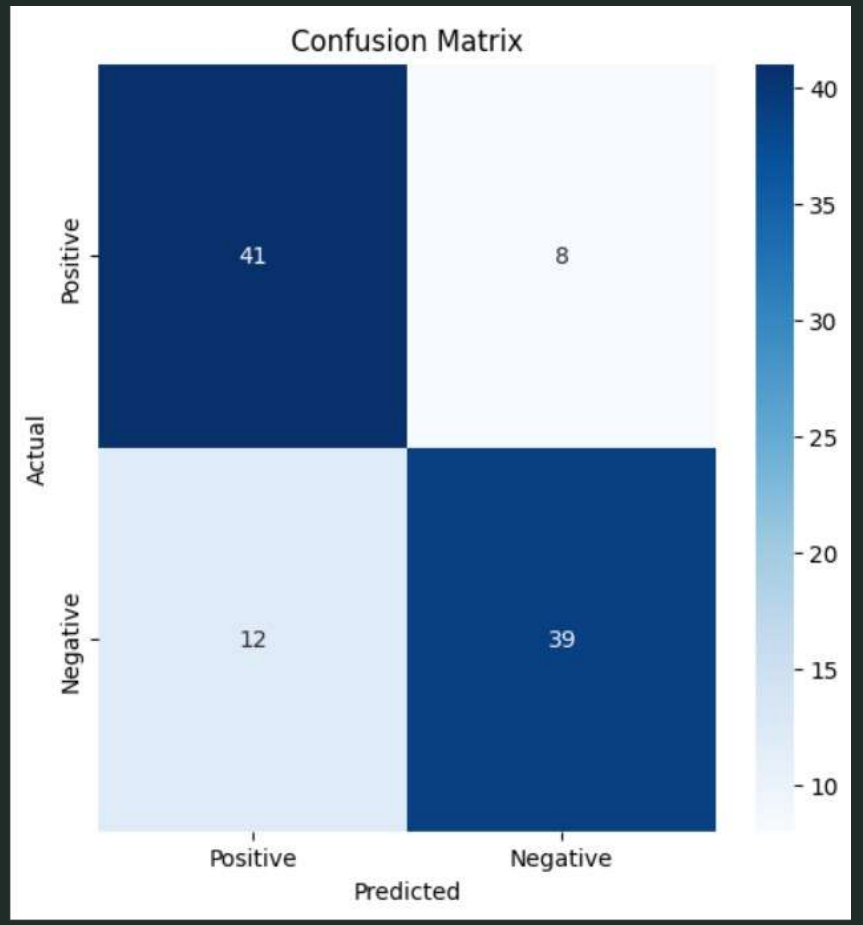
CHART ANALYSIS..





PIE CHARTS..


Positive Sentiment Counts Over Time
Count
0
20
40
60
80
100
120
Jan
Feb
Mar
Apr
May


Confusion Matrix
Actual
Positive
Negative
41
8
12
39
Positive
Negative
Predicted
40
35
30
25
20
15
10

Summary

1. Understanding Audience Emotions- Sentiment analysis helps identify how audiences *feel* about a movie — whether they love it, hate it, or feel neutral.

2. Saves Time and Effort-Manually reading thousands of reviews is impossible.

- Sentiment analysis automates this process, giving quick overviews and trends in minutes.

3. Useful for Decision-Making-Studios, marketers, and streaming platforms can use this data to:

- Improve future content
- Tailor marketing campaigns
- Recommend movies based on mood or themes

4. Enhances User Experience-Recommendation systems become smarter and more personalized.

Viewers can get movie suggestions that match their tastes and emotional preferences.

5. Expanding Use in Industry-This technique isn't just for movies. It's also used in:

- Product reviews
- Social media monitoring
- Brand analysis
- Customer feedback analysis

Thank you