Enhancing Review Processing in the Video Game Adaptation Domain through VADER and Rating-Based Labeling using SVM

Danita Divka Sajmira[1], Khothibul Umam[2]*, Maya Rini Handayani[3]
Department of Information Technology, Science and Technology Faculty [1], [2]
Department of Da'wah and Communication, Da'wah Faculty [3]
Universitas Islam Negeri Walisongo Semarang,
Semarang, Indonesia
2208096038@student.walisongo.ac.id[1], khothibul_umam@walisongo.ac.id[2], maya@walisongo.ac.id[3]

***Abstract*— The adaptation of video games into films or television series has increasingly become a prominent trend in the entertainment sector, often eliciting diverse reactions from audiences. A prime example is *The Last of Us*, a video game adaptation series that generated substantial online discussions and sentiment, and serves as the specific case study in this research. Sentiment patterns found in audience reviews of *The Last of Us* on IMDb are analyzed using a domain-specific classification framework tailored to the language characteristics of entertainment media. A key issue addressed is the discrepancy between numerical ratings and the sentiment conveyed in review texts, which may lead to inconsistent labeling. The study employs a machine learning technique, Support Vector Machine (SVM), coupled with two distinct labeling methods: manual labeling based on IMDb ratings, and automatic labeling using the lexicon-driven VADER tool. A total of 2,017 English reviews of *The Last of Us* were gathered via web scraping from IMDb, followed by preprocessing, TF-IDF feature extraction, and hyperparameter optimization using RandomizedSearchCV. These results show that the SVM model trained on VADER-labeled data achieved an accuracy of 0.97, outperforming the model trained on manually labeled data at 0.79. Lexicon-based automatic labeling provides more consistent and reliable sentiment classification, particularly in specialized domains like video game adaptation reviews. Integrating VADER labeling with SVM enhances sentiment analysis effectiveness and offers practical value for media analytics, content creation, and audience insight research.**

Keywords*— SVM, VADER, Sentiment Analysis, Video Game Adaptation, IMDb Reviews

I. INTRODUCTION

In the digital era, user-generated content has grown exponentially and significantly impacts various industries, including the entertainment industry. The development of films and television series has accelerated, marked by the production of diverse content aimed at entering international markets, entertaining audiences, achieving high ratings, and maximizing profit [1]. In the past, one of the rising trends in this industry has been adaptations from various sources, particularly literature and video games [2]. Video game adaptations into television series or films have become a rapidly growing phenomenon in the global entertainment industry, often receiving varied responses from audiences, critics, and gaming communities. One notable adaptation that gained widespread attention is *The Last of Us*, a post-apocalyptic drama series adapted from the popular video game of the same name, developed by Naughty Dog.

To systematically understand audience perceptions of this series, a method capable of automatically analyzing sentiment in textual reviews is needed. In this context, sentiment analysis based on Natural Language Processing (NLP) has become a widely used solution, especially for extracting positive and negative opinions from review texts. Sentiment analysis is an automated process used to interpret, extract, and process text-based data to identify the underlying opinions or emotions expressed in a statement [3].

The dataset used in this research was obtained from user-generated reviews of *The Last of Us* series available on the IMDb platform. IMDb (Internet Movie Database), a comprehensive film and television database established in 1990 and currently operated under Amazon, serves as a widely recognized source for audience feedback, is a digital platform that provides comprehensive information about films, TV series, and the people involved in their production [4]. In addition to film and series information, IMDb features user-generated reviews and ratings. Although it offers a rating system, the content of the reviews often does not align with the given scores. For example, viewers may give high ratings while still criticizing the film. Therefore, textual sentiment analysis is necessary to better understand audience opinions and perceptions from various backgrounds, going beyond the surface-level numerical ratings.

Sentiment analysis typically utilizes two main approaches: the lexicon-based technique and the Support Vector Machine (SVM) algorithm. SVM is well-known for its strong performance in classification and regression tasks, handling both linear and non-linear data effectively [5]. Its ability to identify patterns in textual data makes SVM one of the most

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i3.2409, Copyright ©2025
Submitted : June 2, 2025, Revised : June 25, 2025, Accepted : July 1, 2025, Published : July 28, 2025

effective algorithms for text-based classification [6]. Moreover, SVM can handle high-dimensional data and offers flexibility through kernel options that enable the modeling of both linear and non-linear patterns.

Meanwhile, one of the most effective lexicon-based methods is VADER (Valence Aware Dictionary and Sentiment Reasoner), which is specifically designed for sentiment analysis by leveraging an English-language sentiment lexicon. It is used to detect sentiment polarity efficiently by utilizing a predefined dictionary of sentiment-laden words [7].

This study integrates both approaches by using SVM algorithm as the classification model, trained on two different types of sentiment labels. The first set of labels was obtained manually by converting IMDb ratings into sentiment categories (positive and negative), while the second set was generated automatically using the VADER method.

A previous study conducted by Siti and Risa, titled "Evaluasi Performa Kernel SVM dalam Analisis Sentimen Review Aplikasi ChatGPT Menggunakan Hyperparameter dan VADER Lexicon", applied a 90:10 data split scheme and found that kernel selection significantly affected classification accuracy. The linear and sigmoid kernels achieved an accuracy of 92.23%, followed by the polynomial kernel with 90.78%, while the RBF kernel delivered the best performance at 92.72%. Additionally, sentiment labeling using the VADER lexicon indicated that positive reviews were more dominant than negative ones [8]. Another relevant research conducted by Viviane and Zuriani, titled "Vader Lexicon and Support Vector Machine Algorithm to Detect Customer Sentiment Orientation" showed that utilizing VADER in conjunction with the SVM algorithm proved to be an effective method for sentiment analysis, reaching an accuracy rate of 88.57% [9].

Meanwhile, Anwar and Permana, through their research entitled *"*Analisis Sentimen Masyarakat Indonesia Terhadap Produk Kendaraan Listrik Menggunakan VADER*"*, a key limitation of VADER was identified in its handling of informal language. Manual inspection of several tweets labeled as negative revealed misclassifications, such as "gila cute bgt gk da obat" and "emang keren parah si berkualitas bgt." These tweets were classified as negative due to the presence of words like *"gila"* and *"parah"*, which VADER interpreted literally as negative, although they were used to emphasize positive sentiments. This finding highlights VADER's limited ability to understand contextual nuance and sarcasm [10].

In a separate investigation, Zhafira et al. examined sentiment classification on IMDb film reviews using the Naive Bayes algorithm. Their study demonstrated that the model achieved a classification accuracy of 86.25%. However, the researchers also highlighted key limitations of Naïve Bayes, particularly its strong assumption of feature independence, which restricts its ability to interpret contextual nuances such as sarcasm or irony. To address these limitations, they proposed the use of more sophisticated models, including SVM, Random Forest, and deep learning techniques like LSTM [11]. A related study conducted by Ramadhan et al. explored sentiment classification of IMDb movie reviews by evaluating the performance of different algorithms. The results indicated that the Decision Tree algorithm yielded better accuracy. While their primary focus was on comparing classifiers, the study underscores the importance of sentiment analysis on IMDb reviews for capturing audience opinion [12].

While VADER and machine learning models such as SVM have been widely adopted in sentiment analysis tasks, comparative approaches that examine lexicon-based and rating-based labeling remain underexplored particularly in high-context domains, which are characterized by emotionally charged language and informal expressions, such as reviews of video game adaptation series. Previous studies have primarily focused on algorithmic performance or general sentiment classification, often overlooking the potential impact of different labeling strategies on classification accuracy. This study, therefore, seeks to evaluate and contrast the effectiveness of the SVM algorithm in analyzing sentiment from IMDb reviews of *The Last of Us* by employing two distinct labeling techniques: automatic labeling using the VADER lexicon and manual labeling derived from rating-to-sentiment conversion.

II. Research Metodology

The research methodology follows a structured sequence of stages to ensure that the expected outcomes align with the initial objectives [13]. In this study, the methodology is systematically designed to facilitate the analysis and evaluation of model performance. In support of an effective sentiment classification approach, each step has its specific function. The entire process is visually summarized through Figure 1.

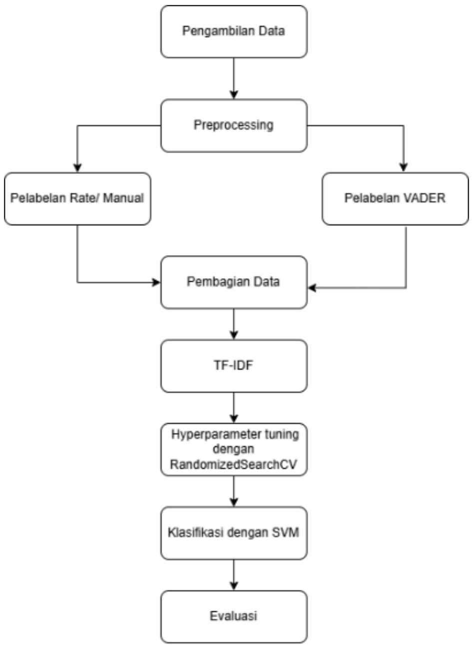

Fig 1. Research Workflow

A. *Data Collection*

Review data from *The Last of Us* series on IMDb was collected using web scraping techniques with the help of the Easy Scraper browser extension. Web scraping is an automated method for extracting data from websites by

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i3.2409, Copyright ©2025
Submitted : June 2, 2025, Revised : June 25, 2025, Accepted : July 1, 2025, Published : July 28, 2025

systematically accessing specific pages and retrieving the required information. Through this process, several data points were obtained, including rating, review link, review text, number of likes, username link, username, and review timestamp. All reviews collected were written in English. The scraped data was saved in CSV format; however, only two main attributes were used in this study: rating and review text.

B. Preprocessing

The preprocessing stage aims to structure and clean raw data to ensure it is optimally prepared for subsequent analysis. This process removes issues such as duplicates, irrelevant characters, and other non-contributing components that may interfere with sentiment analysis [14]. Furthermore, preprocessing enhances the quality of text representation, making the extracted features more relevant and effective when processed by machine learning algorithms such as SVM. The following steps outline the preprocessing procedures applied in this study.

- Case Folding and Cleaning Text

  This stage involves two essential processes in text preprocessing. First, case folding aims to convert all text into lowercase, ensuring that each word has a uniform and consistent form during data processing [15]. Second, cleaning text involves eliminating unnecessary components like hyperlinks, punctuation marks, numerical characters, and symbols that hold no value for sentiment analysis.
- Tokenization

  The tokenization stage involves breaking the text into smaller units known as tokens [16]. This segmentation is typically performed by splitting the text based on non-alphanumeric characters to produce a more structured representation of each review.
- Stopword Removal

  This step is used to eliminate common words that frequently appear in text such as “the,” “and,” or “is” as they do not contribute significantly to the sentiment classification process. Stopword removal helps reduce noise and focus on more relevant information [17].
- Lemmatization

  Lemmatization involves transforming words into their root or dictionary forms. In contrast to stemming, it delivers more precise outcomes by taking the grammatical context into account, thereby maintaining the word’s original meaning [18].

C. Labelling

In this study, sentiment labeling was conducted using two approaches: rating-based (manual) labeling and automatic labeling using the VADER method. The results from both labeling approaches were later compared based on the classification accuracy achieved by the model.

- VADER Labelling

  Sentiment labeling using VADER was performed automatically based on the compound score generated from text analysis. A review is categorized as positive when its compound score is equal to or above 0, and classified as negative if the score falls below 0 [19]. Common words associated with positive sentiment include *amazing*, *happy*, *love*, and *great*, while negative sentiment is typically indicated by words such as *bad*, *terrible*, *hate*, and *sad*.
- Rating-Based (Manual) Labeling

  Manual labeling was conducted by categorizing reviews into two sentiment classes. Reviews with ratings from 6 to 10 were classified as positive, while those with ratings from 1 to 5 were classified as negative. This approach is based on the assumption that the rating score reflects the user's satisfaction or dissatisfaction with the series being reviewed.

D. Data Splitting

The dataset was divided into two separate segments: a training set and a testing set. The training data was utilized to construct and optimize the model, whereas the testing data was allocated to evaluate its ability to perform on previously unseen instances. This division plays a vital role in ensuring unbiased performance assessment and in reducing the likelihood of overfitting due to the model’s exposure to known data.

E. TF-IDF

Term Frequency–Inverse Document Frequency (TF-IDF) is a commonly used technique in text mining that quantifies the significance of a term within a particular document in relation to its frequency across an entire corpus [20]. This approach highlights unique terms within each document while downplaying the influence of commonly occurring words shared across multiple documents.

F. Hyperparameter Tuning using RandomizedSearchCV

This study utilized RandomizedSearchCV to optimize model performance by randomly selecting hyperparameter values. This approach enables broader and more efficient exploration of the hyperparameter space compared to manual search methods, as combinations are sampled randomly from predefined distributions [21].

G. Classification

The classification was performed using the SVM algorithm, where each review was evaluated to identify whether it expressed a positive or negative sentiment. The trained SVM model classified the reviews based on patterns it learned during the training phase.

H. Evaluation

In this stage, the effectiveness of the previously developed classification model was evaluated. The primary instrument used was the confusion matrix, which formed the foundation for calculating essential performance indicators such as accuracy, precision, recall, and F1-score. These metrics offer a comprehensive understanding of the

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i3.2409, Copyright ©2025
Submitted : June 2, 2025, Revised : June 25, 2025, Accepted : July 1, 2025, Published : July 28, 2025

model's ability to accurately classify data and generalize well to new, unseen samples. The following are equations (1), (2), (3), and (4) used to calculate each performance metric [22].

- Accuracy

  Accuracy indicates the ratio of correctly predicted instances to the total number of samples evaluated by the model.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

- Precision

  Precision measures the extent to which the model's positive predictions are relevant to the actual positive instances in the data.

$$\text{Precision} = \frac{TP}{TP + FP} \quad (2)$$

- Recall

  Recall indicates the extent to which the model is able to find and correctly identify all relevant data instances.

$$\text{Recall} = \frac{TP}{TP + FN} \quad (3)$$

- f1-score

  The F1-score merges precision and recall into one metric, providing a balanced measure of the model's overall effectiveness.

$$\text{f-1} = \frac{2 \times \text{Recall} \times \text{Precision}}{\text{Recall} + \text{Precision}} \quad (4)$$

III. RESULTS

At this stage, the processes and results of the procedures carried out during the study are explained, as previously described in the methodology chapter.

A. *Data Collection*

Review data for *The Last of Us* series was obtained from the IMDb website using the Easy Scraper extension. The data collection was conducted on May 21, 2025, resulting in a total of 2.017 review entries. Figure 2 displays the interface during the scraping process.

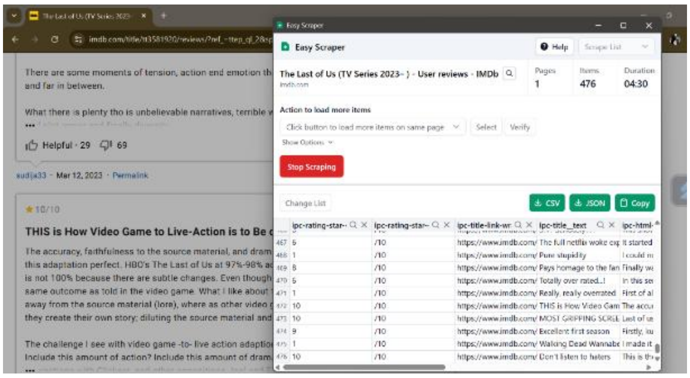

Fig 2. Data Collection Process Using Extension

B. *Preprocessing*

To clarify the preprocessing steps conducted in this study, a sample of the data at each transformation stage from raw input to the final processed form is presented. Table I shows an example of raw review data scraped from IMDb.

TABLE I. SAMPLE DATA

Rate	Review
2	Bad Writing Worse Casting I really tried to gi...
6	Here's a warning
10	Amazing, absolutely amazing Bella Ramsey is pe...

- Case Folding and Cleaning Text

  To achieve consistency in textual data, all letters were standardized to lowercase using the case folding technique. Subsequently, unnecessary elements including punctuation, digits, and special characters were eliminated as part of the cleaning process. The output from these steps is presented in Table II.

TABLE II.
RESULT OF CASE FOLDING AND CLEANING TEXT

Rate	Review	Text
2	Bad Writing Worse Casting I really tried to gi...	bad writing worse casting i really tried to gi...
6	Here's a warning	here is a warning
10	Amazing, absolutely amazing Bella Ramsey is pe...	amazing absolutely amazing bella ramsey is per...

- Tokenization

  After text cleaning through case folding and character removal, the tokenization stage breaks each sentence into individual words. The output of this step is presented in Table III.

TABLE III. RESULT OF TOKENIZED

Text	Tokenized
bad writing worse casting i really tried to gi...	[bad, writing, worse, casting, i, really, trie...
here is a warning	[here, is, a, warning]
amazing absolutely amazing bella ramsey is per...	[amazing, absolutely, amazing, bella, ramsey, ...

- Stopword Removal

  Next, commonly used words categorized as stopwords were removed, as they do not contribute significantly to the sentiment analysis. Table IV presents the outcome of the stopword removal process.

TABLE IV. RESULT OF STOPWORD

Tokenized	Stopword
[bad, writing, worse, casting, i, really, trie...	[bad, writing, worse, casting, really, tried, ...
[here, is, a, warning]	[warning]
[amazing, absolutely, amazing, bella, ramsey, ...	[amazing, absolutely, amazing, bella, ramsey, ...

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i3.2409, Copyright ©2025
Submitted : June 2, 2025, Revised : June 25, 2025, Accepted : July 1, 2025, Published : July 28, 2025

- Lemmatization

  The final step, lemmatization, transforms words into their root forms to improve consistency and contextual relevance in sentiment analysis. Table V presents the outcome of the lemmatization process.

TABLE V. RESULT OF LEMMATIZATION

Stopword	Lemmatized
[bad, writing, worse, casting, really, tried, ...	[bad, writing, worse, casting, really, tried, ...
[warning]	[warning]
[amazing, absolutely, amazing, bella, ramsey, ...	[amazing, absolutely, amazing, bella, ramsey, …

C. *Labelling*

Labeling is the process of classifying sentiment into predefined categories. In this study, two labeling approaches were applied: manual labeling based on ratings and automatic labeling using the VADER lexicon. All data were classified into two sentiment categories: positive and negative.

- VADER Lexicon

  This approach categorized reviews as either positive or negative based on the compound score produced by VADER. The compound score represents an aggregated measure that indicates the overall sentiment polarity of a review. A review is classified as positive if the compound score is ≥ 0, and negative if the score is < 0. Figure 4 demonstrates the results obtained from the labeling process.

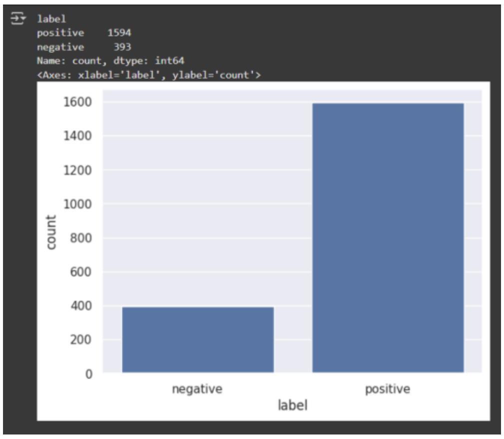

Fig 3. VADER labeling result

  As shown in Figure 3, the sentiment labeling results from VADER indicate that the majority of reviews 1.594 fall under the positive category, whereas 393 reviews are categorized as negative.

- Manual (Rating-Based Labeling)

  Manual labeling was conducted based on the predefined rating range. Figure 4 demonstrates the results obtained from the labeling process.

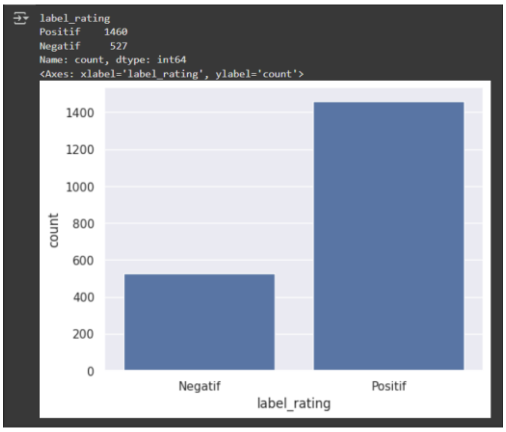

Fig 4. Manual labeling result

  From these results, it is also evident that the number of reviews labeled as positive sentiment (1,460 entries) exceeds those labeled as negative sentiment (527 entries).

D. *Data Splitting*

After the labeling phase, the dataset was partitioned into 80% for training and 20% for testing. This partitioning was intended to facilitate effective model training while ensuring an unbiased evaluation of its performance on data that had not been encountered during training, especially in the context of sentiment prediction for video game adaptation reviews.

E. *TF-IDF*

After the data splitting process, feature weighting was conducted using the TF-IDF approach. This step aimed to transform the text into a numerical format, allowing the model to recognize the most significant terms contributing to sentiment classification. As a result, the classification process becomes more effective, as the model can capture important patterns in reviews related to video game adaptations.

F. *Hyperparameter Tuning using RandomizedSearchCV*

Once feature weighting was completed, the next step involved tuning the hyperparameters of the SVM model. This study utilized RandomizedSearchCV to efficiently search for the best combination of parameters. The parameters adjusted included the kernel type (linear and RBF), the regularization parameter C, and gamma, each within a logarithmic range. The tuning process was carried out using 4-fold cross-validation, evaluating 50 random combinations. The result of this process yielded the optimal parameter configuration, which was then used to train the final model. The following Table VI presents the best hyperparameter combinations based on the labeling approach used.

TABLE VI. HYPERPARAMETER TUNING RESULT

Pelabelan	Algo Kernel	Algo Gamma	Algo C
VADER	RBF	1.0	1.0
Manual	RBF	0.01	100.0

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i3.2409, Copyright ©2025
Submitted : June 2, 2025, Revised : June 25, 2025, Accepted : July 1, 2025, Published : July 28, 2025

G. Classification

The next stage involved performing classification using the SVM algorithm with the optimal parameters obtained from the Randomized Search CV process. The model was trained using the training dataset and tested on the testing dataset to assess its ability to accurately classify sentiment in reviews.

H. Evaluation

The performance of the model was evaluated using a confusion matrix, which illustrates the distribution of the model's predictions across the two sentiment classes: positive and negative, as previously defined by both VADER-based and manual labeling approaches. Figure 5 presents the confusion matrix of the SVM model evaluated using VADER-based sentiment labeling.

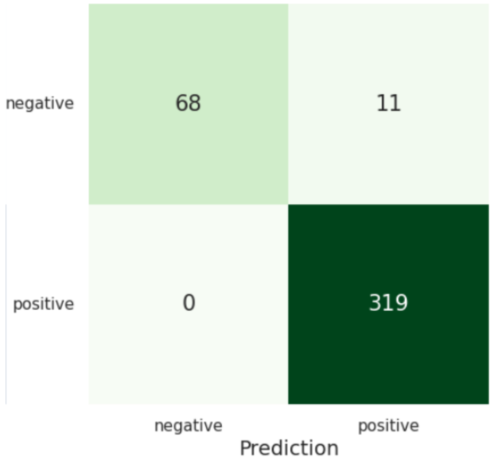

Fig 5. Confusion Matrix VADER labeling

This information is summarized in Table VII

TABLE VII. CONFUSION MATRIX

True Positive (TP)	True Negative (TN)	False Positive (FP)	True Negative (FN)
319	68	11	0

Referring to the confusion matrix shown in Table VII, performance indicators such as accuracy, precision, recall, and F1-score were calculated using the formulas outlined earlier in the methodology section. These metrics provide a measure of how well the model can differentiate between positive and negative reviews based on lexicon-based labeling. A summary of these evaluation results is presented in Table VIII.

TABLE VIII. EVALUATION RESULTS OF SVM WITH VADER LABELING

	precision	recall	f1-score	support
Negatif	1.00	0.86	0.93	79
Positif	0.97	1.00	0.98	319
Accuracy			0.97	398
Macro avg	0.98	0.93	0.95	398
Weighted avg	0.97	0.97	0.97	398

In addition to VADER-based lexicon labeling, the SVM model was also evaluated using data labeled manually. The evaluation followed the same methodology, using a confusion matrix to display how the model's predictions were distributed between the two sentiment classes. Figure 6 shows the confusion matrix for the SVM model based on manual labeling.

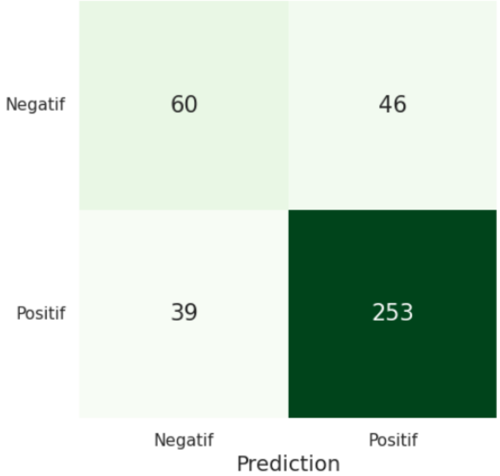

Fig 6. Confusion Matrix manual labeling

This information is summarized in Table IX.

TABLE IX. CONFUSION MATRIX

True Positive (TP)	True Negative (TN)	False Positive (FP)	True Negative (FN)
253	60	46	39

Based on the confusion matrix generated from manual labeling, accuracy, precision, recall, and F1-score were calculated to measure the model's performance. The evaluation results are presented in X.

TABLE X. EVALUATION RESULTS OF SVM WITH MANUAL LABELING

	precision	recall	f1-score	support
Negative	0.61	0.57	0.59	106
Positive	0.85	0.87	0.86	292
Accuracy			0.79	398
Macro avg	0.73	0.72	0.72	398
Weighted avg	0.78	0.79	0.78	398

Based on the evaluation results of the SVM model using two labeling approaches VADER-based lexicon labeling table VIII and manual labeling table X a significant difference in model performance was observed. The results of both labeling strategies demonstrate a substantial performance gap, which is analyzed in further

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i3.2409, Copyright ©2025
Submitted : June 2, 2025, Revised : June 25, 2025, Accepted : July 1, 2025, Published : July 28, 2025

detail in the following section.

IV. DISCUSSION

The results demonstrate a clear performance advantage of the SVM model trained using VADER-labeled data, with an accuracy of 0.97. The model performed consistently well in both sentiment classes, achieving F1-scores of 0.93 for negative and 0.98 for positive sentiment. This high performance can be attributed to the consistency and objectivity of VADER's lexicon-based labeling, which applies sentiment rules uniformly across texts. As a result, the training data contains clearer sentiment signals, enabling the classifier to detect patterns more effectively.

In contrast, the model trained with manually labeled data derived from IMDb rating conversions showed significantly lower performance, especially in classifying negative sentiment F1-score: 0.59. This may be due to inconsistencies in human interpretation, where users might assign high ratings but express mixed or negative opinions in text. Such subjectivity introduces noise in the labeling process, affecting the model's ability to generalize.

These findings emphasize that automatic lexicon-based labeling, despite its limitations, offers greater consistency and scalability for sentiment analysis tasks. In domains like video game adaptation reviews, where fans often use hyperbolic or emotionally charged language, VADER's ability to capture sentiment through rule-based processing becomes particularly advantageous. This also reduces the burden of manual annotation, allowing researchers or businesses to process large datasets efficiently without sacrificing reliability.

From a practical standpoint, these results can be applied in several areas. For content producers and streaming platforms, understanding sentiment trends with higher accuracy enables better targeting and personalization strategies. For media analysts, reliable sentiment classification helps assess audience reception in real time, especially during the launch phase of a show or film. Additionally, the approach demonstrated here can be adopted as a fast and cost-effective framework for early-stage sentiment screening in marketing, entertainment, or even political domains where public perception shifts rapidly.

V. CONCLUSION

This research demonstrates that combining VADER-based sentiment labeling with the Support Vector Machine (SVM) algorithm substantially enhances the accuracy of sentiment classification in the context of video game adaptation reviews. The model trained using VADER-generated labels achieved a performance accuracy of 0.97, notably higher than the 0.79 accuracy recorded by the model trained with sentiment labels converted manually from IMDb ratings. These results highlight the advantages of lexicon-based automatic labeling in providing consistent and reliable sentiment detection while minimizing subjectivity.

The practical implication of these findings is that content creators, streaming services, and media analysts can adopt VADER-SVM as an effective and scalable tool for analyzing audience sentiment without relying on manual review. Future research could explore whether combining lexicon-based methods like VADER with advanced deep learning models for example, a customized BERT model could further enhance the detection of contextual nuances, sarcasm, and domain-specific language. It would also be valuable to apply this approach to other niche entertainment domains or multilingual datasets to broaden its applicability.

REFERENCES

[1] R. R. Ilmi, F. Kurniawan, and S. Harini, "PREDIKSI RATING FILM IMDb MENGGUNAKAN DECISION TREE IMDb MOVIE RATING PREDICTION USING DECISION TREE," *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 10, no. 4, 2023, doi: 10.25126/jtiik.2023106615.

[2] A. Saputra, S. Alfarisi, and R. N. Kholifah, "ADAPTATION ANALYSIS OF AMERICAN VIDEO GAMES MORTAL KOMBAT INTO FILM MORTAL KOMBAT REBOOT 2021," *Journal of Adaptation Studies*, vol. 2, no. 2, 2022.

[3] S. D. Wahyuni and R. H. Kusumodestoni, "Optimalisasi Algoritma Support Vector Machine (SVM) Dalam Klasifikasi Kejadian Data Stunting," *Bulletin of Information Technology (BIT)*, vol. 5, no. 2, pp. 56–64, 2024, doi: 10.47065/bit.v5i2.1247.

[4] R. P. Kurnia and Y. A. Atma, "ANALISIS REKOMENDASI FILM DARI DATA IMDB MENGGUNAKAN PYTHON," *Journal Of Information System, Computer Science and Information Technology*, vol. 3, no. 2, pp. 23–28, 2022.

[5] L. M. Huizen, M. Basyier, and R. Artikel, "Meningkatkan kinerja SVM: Dampak berbagai teknik seleksi fitur pada akurasi prediksi," *AITI: Jurnal Teknologi Informasi*, vol. 22, no. Maret, pp. 1–14, 2025.

[6] M. H. Aufan, M. R. Handayani, A. B. Nurjanna, N. C. H. Wibowo, and K. Umam, "THE PERCEPTIONS OF SEMARANG FIVE STAR HOTEL TOURISTS WITH SUPPORT VECTOR MACHINE ON GOOGLE REVIEWS," *Jurnal Teknik Informatika (Jutif)*, vol. 5, no. 5, pp. 1241–1247, Oct. 2024, doi: 10.52436/1.jutif.2024.5.5.2025.

[7] D. Abimanyu, E. Budianita, E. P. Cynthia, F. Yanto, and Yusra, "Analisis Sentimen Akun Twitter Apex Legends Menggunakan VADER," *Jurnal Nasional Komputasi dan Teknologi Informasi*, vol. 5, no. 3, 2022.

[8] S. Ernawati and R. Wati, "Evaluasi Performa Kernel SVM dalam Analisis Sentimen Review Aplikasi ChatGPT Menggunakan Hyperparameter dan VADER Lexicon," 2024.

[9] V. Nurcahyawati and Z. Mustaffa, "Vader Lexicon and Support Vector Machine Algorithm to Detect Customer Sentiment Orientation," *Journal of Information Systems Engineering and Business Intelligence*, vol. 9, no. 1, 2023, doi: 10.20473/jisebi.9.1.109-119.

[10] M. Taufiq Anwar and D. Riandhita Arief Permana, "Analisis Sentimen Masyarakat Indonesia Terhadap Produk Kendaraan Listrik Menggunakan VADER," *Jurnal Teknik Informatika dan Sistem Informasi*, vol. 10, no. 1, pp. 783–792, 2023.

[11] A. N. Zhafira, N. Afifah, S. A. Putri, V. Marhalatun, and D. I. S. S. Saputra, "ANALISIS SENTIMEN ULASAN FILM PADA DATASET IMDB MENGGUNAKAN ALGORITMA NAIVE BAYES INFO ARTIKEL," vol. 4, no. 1Juni, pp. 373–383, 2025, doi: 10.70247/jumistik.v4i1.139.

[12] N. J. Ramadhan, A. P. P. Putri, and D. A. Riyadi, "Eksplorasi Analisis Sentimen pada Rating Film IMDb: Pendekatan Perbandingan menggunakan Decision Tree dan Naive Bayes," *Journal Of Social Science Research*, vol. 4.

[13] F. Amaliah and I. K. D. Nuryana, "Perbandingan Akurasi Metode Lexicon Based Dan Naive Bayes Classifier Pada Analisis Sentimen Pendapat Masyarakat Terhadap Aplikasi Investasi Pada Media Twitter," *Journal of Informatics and Computer Science*, vol. 03, 2022.

[14] F. Alghifari and D. Juardi, "Penerapan Data Mining Pada Penerapan Data Mining Pada Penjualan Makanan Dan Minuman Menggunakan Metode Algoritma Naïve Bayes," *Jurnal Ilmiah Informatika*, vol. 9, 2021.

[15] M. U. Albab, Y. K. P., and M. N. Fawaiq, "Optimization of the Stemming Technique on Text Preprocessing President 3 Periods Topic," *Jurnal Transformatika*, vol. 20, no. 2, pp. 1–12, Jan. 2023, doi: 10.26623/transformatika.v20i2.5374.

[16] M. Azhari and Parjito, "Analisis Sentimen Opini Publik Program Makan Siang Gratis dengan Random Forest Pada Media," *Building of*

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i3.2409, Copyright ©2025
Submitted : June 2, 2025, Revised : June 25, 2025, Accepted : July 1, 2025, Published : July 28, 2025

Informatics, Technology and Science (BITS), vol. 6, no. 3, pp. 1932–1942, Dec. 2024, doi: 10.47065/bits.v6i3.6423.

[17] A. Daud, D. Irwanto, M. Said, M. Mubyl, and Mustamin, “Identification Of Chatbot Usage In Online Store Services Using Natural Language Processing Methods,” *Advance Sustainable Science, Engineering and Technology*, vol. 6, no. 2, Feb. 2024, doi: 10.26877/asset.v6i2.18309.

[18] D. Khyani, S. BS, N. NM, D. BM, and M. YM, “An Interpretation of Lemmatization and Stemming in Natural Language Processing,” *Journal of University of Shanghai for Science and Technology*, vol. 22, 2020.

[19] T. R. Salsabilla and N. Pratiwi, “Penerapan Support Vector Machine Untuk Analisis Sentimen pada X (Twitter) Mengenai Obat Penyebab Gagal Ginjal Akut pada Anak,” *Jurnal Teknik Informatika dan Komputer*, vol. 3, 2024.

[20] D. Septiani and I. Isabela, “ANALISIS TERM FREQUENCY INVERSE DOCUMENT FREQUENCY (TF-IDF) DALAM TEMU KEMBALI INFORMASI PADA DOKUMEN TEKS,” *SINTESIA: Jurnal Sistem dan Teknologi Informasi Indonesia*, vol. 1, 2022.

[21] T. A. E. Putri, T. Widiharih, and R. Santoso, “PENERAPAN TUNING HYPERPARAMETER RANDOMSEARCHCV PADA ADAPTIVE BOOSTING UNTUK PREDIKSI KELANGSUNGAN HIDUP PASIEN GAGAL JANTUNG,” *Jurnal Gaussian*, vol. 11, no. 3, pp. 397–406, Jan. 2023, doi: 10.14710/j.gauss.11.3.397-406.

[22] P. L. Romadloni, B. A. Kusuma, and W. M. Baihaqi, “KOMPARASI METODE PEMBELAJARAN MESIN UNTUK IMPLEMENTASI PENGAMBILAN KEPUTUSAN DALAM MENENTUKAN PROMOSI JABATAN KARYAWAN,” *Jurnal Mahasiswa Teknik Informatika*, vol. 6, 2022.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i3.2409, Copyright ©2025
Submitted : June 2, 2025, Revised : June 25, 2025, Accepted : July 1, 2025, Published : July 28, 2025