Sentiment Analysis

Course: Data Analysis and Machine Learning with Python

# Sentiment Analysis of popular Sequels based on Reddit Comments:

## Attempting to understand the performance of sequels over time

**Group A**

**Lillian, Anh, Maëlys, Ewald, Sora**

# Table of Content

# Introduction

**Research Question**

How have **sentiments** towards the Pirates of the Caribbean and Twilight movies **changed overtime** and does it have any **correlation with box office** overall ratings?

**Research Rationale**

We focused on Pirates of the Caribbean and Twilight due to their five movies each and **popularity, ensuring sufficient data**.



**Research Goal**

Our goal was to understand how **sequels are viewed** today **and if there's a clear change in sentiments over time.**

# Related Works

Sentiment analysis can be approached in two main ways: **lexicon-based tools** or **machine learning models**.

**Tonia & Rasha (2020)**
**Ricardo et al. (2024)**

We decided to use **VADER** as our lexicon-based baseline model, based on existing research and examples and due to its **consistency** in "microlog-like contexts" **like Reddit comments.**

**Hossen & Dev (2021)**

Highlighted that traditional lexicon-based tools, while simple, lack nuance and context, **recommending more advanced models.**

**Chang et al. (2018)**

Pre-trained BERT models can be fine-tuned with only an additional output layer. **Advancements in transformer-based pre-trained models show that they have received better results in text classification tasks**

**Chedia Dhaoui et al. (2017)**

Found no **significant difference in accuracy** between lexicon-based and machine learning approaches, **suggesting a hybrid approach.**

**J. Sai Teja (2018)**

Studies on movie reviews using Twitter's API showed **varied effectiveness for machine learning models, with SVM and Logistic Regression** being identified as effective in different contexts

# Data Description

## Data Collection

**Reddit comments for each movie** in the chosen sequels using **Reddit API**, focusing on **precise subreddits**.

**Box office** gross revenues **for each movie in the sequels. Required for** correlation analysis

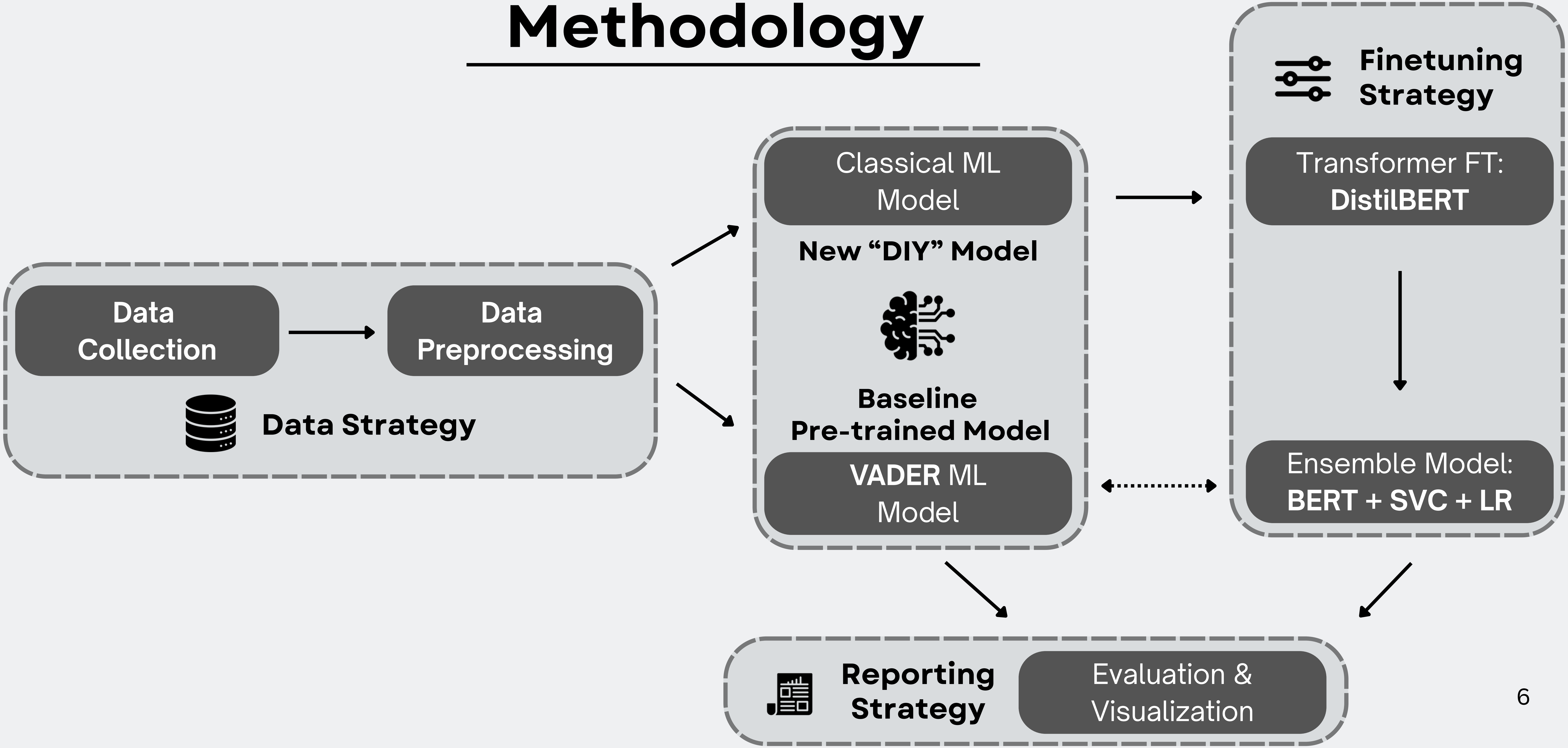
**Rotten Tomatoes' film ratings** for subsequent analysis

## Data Preprocessing

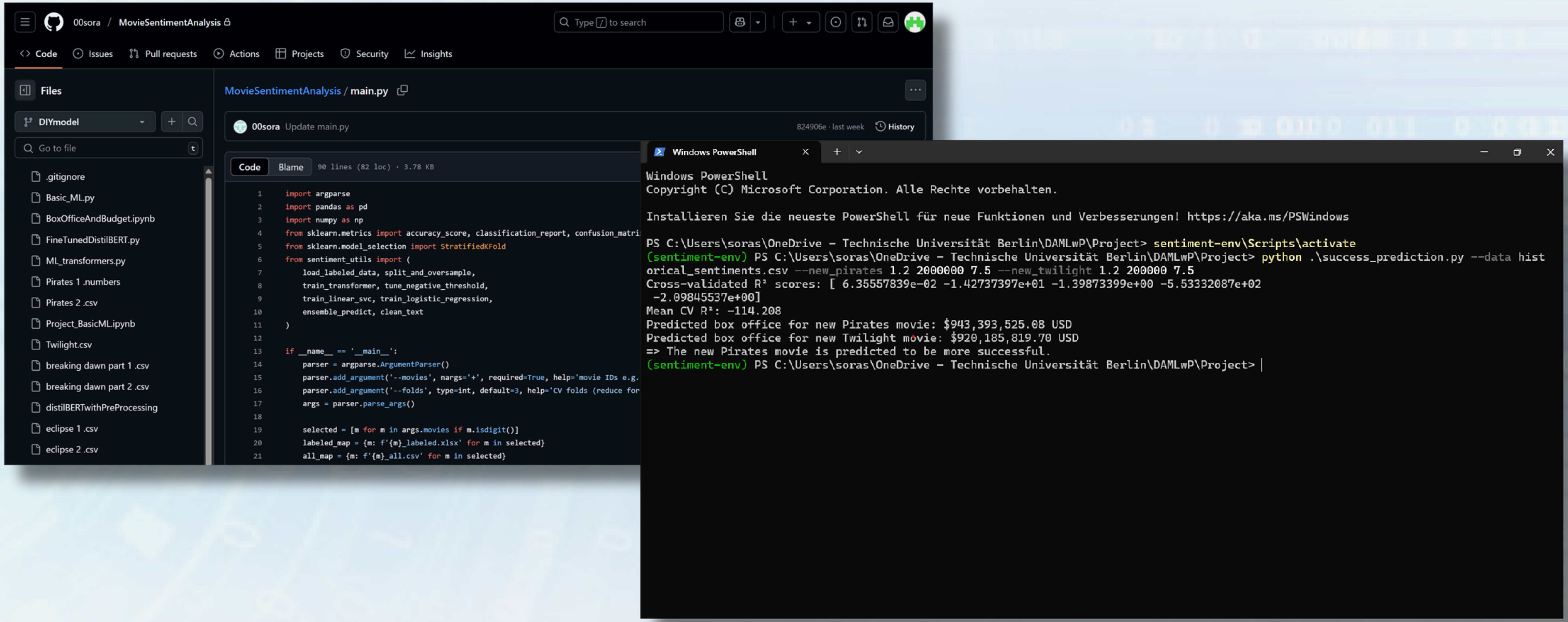
- **Remove** URL links, symbols, stop words or common words, **translate text** into English language
- **Split dataset** into train set and test set, including **oversampling** in order to **address imbalance**
- Text comments are **tokenized, trained, and tuned** in order to fit the classic model training process.

## Metho

# Methodology



# Demonstration



# Results: Model Performance Summary

## Custom Ensemble Model

| | Pred: Negative | Pred: Neutral | Pred: Positive |
|---|---|---|---|
| Actual Neg | 72 | 42 | 27 |
| Actual Neu | 38 | 108 | 36 |
| Actual Pos | 46 | 43 | 99 |

**Best overall performance on neutral sentiment**

| Class | Precision | Recall | F1-score | Support |
|---|---|---|---|---|
| Negative | 0.46 | 0.51 | 0.48 | 141 |
| Neutral | 0.56 | 0.59 | 0.58 | 182 |
| Positive | 0.61 | 0.53 | 0.57 | 188 |

**Overall Accuracy: 54.6%**

**Precision is highest in predicting positive and lowest precision in negative**

## VADER Baseline Model

| | Pred: Positive | Pred: Neutral | Pred: Negative |
|---|---|---|---|
| Actual Pos | 778 | 89 | 197 |
| Actual Neu | 456 | 339 | 248 |
| Actual Neg | 354 | 73 | 341 |

**Bias toward predicting positive sentiment**

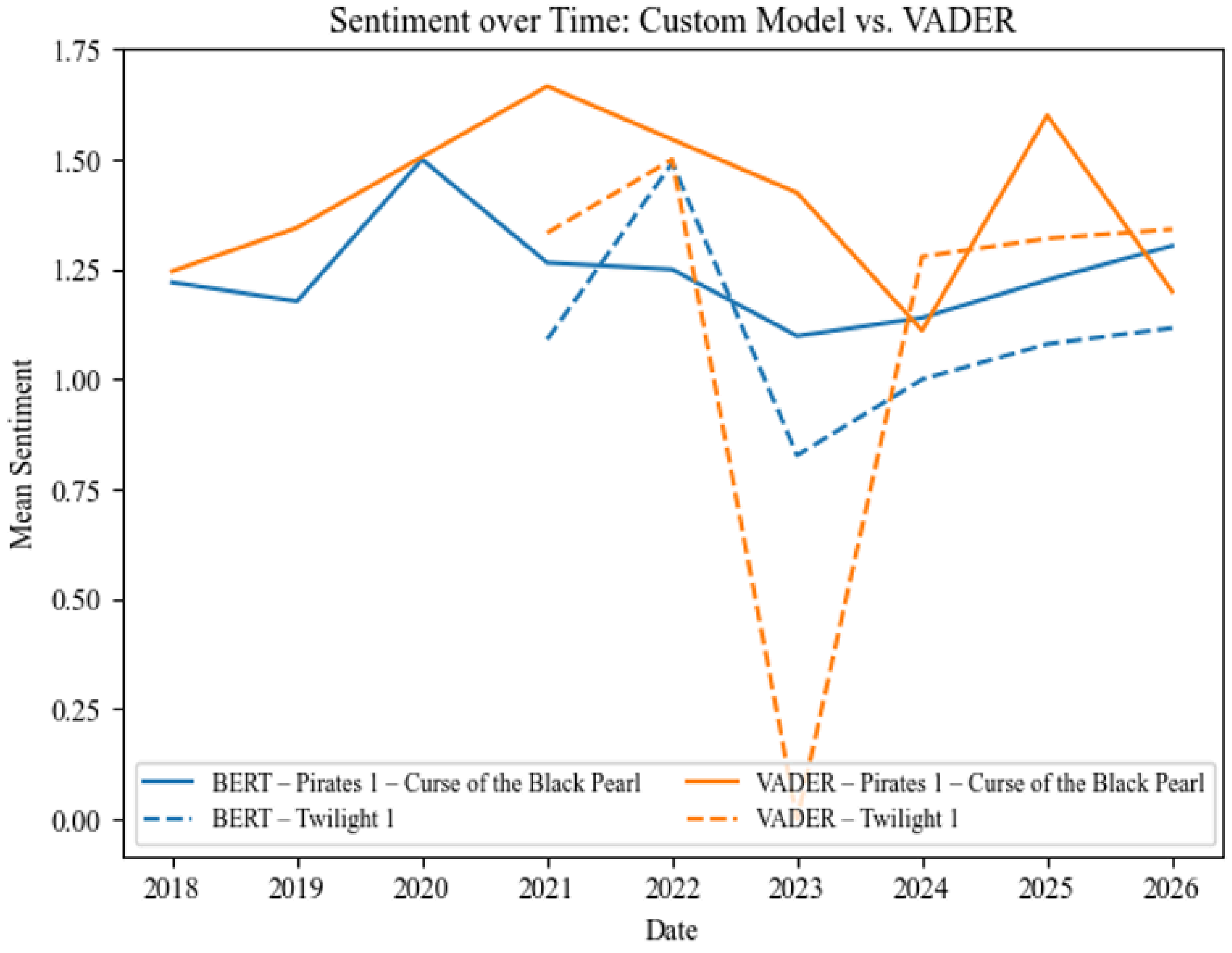
| Class | Precision | Recall | F1-Score | Support |
|---|---|---|---|---|
| Negative | 0.43 | 0.44 | 0.44 | 768 |
| Neutral | 0.68 | 0.33 | 0.44 | 1043 |
| Positive | 0.49 | 0.73 | 0.59 | 1064 |

**Overall Accuracy: 50.7%**

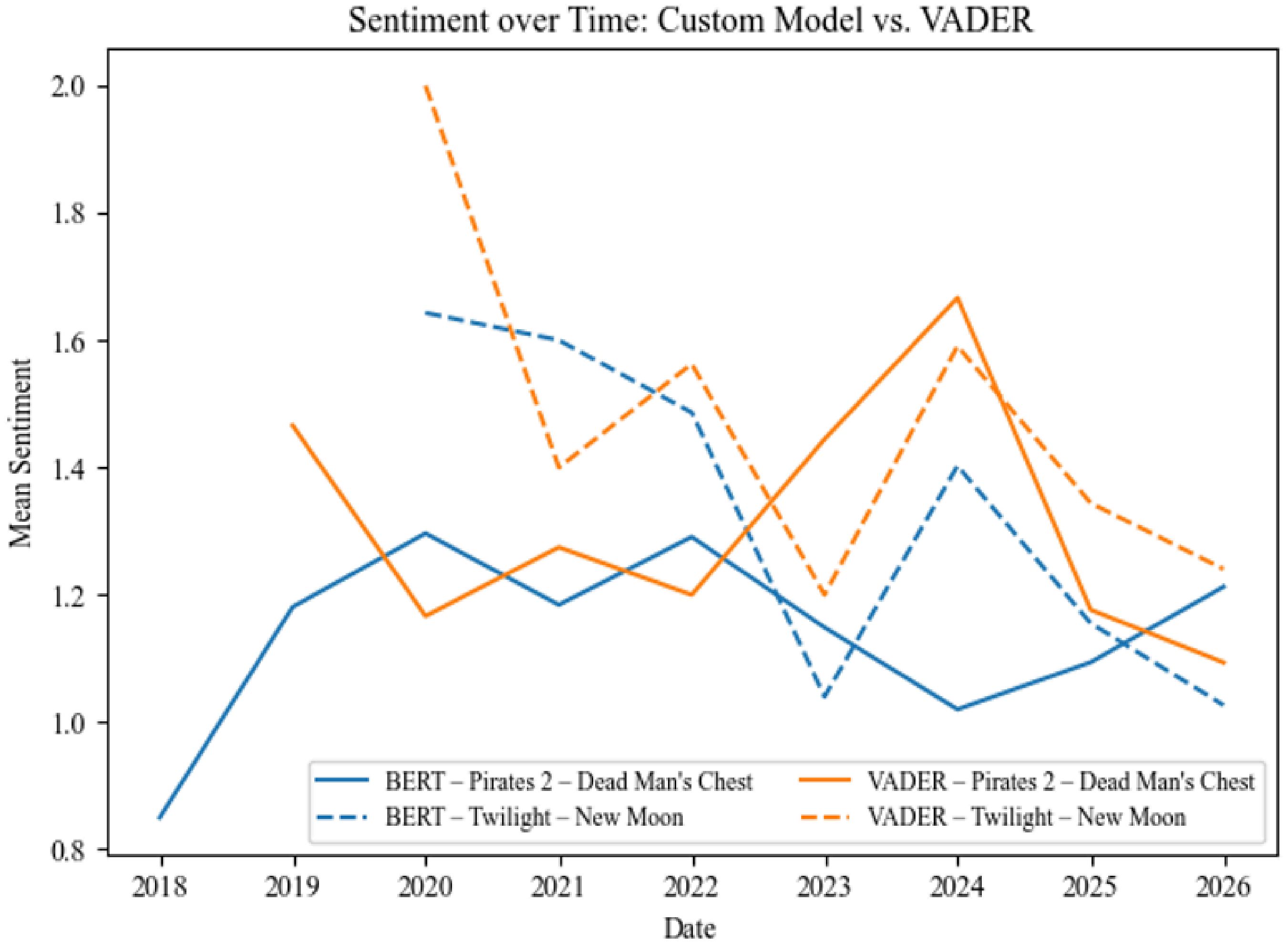
**Precision is highest for predicting neutral**

# Results: Sentiment Over Time (1/3)

**Comparison of Sentiment for each Franchise's Corresponding Movie Releases Over Time**




Movie 1




Movie 2

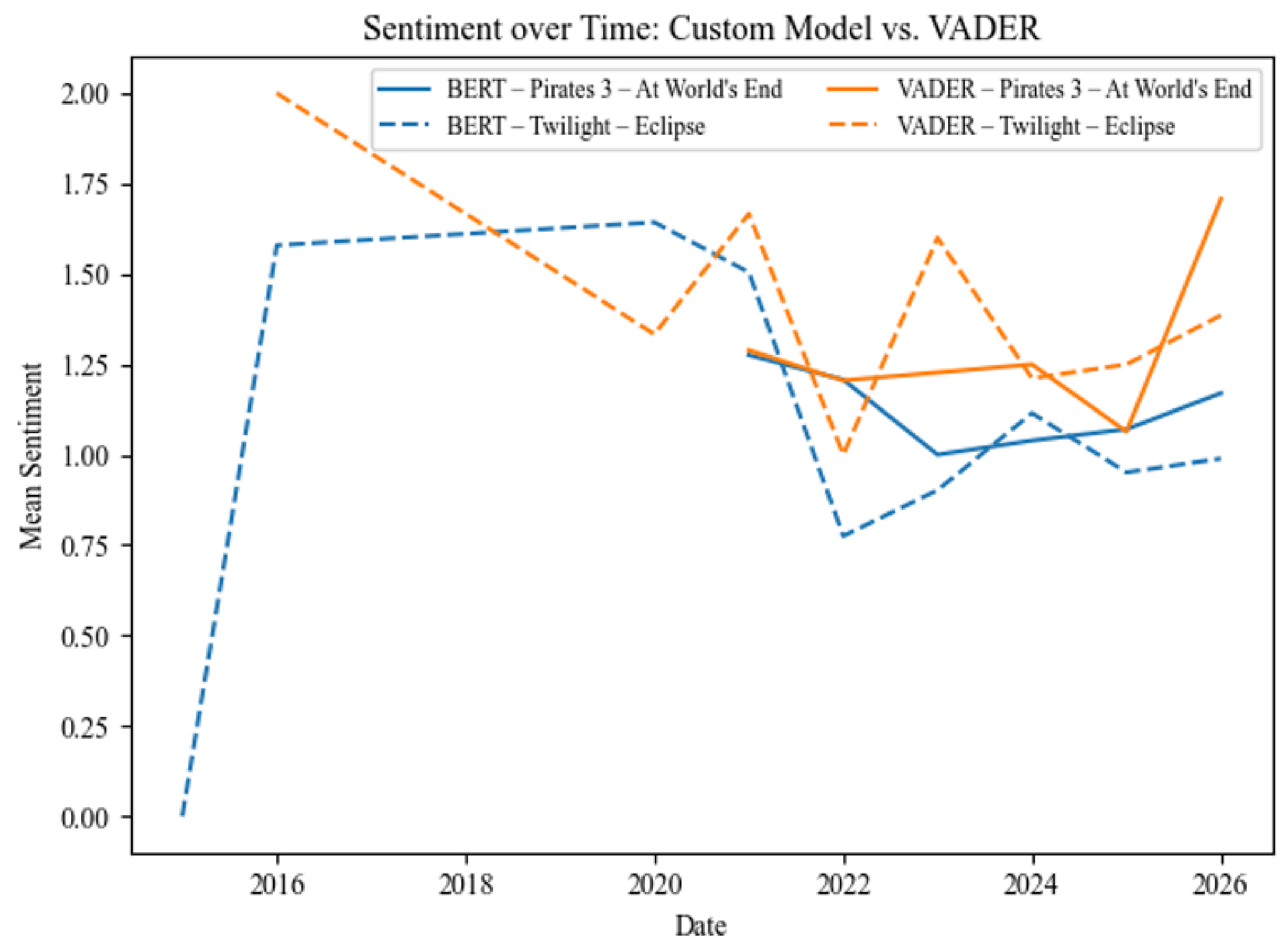
**Legend**

GUIDE

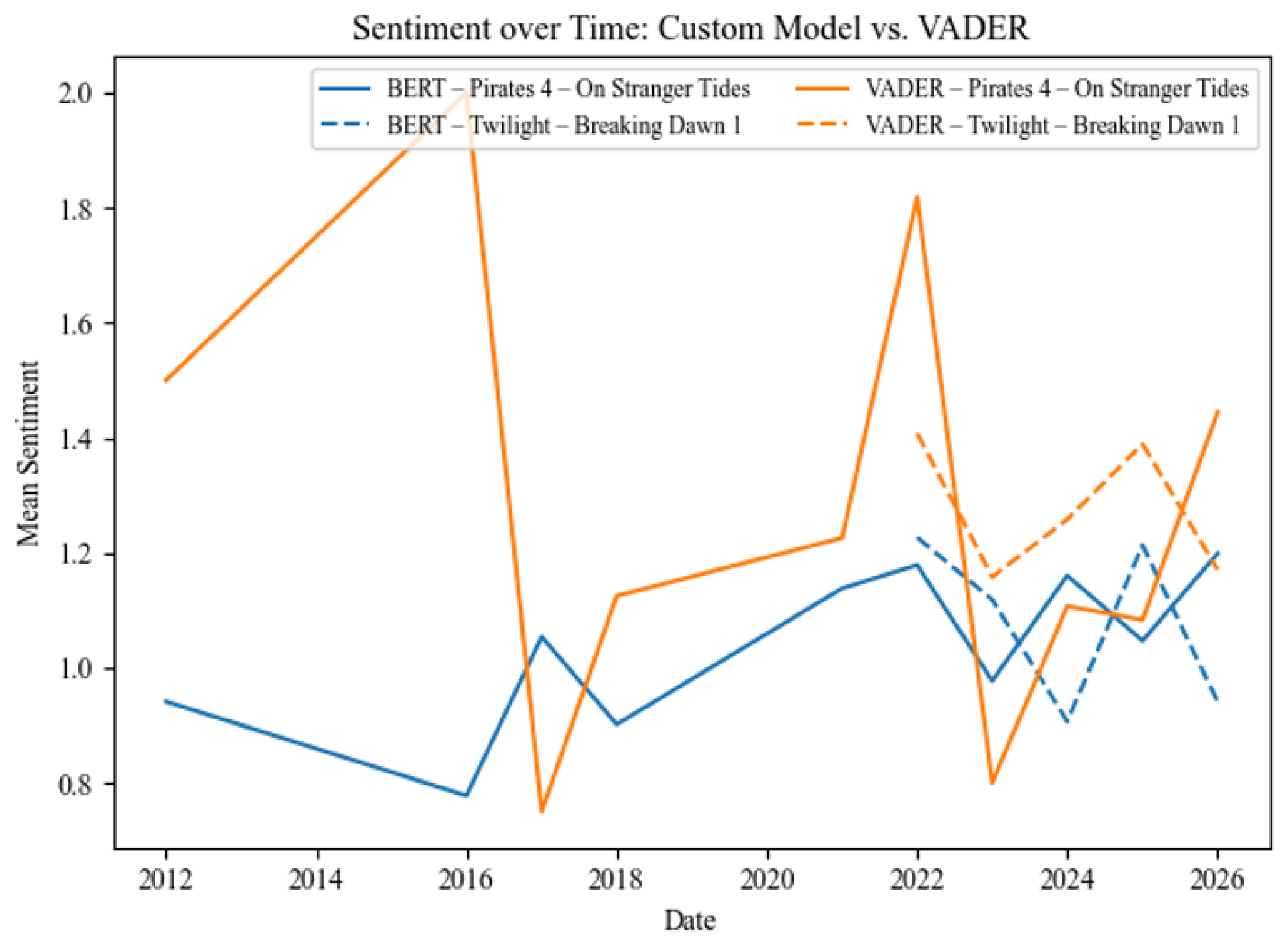
- BERT Pirates
- BERT Twilight
- VADER Pirates
- VADER Twilight

# Results: Sentiment Over Time (2/3)

**Comparison of Sentiment for each Franchise's Corresponding Movie Releases Over Time**




Movie 3




Movie 4

**Legend**

BERT Pirates

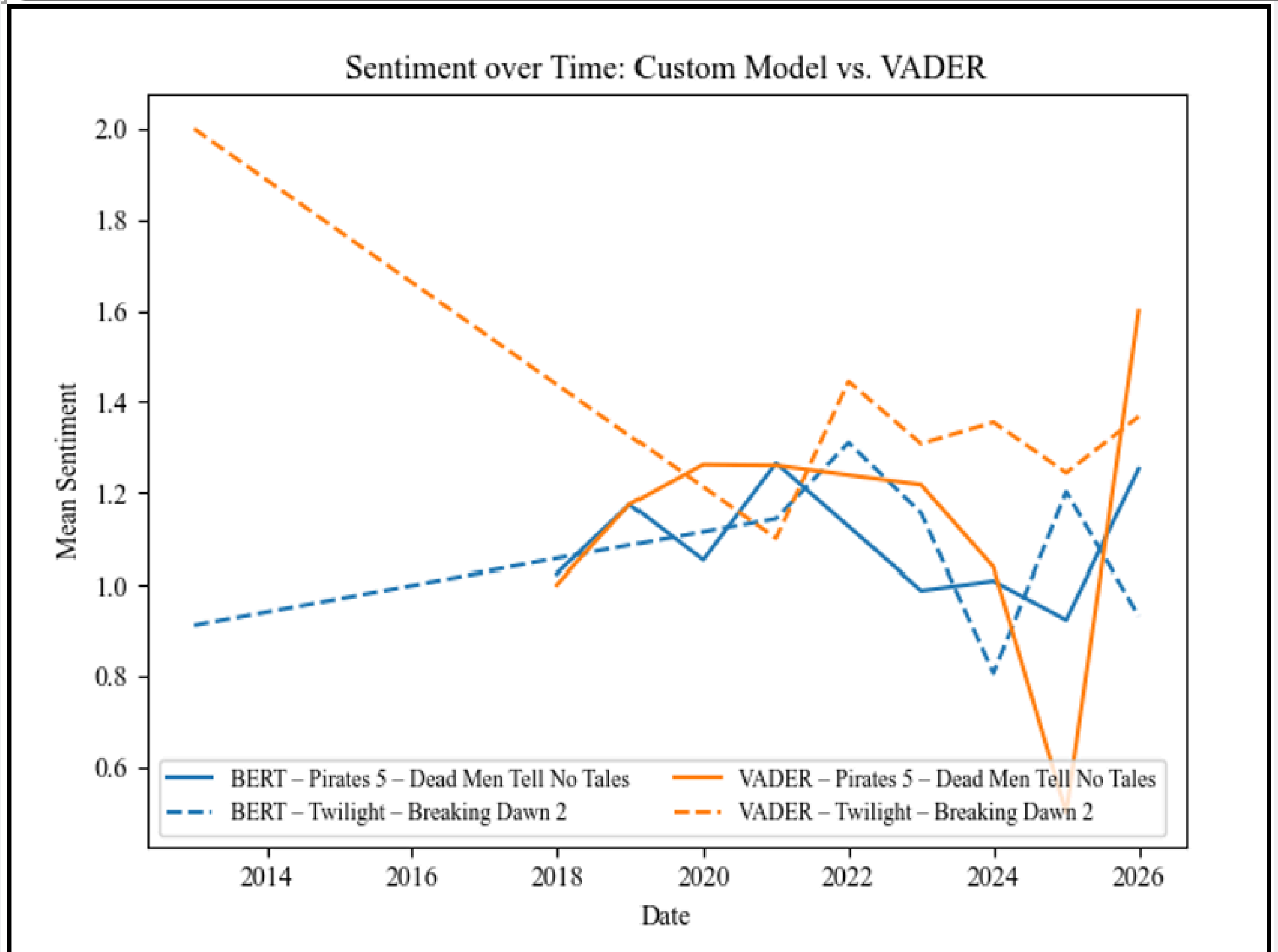
BERT Twilight

VADER Pirates

VADER Twilight

# Results: Sentiment Over Time (3/3)

## Comparison of Sentiment for each Franchise's Corresponding Movie Releases Over Time




Movie 5

## Summary

- Twilight:

BERT and VADER show **similar sentiment trends** across movies

- Pirates:

**Greater difference** between BERT and VADER predictions

➡ Suggests **Pirates comments are more complex** and nuanced, harder for VADER to analyze

➡ Twilight comments are more direct and **emotionally charged**, making VADER and **BERT align more closely**

## Legend

GUIDE

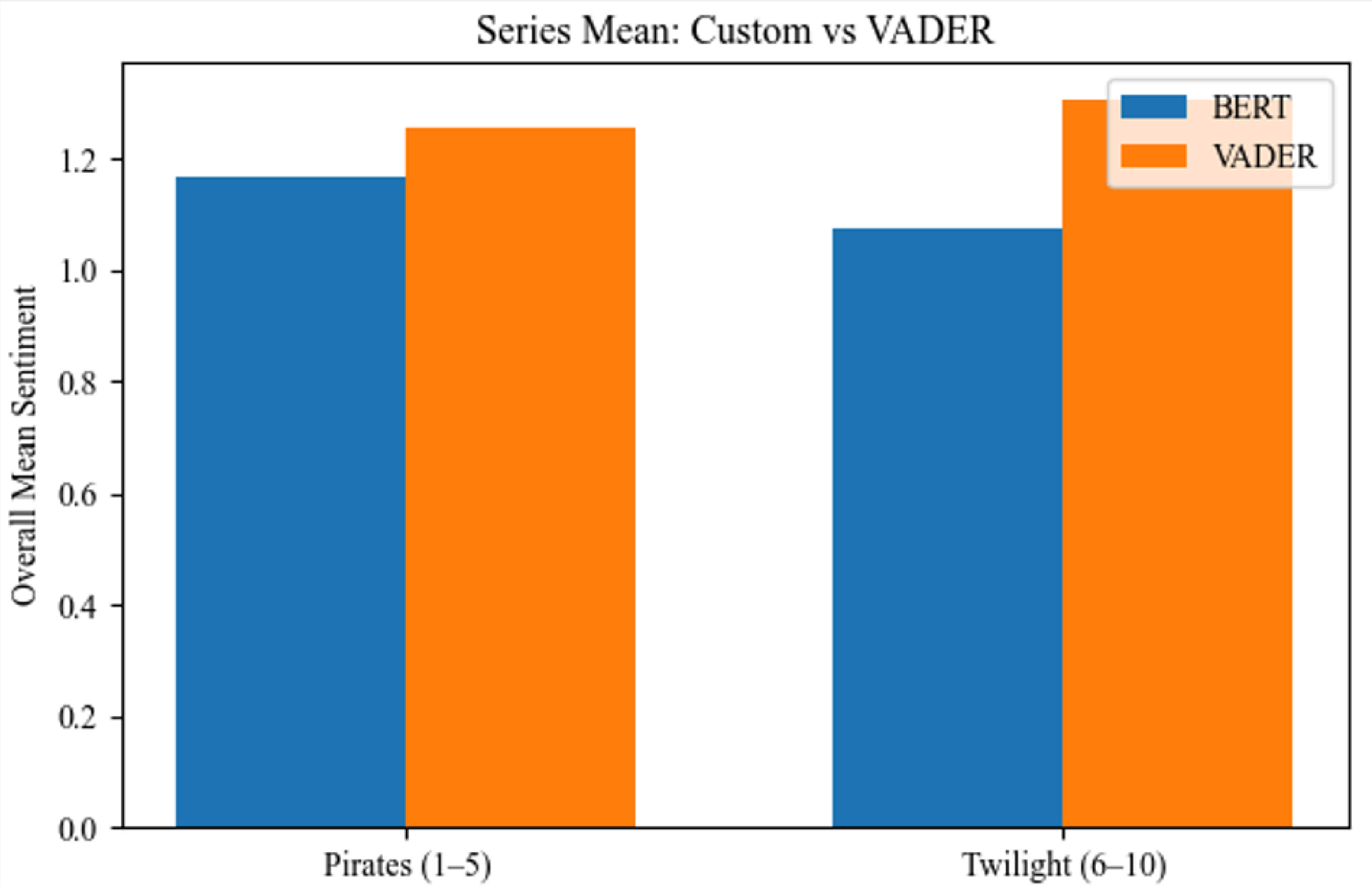
BERT Pirates

BERT Twilight

VADER Pirates

VADER Twilight

# Results: Series Mean Values

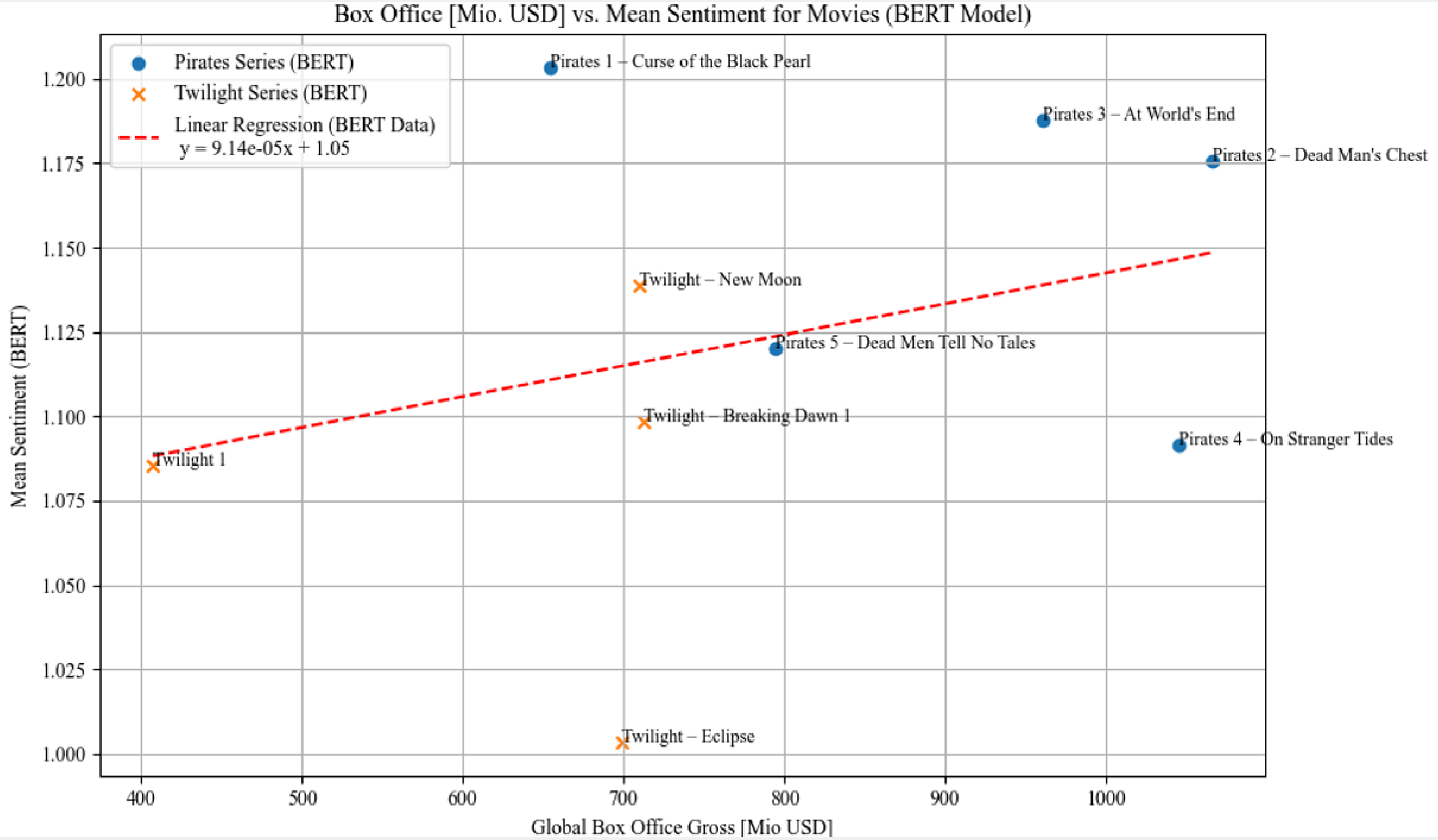



## Key Take-aways

- VADER produces higher overall sentiment scores than BERT for both franchises
- BERT shows Pirates sequels with slightly higher sentiment than Twilight
- VADER suggests the opposite, Twilight shows slightly higher sentiment
- Indicates that VADER may be biased toward overly positive interpretations

# Results: Correlation Analysis

## BoxOffice's Gross and Sentiment Score

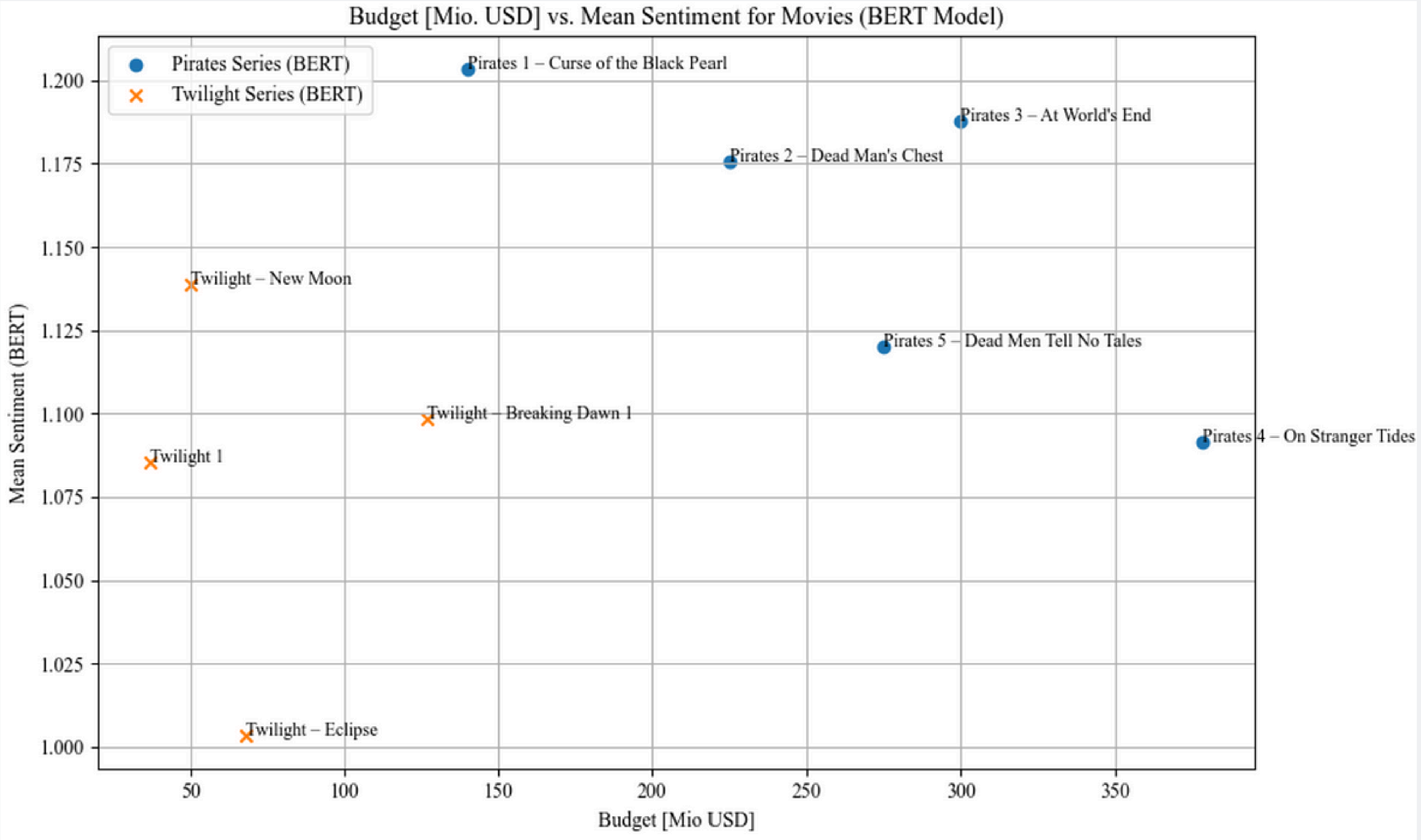



### Key Take-aways

- Scatterplot shows **positive correlation** between sentiment and box office gross
- Movies with **higher sentiment scores tend to earn more globally**
- Suggests **positive audience engagement reflects in commercial success**

# Results: Scatterplot Analysis

## Scatterplot of Budget and Mean Sentiment Score of Both Sequels




### Key Take-aways

- Pirates had much **higher budgets** than Twilight
- Later Pirates movies had **lower sentiment despite higher spending**
- **Budget doesn't guarantee audience approval**

# Results: Scaled Rating Analysis (1/2)

## Comparison between Rating Scores and Predicted Sentiment Score by Two Models

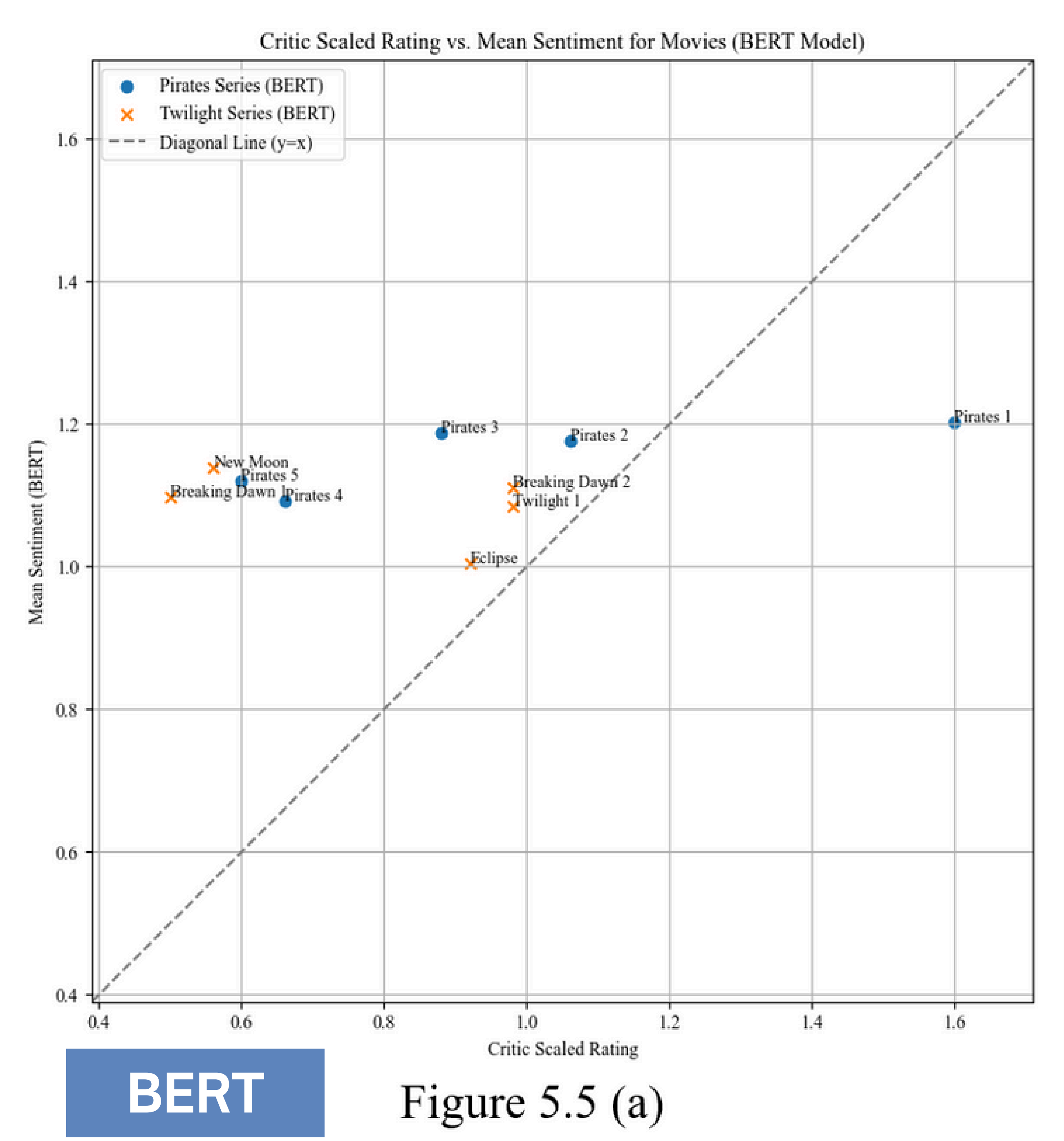



Figure 5.5 (a)

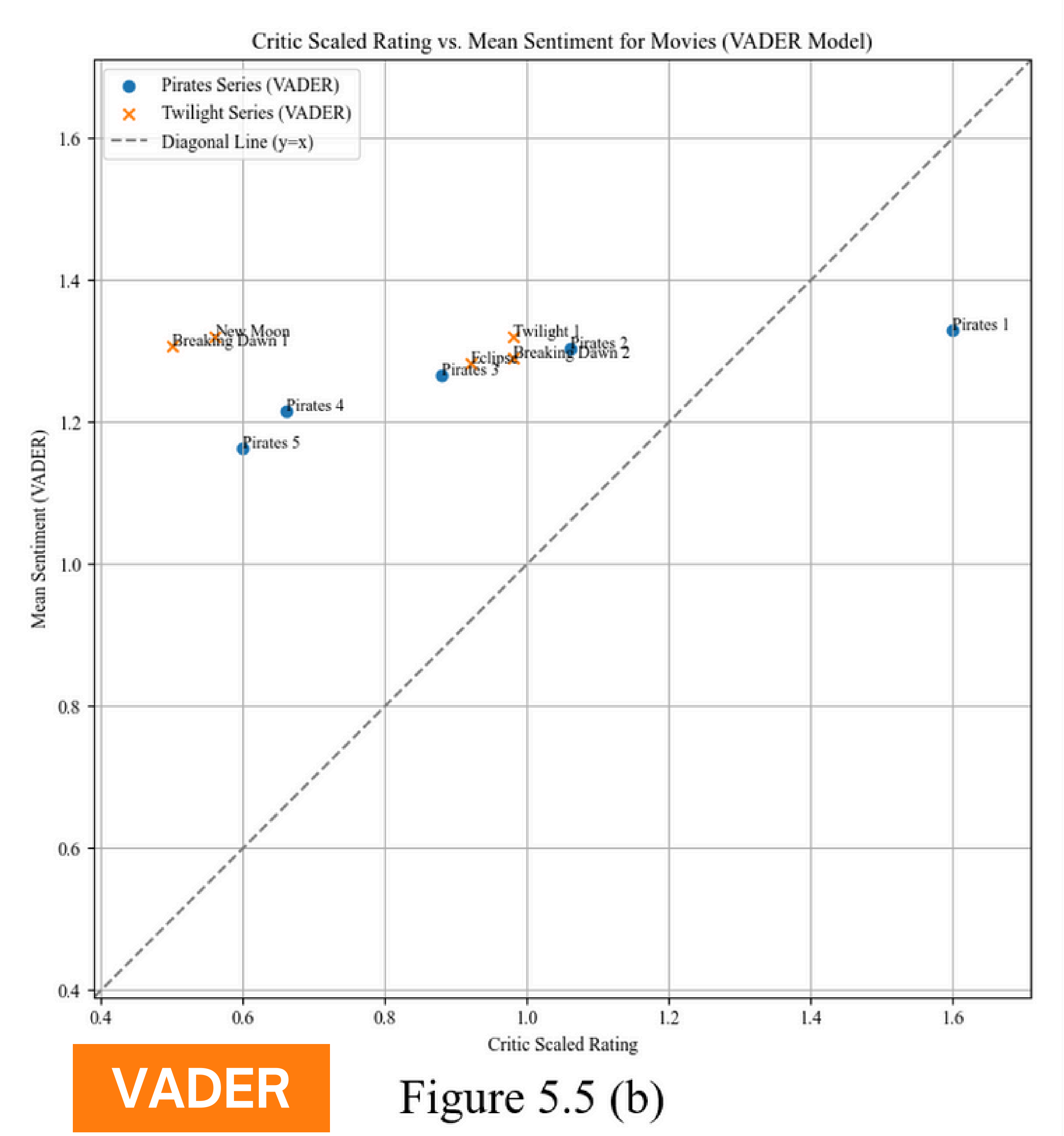


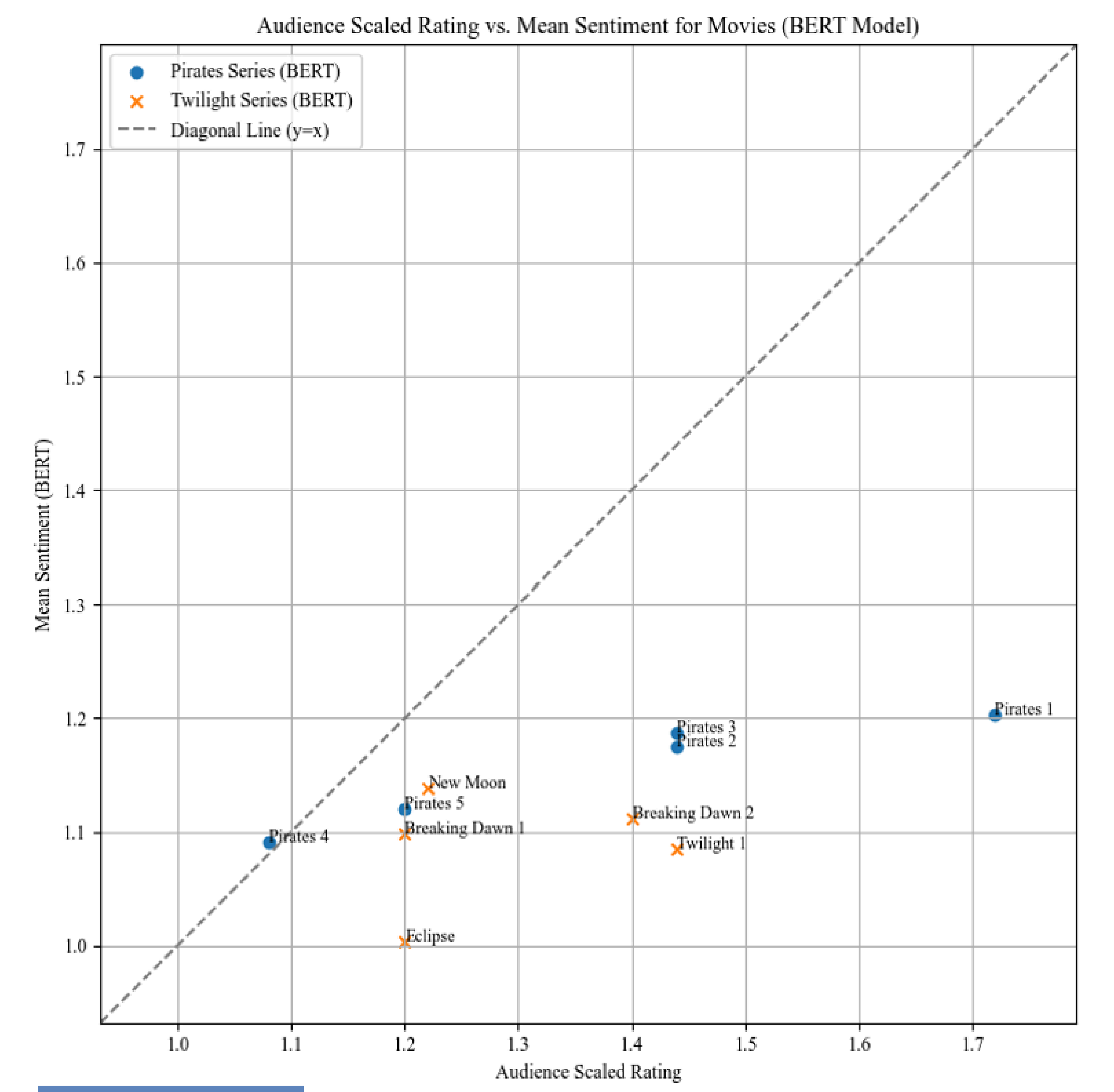

Figure 5.5 (b)

## Key Take-aways

- BERT aligns better with critic ratings than VADER
- **Audience ratings are often higher** than predicted sentiment
- Shows that **BERT is better for modeling nuanced reactions**

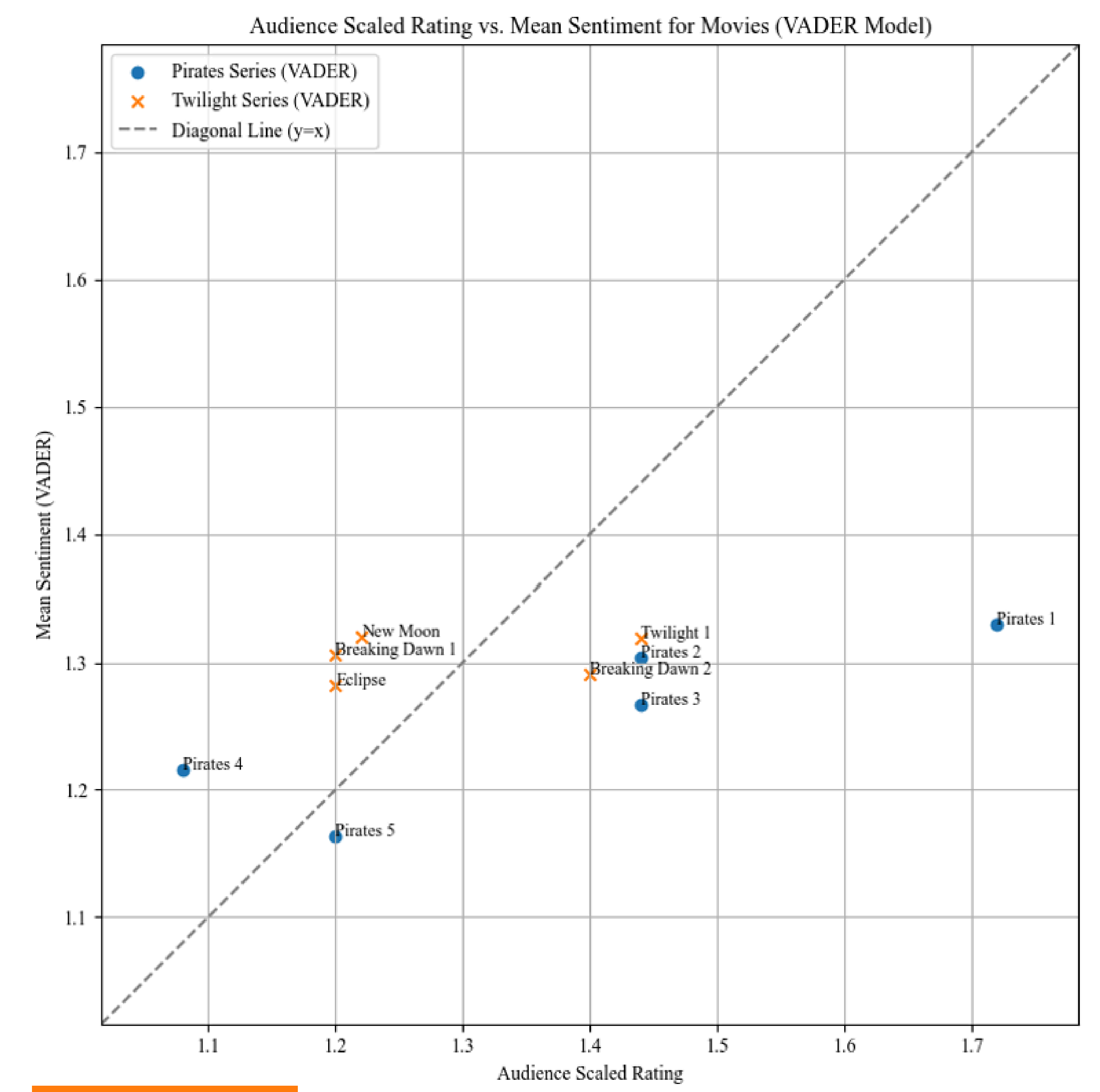
# Results: Scaled Rating Analysis (2/2)

## Comparison between Audience Scaled Ratings and Predicted Sentiment Score by Two Models




BERT

Figure 5.5 (c)




VADER

Figure 5.5 (d)

## Key Take-aways

- In the BERT model, most movies have **lower sentiment scores** than their audience ratings.
- **Pirates 4 is the only case** where sentiment prediction matches audience rating.
- In the VADER model, Twilight movies show **higher sentiment scores than their audience ratings**.
- **Pirates movies generally receive higher audience scores** than what VADER predicts.

# Conclusion and Future Work

1. VADER often overestimates positivity and struggles with context and sarcasm.
2. BERT performs better with nuanced Reddit comments but still limited.
3. Pirates had higher early sentiment, but later films dropped despite big budgets
4. Sentiment ≠ guaranteed box office success --> Factors like timing and fan loyalty matter
5. Labeling was hard: Reddit comments often mixed or off-topic.
6. Accuracy stayed modest, even after using advanced models.
7. Reddit may not reflect the broader public, limiting generalizability.
8. Future models should include more data and context (e.g., release dates, trailers).

Data Analysis and Machine Learning with Python Course

# Thank you

**Group A**

**Lilian, Anh, Maëlys, Ewald, Sora**

Sentiment Analysis

References used for the final group project


- **Batra, H., Punn, N. S., Sonbhadra, S. K., Agarwal, S., Kotsis, G., Strauss, C., Khalil, I., & Tjoa, A. M. (2021).** BERT-based sentiment analysis: A software engineering perspective. In Database and Expert Systems Applications (Vol. 12923, pp. 138–148). Springer International Publishing AG. https://doi.org/10.1007/978-3-030-86472-9_13

- **Chang, M. W., Devin, J., Lee, K., & Toutanova, K. (2018, October 11)**. [1810.04805] BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. arXiv. Retrieved May 27, 2025, from https://arxiv.org/abs/1810.04805

- **Dhaoui, C., Webster, C. M., & Tan, L. P. (2017)**. Social media sentiment analysis: Lexicon versus machine learning. The Journal of Consumer Marketing, 34(6), 480–488. https://doi.org/10.1108/JCM-03-2017-2141

- **Haanaes, K. B. S. (2025, January 7).** Déjà vu: Is the film industry's sequel and remake addiction a sign of the end? IMD Business School for Management and Leadership Courses. https://www.imd.org/research-knowledge/economics/articles/deja-vu-is-the-film-industrys-sequel-and-remake-addiction-a-sign-of-the-end/

- **Hernández-Pérez, R., Lara-Martínez, P., Obregón-Quintana, B., Liebovitch, L. S., & Guzmán-Vargas, L. (2024)**. Correlations and fractality in sentence-level sentiment analysis based on VADER for literary texts. Information (Basel), 15(11), 698. https://doi.org/10.3390/info15110698

- **Hossen, Md. S., & Dev, N. R. (2021).** An improved lexicon-based model for efficient sentiment analysis on movie review data. Wireless Personal Communications, 120(1), 535–544. https://doi.org/10.1007/s11277-021-08474-4

- **Manuel-Ilie, D., Gabriel, P. A., & George, C. R. (2023)**. Sentiment analysis using BERT model. International Journal of Advanced Statistics and IT&C for Economics and Life Sciences, 13(1), 59–66. https://doi.org/10.2478/ijasitels-2023-0007

Sentiment Analysis

References used for the final group project


- **Nkhata, G. (2022).** Movie reviews sentiment analysis using BERT [Master's thesis, University of Arkansas]. ScholarWorks@UARK. https://scholarworks.uark.edu/etd/4768

- **Pano, T., & Kashef, R. (2020).** A complete VADER-based sentiment analysis of Bitcoin (BTC) tweets during the era of COVID-19. Big Data and Cognitive Computing, 4(4), 33. https://doi.org/10.3390/bdcc4040033

- **Samuel, S. T., & Marikannan, B. P. (2020).** Sentiment analysis on movie reviews using Twitter. Journal of Computational and Theoretical Nanoscience, 17(7), 2869–2875. https://doi.org/10.1166/jctn.2020.9326

- **Teja, J. S., Sai, G. K., Kumar, M. D., & Manikandan, R. (2018).** Sentiment analysis of movie reviews using machine learning algorithms: A survey. International Journal of Pure and Applied Mathematics, 118(20), 3277–3284. http://acadpubl.eu/hub/2018-118-21/articles/21d/44.pdf

- **White, A. (2024, September 23).** Hollywood's dramatic decline results in sequels – is this the end? The Torch. https://www.valpotorch.com/arts_and_entertainment/article_b0d68e30-7701-11ef-91fe-e76fb9691cf0.html

- **Youvan, D. (2024, June).** Understanding sentiment analysis with VADER: A comprehensive overview and application. ResearchGate. https://www.researchgate.net/publication/381650914_Understanding_Sentiment_Analysis_with_VADER_A_Comprehensive_Overview_and_Application