International Journal of Innovative Research and Scientific Studies, 8(7) 2025, pages: 64-71

IJIRSS
International Journal of Innovative Research & Scientific Studies

ISSN: 2617-6548

URL: www.ijirss.com

Comparison of machine learning and deep learning algorithms on sentiment analysis in game reviews

Albertus Arga Soetasad[1], Erick Fernando[2*]

[1,2]*Department of Information System, Faculty of Engineering and Informatics, Universitas Multimedia Nusantara, Tangerang, Indonesia.*

Corresponding author: Erick Fernando (*Email: Erick.fernando@umn.ac.id*)

Abstract

This study aims to analyze player review sentiment for the Marvel Rivals game on Steam using machine learning and deep learning algorithms, including Random Forest, Naive Bayes, XGBoost, and Bi-LSTM. The research was conducted within the CRISP-DM framework, which encompasses understanding the business problem, data exploration, data preparation, model building, and evaluation and implementation. Player review data was collected through web scraping, then preprocessed to clean and reformat the text before being used to train a sentiment classification model. Model evaluation was performed using metrics such as accuracy, precision, recall, and F1-score to identify the most effective model. The results indicated that Bi-LSTM was the best performing model, achieving an accuracy of 89% and an F1-score of 0.72 for negative sentiment. Hyperparameter tuning on real data contributed significantly to this performance. Conversely, applying SMOTE to balance the dataset actually reduced the performance of the Bi-LSTM model, suggesting that parameter optimization is more effective than synthetic data balancing, particularly for deep learning models.

Keywords: BiLSTM, Game review, Multinomial Naïve Bayes, Random forest, Sentiment analysis, XGBoost.

DOI: 10.53894/ijirss.v8i7.10401
Funding: This study received no specific financial support.
History: Received: 18 August 2025 / **Revised:** 22 September 2025 / **Accepted:** 24 September 2025 / **Published:** 1 October 2025

Competing Interests: The authors declare that they have no competing interests.
Authors' Contributions: All authors contributed equally to the conception and design of the study. All authors have read and agreed to the published version of the manuscript.
Transparency: The authors confirm that the manuscript is an honest, accurate, and transparent account of the study; that no vital features of the study have been omitted; and that any discrepancies from the study as planned have been explained. This study followed all ethical practices during writing.
Acknowledgements: Thank you to Multimedia Nusantara University for providing space and time for this research process.
Publisher: Innovative Research Publishing

1. Introduction

The digital gaming industry has seen rapid growth in recent years, with an increasing number of games being released and a growing player community. One game that has attracted attention is Marvel Rivals, a hero-based shooter that combines characters from Marvel comics into a competitive and dynamic gaming experience. Since its launch on December 5, 2024, Marvel Rivals has garnered widespread attention, reaching 10 million players within 72 hours [1, 2]. Since then, many people have voiced their opinions about Marvel Rivals through various online platforms.

Player opinions about a game are important to both game developers and other players. Developers can learn what players like and dislike and consider improving game features, such as gameplay, story, graphic quality, and optimization [3, 4]. For new players, understanding other players' sentiments can help them decide whether to invest time or money in the game [5, 6]. One approach to understanding the voice of consumers or users is sentiment analysis. Sentiment analysis is a technique for categorizing a person's emotions as negative (-1), neutral (0), or positive (1) [7, 8]. The data commonly used for sentiment analysis is text, which can be collected from various sources [9, 10]. In the case of Marvel Rivals, many players voice their opinions through the Steam platform. Sentiment analysis can provide significant benefits in the gaming industry, allowing developers and researchers to gain in-depth insights into player preferences and sentiments toward a game [3, 11, 12]. By monitoring sentiment trends over time, developers can anticipate changes in player satisfaction and engagement, which is crucial in a live service environment where continuous content updates and player interaction are crucial [13]. The used an aspect-based sentiment analysis framework to extract insights from esports game reviews, uncovering player preferences and areas for improvement in game design. They identified 16 topics for each game, categorized into gameplay-related topics (GRTs) and player-related topics (PRTs). GRT topics include graphics, characters, maps, optimization, updates, and gameplay, while PRT topics include skills, cheating, servers, rankings, and community. Analysis of topic prevalence indicates that players value community, skills, and server performance, while also having concerns about issues such as cheating and matchmaking. Several studies have applied machine learning models to classify sentences into different sentiments, such as positive and negative. Setiawan, et al. [14] analyzed 2,000 user reviews of Growtopia using Naive Bayes, resulting in an accuracy of 87%. Purbolaksono [15] used Random Forest to analyze game reviews on Steam using TF-IDF Bigrams, resulting in an F1-score of 62%.

This study aims to analyze the sentiment of player reviews of Marvel Rivals using machine learning, deep learning, and NLP techniques. By understanding the sentiment patterns that emerge in reviews, this research is expected to contribute to the gaming industry in optimizing the gaming experience.

2. Methods

This section systematically explains the stages taken in the research to classify user sentiment towards the Marvel Rivals game on the Steam platform. This research uses a quantitative approach by applying Natural Language Processing (NLP)-based data mining techniques to analyze user review text data with machine learning and deep learning algorithms.

One of the main advantages of using machine learning algorithms in sentiment analysis is their ability to efficiently handle large datasets. Modern algorithms, including Naïve Bayes, Support Vector Machines, and Logistic Regression (LR), have demonstrated high levels of accuracy when applied to the task of classifying game review sentiment [16]. One popular machine learning model used in sentiment analysis is the Support Vector Machine (SVM) [17-19]. The SVM algorithm aims to find a hyperplane that best separates different sentiment classes in a feature space [4]. Another common model is the Naïve Bayes classifier, known for its simplicity and efficiency [14, 20, 21]. This probabilistic model applies Bayes' theorem, assuming independence among features, which makes it efficient to train [22]. Random Forest (RF), another ensemble learning method, operates by constructing multiple decision trees to make predictions [23]. RF has become a popular algorithm in sentiment analysis [16] because its ensemble nature, coupled with Random Forest, increases its resilience to overfitting, especially when dealing with complex datasets caused by noise and high variance [24]. Other ensemble methods such as XGBoost are also being explored, combining predictions from multiple weak learners to improve overall prediction performance [16, 25]. This algorithm focuses on errors made by previous models, progressively improving prediction accuracy [26] but its use in game review sentiment classification is still limited.

One popular deep learning algorithm in sentiment analysis is the Convolutional Neural Network (CNN). CNNs are used for their strength in feature extraction, employing multiple layers of convolutional filters to detect patterns in textual data [27-29] highlights that CNNs can achieve significant performance improvements over traditional models in sentiment classification tasks by efficiently recognizing contextual clues and sentiment-containing phrases. Another popular algorithm is Long Short-Term Memory (LSTM), which overcomes some of the limitations of Recurrent Neural Networks (RNNs) [30]. LSTMs are adept at capturing information in sequential data, ensuring that contextual information across sentences or phrases is taken into account in predictions Al-Selwi, et al. [31]. Dashtipour, et al. [32] LSTMs have been shown to outperform several other conventional models such as Multilayer Perceptron (MLP), SVM, LR, and CNN. Bidirectional Long Short-Term Memory (BiLSTM) networks enhance the capabilities of LSTMs by processing data in both forward and backward directions [33]. This bidirectional approach allows the model to consider both past and future context in making predictions, making it capable of understanding complex sequential data [34] often found in natural language. The research process is designed following the CRISP-DM (Cross-Industry Standard Process for Data Mining) framework, which consists of six main stages: business understanding, data understanding, data preparation, modeling, evaluation, and deployment.

2.1. Research Framework

This study uses a quantitative approach with natural language processing (NLP)-based data mining methods to analyze user sentiment toward the Marvel Rivals game on the Steam platform. The primary objective of this study is to build a classification model capable of identifying positive and negative sentiment based on text reviews provided by users.

The research process follows the CRISP-DM (Cross-Industry Standard Process for Data Mining) framework, which consists of six main stages: business understanding, data understanding, data preparation, modeling, evaluation, and deployment [35]. This framework was chosen because it provides a structured and systematic approach to guide the entire data mining process, from understanding the problem to implementing the final solution [36].

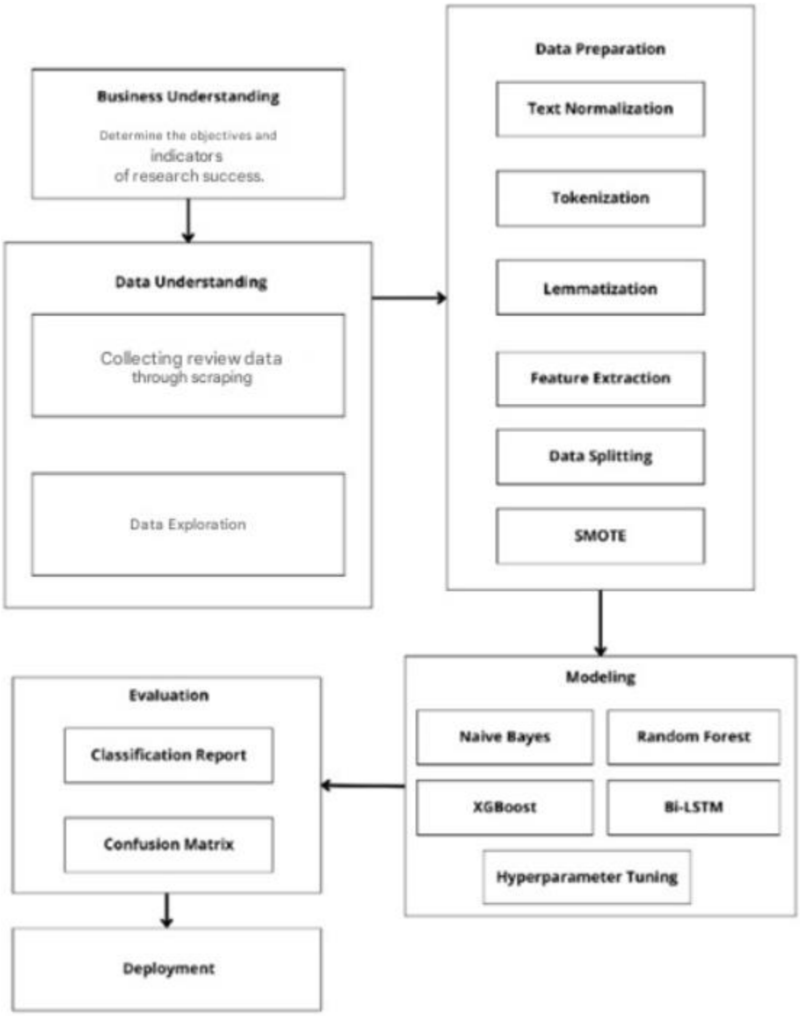

Figure 1.
Research Workflow.

The data used was collected through web scraping from Marvel Rivals review pages on Steam. It then underwent text preprocessing processes such as data cleaning, tokenization, normalization, and lemmatization before being used in model training. During the modeling process, several machine learning and deep learning algorithms were used to compare sentiment classification performance. These algorithms included Random Forest, Multinomial Naive Bayes (MNB), Extreme Gradient Boosting (XGBoost), and Bidirectional Long Short-Term Memory (Bi-LSTM). To determine the best model, experiments were conducted using Synthetic Minority Oversampling (SMOTE) and hyperparameter tuning.

2.2. Business Understanding

Marvel Rivals is a multiplayer hero shooter game developed by NetEase Games in collaboration with Marvel Entertainment. The game pits various characters from the Marvel universe against each other in 6v6 team battles. As a relatively new game, understanding player responses and sentiment through their reviews is crucial for ongoing development and marketing strategy. These reviews reflect players' experiences with key aspects of the game, such as gameplay mechanics, character balance, graphics, and the monetization system. The goal of this research, from a developer perspective, is to understand players' perceptions of Marvel Rivals based on their reviews, identify aspects of the game that received positive and negative responses, and provide recommendations for game improvements and marketing strategies based on sentiment analysis. The success criteria for this research are a sentiment classification model accuracy of at least 80% and concrete, actionable recommendations for game improvements based on the analysis.

2.3. Data Understanding

The data used in this study comes from the Marvel Rivals game review page on the Steam platform. This data was obtained using web scraping techniques assisted by the website Apify. The columns used in this analysis are reviews, which contain player feedback, and voted_up, where a value of 'True' indicates a player's liking of Marvel Rivals, and a value of 'False' indicates a player's disliking of it. There are 166,377 rows of data.

Table 1.
Example of Marvel Rivals Review.

	Review	Voted_Up
1	Pretty Good Unless enemy team goes Triple Support...	True
2	The matchmaking in this game is so broken. There should absolutely be placement matches at the beginning of your season	False
3	Extremely toxic player base. The games without a fuming playing who crashes out on their first death…	False

After removing data containing empty and duplicate reviews, there were 98,808 positive reviews and 25,818 negative reviews, indicating an imbalance in the dataset. This prompted a consideration to perform SMOTE before the modeling stage to balance the two classes.

2.4. Data Preparation

The data was processed before being used in the modeling stage. The text was cleaned of unnecessary characters and all words were lowercased. Afterward, the text was tokenized, converting sentences into chunks of words. This was done to allow for stopword removal and lemmatization. Next, sentiment encoding was performed to convert the data into a numeric format. Positive sentiment was converted to '1' and negative sentiment to '0'. The text was then processed into numeric features using Term Frequency-Inverse Document Frequency (TF-IDF). The max_features parameter was set to 5000, meaning only the 5000 most relevant words were retrieved. Due to text processing, some of the data contained empty reviews. Rows with blank reviews were removed, resulting in a final data set of 123,206 rows (97,570 positive and 25,636 negative).

The dataset was divided into a training set, a validation set, and a test set. 80% of the data was used as the training set, 20% as the validation set, and 20% as the test set. The training set was used to train the model, the validation set was used for hyperparameter tuning, and the test set was used to evaluate the model's performance on unseen data.

2.5. Modeling

Four algorithms were tested to classify sentiment from player reviews of Marvel Rivals: Random Forest (RF), Multinomial Naive Bayes (MNB), XGBoost, and Bidirectional Long Short-Term Memory (Bi-LSTM). Each model was evaluated in four different scenarios: without SMOTE and without hyperparameter tuning, with SMOTE without hyperparameter tuning, without SMOTE with hyperparameter tuning, and with SMOTE and hyperparameter tuning.

In the first scenario (without SMOTE and without hyperparameter tuning), a Random Model was created with default parameters of n_estimators = 100, max_depth = None, with an additional random_state = 42; the MNB model was created with default parameters of alpha = 1; and the XGBoost model was created with default parameters of n_estimators = 100. The Bi-LSTM model consists of an Embedding layer, which converts words in the vocabulary (vocab_size) into 100-dimensional vectors (embedding_dim). Then, a Bidirectional Bi-LSTM layer with 64 units was added, allowing the model to understand context in both directions (forward and backward in the text). After that, there is a Dense layer with sigmoid activation, which is used to generate binary predictions (e.g., positive or negative sentiment). The model was compiled using the Adam optimizer, with a binary crossentropy loss function (since the task was binary classification) and an accuracy metric for evaluating model performance. The model had a total of 3,948,409 trainable parameters. The training process lasted for 5 epochs, meaning the model would view the entire dataset 5 times to update its weights. Batch_size = 32 specified that the data would be processed in batches of 32 samples per iteration, rather than the entire dataset at once. In the second scenario (with SMOTE without hyperparameter tuning), the model was initialized with the same parameters as the first scenario, but using a training set that had been balanced with SMOTE. In the third scenario (without SMOTE and with hyperparameter tuning), several variations of the initial model were created to improve model accuracy by changing the values of several parameters. For the Random Forest model, GridSearchCV was performed with a combination of n_estimators (100), max_depth (None), min_samples_split (5, 10), and min_samples_leaf (2, 4). The min_samples_split parameter controls the minimum number of samples required to split a node. For the MNB model, GridSearchCV was performed on the alpha parameter (0.1, 0.5, 1.0, 5, 10). The alpha parameter controls Laplace smoothing, where a positive value is added to avoid zero probabilities. For the XGBoost model, GridSearchCV was performed on the n_estimators (100), maxdepth (3, 6, 10), and learning_rate (0.01, 0.1, 0.3) parameters. The n_estimators parameter controls the number of trees used in the model. The max_depth parameter controls the depth of the trees, and the learning_rate parameter controls how fast the model learns. For the Bi-LSTM model, Keras Tuner was used for hyperparameter tuning. The function is created by defining a model with an Embedding layer, Bi-LSTM, and Dense layer, where the embedding size (50, 100, 200), the number of LSTM units (32, 64, 128), and the learning rate (0.001, 0.003, 0.005) are automatically selected from a list of options. The model is compiled with the Adam optimizer and binary crossentropy loss. The search for the best hyperparameter combination was performed in 10 trials (max_trials = 10), with each trial run once (executions_per_trial = 1). In the fourth scenario (with SMOTE and hyperparameter tuning), hyperparameter tuning was performed as in the third scenario, but with a training set that had been balanced using SMOTE.

2.6. Evaluation

Evaluation was conducted based on accuracy metrics and the F1-score for each sentiment label (positive and negative), to assess the balance between precision and recall. This approach aims to identify the model with the most optimal performance in handling data imbalance and producing reliable sentiment classification. Additionally, a confusion matrix was created to determine the amount of misclassified data.

2.7. Deployment

The deployment phase was carried out using the Python programming language in combination with several supporting libraries such as Streamlit, TensorFlow, NLTK, LangChain, and ChromaDB. The system is developed in the form of an interactive web-based application that is capable of performing the process of retrieving review data from the Steam platform, data processing, sentiment prediction using the Bi-LSTM model, and sentiment explanation using the Retrieval-Augmented Generation (RAG) method based on LLM Llama 3. The embedding model used is sentence-transformers/all-MiniLM-L6-v2 from HuggingFace, which is a pre-trained model that functions to convert review text into a fixed-dimensional vector representation that contains the semantic meaning of the text.

3. Results and Discussion

This section presents the experimental results of four algorithms for classifying sentiment from Marvel Rivals player reviews: RF, MNB, XGBoost, and Bi-LSTM. Each model was evaluated in four different scenarios: without SMOTE and without hyperparameter tuning, with SMOTE but without hyperparameter tuning, without SMOTE but with hyperparameter tuning, and with SMOTE and hyperparameter tuning. These scenarios were designed to assess how data balancing and parameter optimization affect model performance. The goal was to determine which combination yielded the most accurate and reliable sentiment classification.

3.1. Suite the algorithm

The first scenario was without SMOTE and without hyperparameter tuning. The best model among the models without SMOTE and without hyperparameter tuning was the Bi-LSTM model with 88% accuracy. Despite having similar accuracy to RF and XGBoost, Bi-LSTM has the highest F1-score of all models for the 0/negative labels (71%), indicating a better balance between precision and recall.

Table 2.
Confusion Matrix for the First Scenario.

	TP	**TN**	**FP**	**FN**
RF	18.561	3.168	1.989	924
MNB	18.833	2.725	2.423	652
XGBoost	18.799	2.882	2.275	686
Bi-LSTM	18.255	3.525	1.632	1.230

Table 3.
Classification Report First Scenario.

	Accuracy	**0**			**1**		
		Pre	**Rec**	**F1**	**Pre**	**Rec**	**F1**
RF	0.88	0.77	0.61	0.69	0.90	0.95	0.93
MNB	0.87	0.81	0.53	0.64	0.89	0.97	0.92
XGBoost	0.88	0.81	0.56	0.66	0.89	0.96	0.93
Bi-LSTM	0.88	0.74	0.68	0.71	0.92	0.94	0.93

3.2. Algorithm Analysis with Smote and Without Tuning

In the second scenario, using SMOTE without hyperparameter tuning, the best model was RF with an accuracy of 87%. The highest F1-score for label 0 (negative) was 70%.

Table 4.
Confusion Matrix for Second Scenario.

	TP	**TN**	**FP**	**FN**
RF	17.460	3.876	1.281	2025
MNB	15.481	4.359	798	4.004
XGBoost	17.688	3.664	1.493	1.797
Bi-LSTM	14.310	4.206	951	5.175

Table 5.
Classification Report Second Scenario.

		0			1		
	Accuracy	**Pre**	**Rec**	**F1**	**Pre**	**Rec**	**F1**
RF	0.87	0.66	0.75	0.70	0.93	0.90	0.91
MNB	0.81	0.52	0.85	0.64	0.95	0.79	0.87
XGBoost	0.87	0.67	0.71	0.69	0.92	0.91	0.91
Bi-LSTM	0.75	0.45	0.82	0.58	0.94	0.73	0.82

3.3. Algorithm Analysis Without Smote and With Tuning

The third scenario is without SMOTE but with hyperparameter tuning. In this scenario, Bi-LSTM is again the best model, with an accuracy increase of 89%. Furthermore, the F1-score for label 0 (negative) also increases to 72%.

Table 6.
Confusion Matrix for Third Scenario.

	TP	**TN**	**FP**	**FN**
RF	18.724	3.011	2.146	761
MNB	18.763	2.820	2.337	722
XGBoost	18.722	3.063	2.094	763
Bi-LSTM	18.195	3.625	1.532	1.290

Table 7.
Classification Report Third Scenario.

		0			1		
	Accuracy	Pre	Rec	F1	Pre	Rec	F1
RF	0.88	0.80	0.58	0.67	0.90	0.96	0.93
MNB	0.88	0.80	0.55	0.65	0.89	0.96	0.92
XGBoost	0.88	0.80	0.59	0.68	0.90	0.96	0.93
Bi-LSTM	0.89	0.74	0.70	0.72	0.92	0.93	0.93

3.4. Algorithm Analysis with Smote and Tuning

The fourth scenario involves applying SMOTE and hyperparameter tuning. The best model in this scenario is XGBoost with an accuracy of 87% and an F1-score for label 1 (positive) of 92%.

Table 8.
Confusion Matrix for the Fourth Scenario.

	TP	**TN**	**FP**	**FN**
RF	17.443	3.953	1.204	2.042
MNB	15.469	4.343	814	4.016
XGBoost	17.845	3.713	1.444	1.640
Bi-LSTM	14.122	4.310	847	5.363

Table 9.
Classification Report Fourth Scenario.

		0			**1**		
	Accuracy	**Pre**	**Rec**	**F1**	**Pre**	**Rec**	**F1**
RF	0.87	0.66	0.77	0.71	0.94	0.90	0.91
MNB	0.80	0.52	0.84	0.64	0.95	0.79	0.86
XGBoost	0.87	0.69	0.72	0.71	0.93	0.92	0.92
Bi-LSTM	0.75	0.45	0.84	0.58	0.94	0.72	0.82

3.5. Discussion

The experiments in this study were conducted across four scenarios to find the best model. Scenario 1, which used default parameters on the original imbalanced data, served as a baseline, with the Bi-LSTM model showing the best F1-score for the negative class (0.71), although several other models achieved comparable accuracy of 88%. The main limitation of this scenario was the low recall of the models for negative reviews. To address this, Scenario 2 implemented the SMOTE technique, which successfully improved the F1-score of machine learning models like XGBoost (from 0.66 to 0.69), but significantly decreased the performance of the Bi-LSTM (accuracy dropping from 88% to 75%). Given these contrasting results and the continued use of default parameters, Scenario 3 focused on hyperparameter tuning on the original data. This scenario was the most successful, with the optimized Bi-LSTM model achieving peak performance in the study, with 89% accuracy and the highest negative F1-score of 0.72. As a final confirmation step, Scenario 4 combines SMOTE with hyperparameter tuning, but the results are unable to surpass Scenario 3 and again show poor performance on Bi-LSTM (75% accuracy). The performance results of each model in all scenarios can be seen.

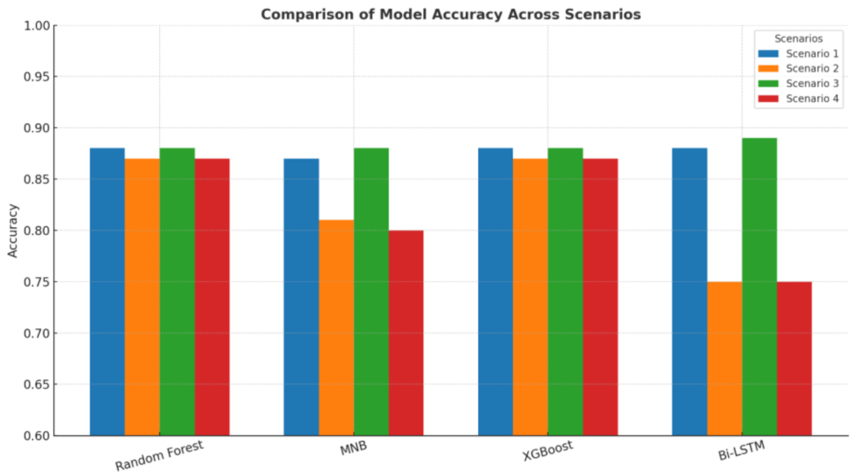

Figure 2.
Comparison of Model Accuracy Across Scenarios.

4. Conclusion

The evaluation results of the four scenarios, it can be concluded that the Bi-LSTM model with hyperparameter tuning is the best model for classifying player review sentiment for Marvel Rivals. This model demonstrated the most balanced performance, with the highest accuracy of 89% and an F1-score of 72% for negative labels and 93% for positive labels. This indicates that deep learning models like Bi-LSTM, especially after hyperparameter tuning, are able to capture sentence context and patterns in text data more effectively than traditional machine learning models. Conversely, machine learning models like Random Forest and XGBoost showed more stable performance improvements when SMOTE was applied, particularly in improving classification ability for minority classes. However, performance improvements were more optimal when SMOTE was combined with tuning, as seen in the fourth scenario, where XGBoost achieved the highest F1-score for the positive class (92%). Using SMOTE can assist with data balancing, but it does not always result in better performance, especially for deep learning models that are sensitive to the distribution of synthetic data. Meanwhile, Multinomial Naive Bayes (MNB) consistently demonstrated the lowest performance among all models, both in terms of accuracy and F1-score. While this model is simple and efficient, its limitations in handling complex features and its reliance on strong data distribution assumptions make it less suitable for more dynamic and imbalanced text data such as gamer reviews. Therefore, an approach that combines model selection appropriate to the data characteristics and appropriate hyperparameter tuning is key to producing an accurate and balanced sentiment classification model.

References

[1] W. Yin-Poole, "Marvel rivals is a smash hit with 10 million players in just 3 days," 2025. https://sea.ign.com/marvel-rivals/223084/news/marvel-rivals-is-a-smash-hit-with-10-million-players-in-just-3-days. [Accessed Feb. 27, 2025]

[2] Marvel Rivals, "Marvel Rivals celebrates a successful launch, with NetEase Games' Overwatch rival hitting over 10 million players in just 3 days | GamesRadar+.. GamesRadar+," 2025. https://www.gamesradar.com/games/third-person-shooter/marvel-rivals-celebrates-a-successful-launch-with-netease-games-overwatch-rival-hitting-over-10-million-players-in-just-3-days/

[3] Y. Yu, D.-T. Dinh, B.-H. Nguyen, F. Yu, and V.-N. Huynh, "Mining insights from esports game reviews with an aspect-based sentiment analysis framework," *Ieee Access,* vol. 11, pp. 61161-61172, 2023. https://doi.org/10.1109/ACCESS.2023.3285864

[4] L. F. Britto and L. Pacífico, "Evaluating video game acceptance in game reviews using sentiment analysis techniques," *Proceedings of SBGames,* pp. 399-402, 2020.

[5] B. Agarwal, N. Mittal, P. Bansal, and S. Garg, "Sentiment analysis using common-sense and context information," *Computational Intelligence and Neuroscience,* vol. 2015, no. 1, p. 715730, 2015. https://doi.org/10.1155/2015/715730

[6] Z. H. Kilimci, H. Yörük, and S. Akyokus, "Sentiment analysis based churn prediction in mobile games using word embedding models and deep learning algorithms," in *2020 International Conference on Innovations in Intelligent Systems and Applications (INISTA) (pp. 1-7). IEEE*, 2020.

[7] H. Taherdoost and M. Madanchian, "Artificial intelligence and sentiment analysis: A review in competitive research," *Computers,* vol. 12, no. 2, p. 37, 2023. https://doi.org/10.3390/computers12020037

[8] I. M. Urriza and M. A. A. Clariño, "Aspect-based sentiment analysis of user created game reviews," in *2021 24th Conference of the Oriental COCOSDA International Committee for the Co-ordination and Standardisation of Speech Databases and Assessment Techniques (O-COCOSDA) (pp. 76-81). IEEE*, 2021.

[9] M. Godsay, "The process of sentiment analysis: A study," *International Journal of Computer Applications,* vol. 126, no. 7, pp. 26-30, 2015. https://doi.org/10.5120/IJCA2015906091

[10] N. Punetha and G. Jain, "Game theory and MCDM-based unsupervised sentiment analysis of restaurant reviews," *Applied Intelligence,* vol. 53, no. 17, pp. 20152-20173, 2023. https://doi.org/10.1007/s10489-023-04471-1

[11] J. Y. Tan, A. S. K. Chow, and C. W. Tan, "A comparative study of machine learning algorithms for sentiment analysis of game reviews," *IEM Journal,* vol. 82, no. 3, 2022. https://doi.org/10.54552/v82i3.101

[12] F. Alzami, E. D. Udayanti, D. P. Prabowo, and R. A. Megantara, "Document preprocessing with TF-IDF to improve the polarity classification performance of unstructured sentiment analysis," *Kinetik: Game Technology, Information System, Computer Network, Computing, Electronics, and Control,* vol. 5, no. 3, pp. 235-242, 2020.

[13] D. Sirbu, A. Secui, M. Dascalu, S. A. Crossley, S. Ruseti, and S. Trausan-Matu, "Extracting gamers' opinions from reviews," in *2016 18th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing (SYNASC) (pp. 227-232). IEEE*, 2016.

[14] R. P. Setiawan, B. Irawan, and W. P. Prihartono, "Sentiment analysis of Growtopia reviews on google play store using naïve bayes classifier to identify user needs," *Jurnal Informatika dan Teknik Elektro Terapan,* vol. 13, no. 2, 2025. https://doi.org/10.23960/jitet.v13i2.6415

[15] M. D. Purbolaksono, "Sentiment analysis of game review in steam platform using random forest," *International Journal on Information and Communication Technology,* vol. 10, no. 2, pp. 161-169, 2024. https://doi.org/10.21108/ijoict.v10i2.1007

[16] J. Y. Tan, A. S. K. Chow, and C. W. Tan, "Sentiment analysis on game reviews: A comparative study of machine learning approaches," in *International Conference on Digital Transformation and Applications (ICDXA2021)*, 2021.

[17] F. Wijaya, "Text-based sentiment analysis of Counter Strike 2 game with classification algorithm," Bachelor's Thesis. Universitas Multimedia Nusantara, Tangerang, Indonesia, 2024.

[18] N. C. Ramadani, "Sentiment analysis to measure mobile legend app user reviews Using Naive Bayes, SVM, random fores, decision tree, and logistic regression algorithms," *Jurnal Sistem Informasi (E-Journal),* vol. 16, no. 1, pp. 123–138, 2024. https://doi.org/10.18495/jsi.v16i1.152

[19] A. F. Panjalu, S. Alam, and M. I. Sulistyo, "Moba game review sentiment analysis using support vector machine algorithm," *Jurnal Informatika dan Komputer,* vol. 6, no. 2, pp. 1–8, 2023. https://doi.org/10.33387/jiko.v6i2.6388

[20] M. Y. Febrianta, S. Widiyanesti, and S. R. Ramadhan, "Analysis of local indie video game reviews on Steam using sentiment analysis and topic modeling based on latent dirichlet allocation," *Journal of Animation and Games Studies,* vol. 7, no. 2, pp. 117-144, 2021. https://doi.org/10.24821/jags.v7i2.5162

[21] B. J. Rizqullah, "Sentiment analysis of Apex Legends game reviews on Steam using the Naïve Bayes classifier," Thesis. Jenderal Soedirman University, 2024.

[22] P. B. Pajila, B. G. Sheena, A. Gayathri, J. Aswini, and M. Nalini, "A comprehensive survey on naive bayes algorithm: Advantages, limitations and applications," in *2023 4th International Conference on Smart Electronics and Communication (ICOSEC) (pp. 1228-1234). IEEE*, 2023.

[23] M. Aria, C. Cuccurullo, and A. Gnasso, "A comparison among interpretative proposals for Random Forests," *Machine Learning with Applications,* vol. 6, p. 100094, 2021. https://doi.org/10.1016/j.mlwa.2021.100094

[24] A. H. Salman and W. A. M. Al-Jawher, "Performance comparison of support vector machines, AdaBoost, and random forest for sentiment text analysis and classification," *Journal Port Science Research,* vol. 7, no. 3, pp. 300-311, 2024. https://doi.org/10.36371/port.2024.3.8

[25] I. Pramudia, D. Thio, A. Hendrawan, and M. Kom, "User sentiment analysis of arena breakout game on Google Playstore using machine learning," presented at the National Seminar on Innovation and Trends in Information Technology (SINATTI), 2024.

[26] K. Afifah, I. N. Yulita, and I. Sarathan, "Sentiment analysis on telemedicine app reviews using xgboost classifier," in *2021 International Conference on Artificial Intelligence and Big Data Analytics (pp. 22-27). IEEE*, 2021.

[27] S. Liao, J. Wang, R. Yu, K. Sato, and Z. Cheng, "CNN for situations understanding based on sentiment analysis of twitter data," *Procedia Computer Science,* vol. 111, pp. 376-381, 2017. https://doi.org/10.1016/j.procs.2017.06.037

[28] X. Ouyang, P. Zhou, C. H. Li, and L. Liu, "Sentiment analysis using convolutional neural network," in *2015 IEEE International Conference on Computer and Information Technology; Ubiquitous Computing and Communications; Dependable, Autonomic and Secure Computing; Pervasive Intelligence and Computing (pp. 2359-2364). IEEE*, 2015.

[29] L. Zhang, S. Wang, and B. Liu, "Deep learning for sentiment analysis: A survey," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery,* vol. 8, no. 4, p. e1253, 2018. https://doi.org/10.1002/widm.1253

[30] F. M. Shiri, T. Perumal, N. Mustapha, and R. Mohamed, "A comprehensive overview and comparative analysis on deep learning models," *Journal on Artificial Intelligence,* vol. 6, no. 1, pp. 301–360, 2024. https://doi.org/10.32604/jai.2024.054314

[31] S. M. Al-Selwi *et al.*, "RNN-LSTM: From applications to modeling techniques and beyond—Systematic review," *Journal of King Saud University-Computer and Information Sciences,* vol. 36, no. 5, p. 102068, 2024. https://doi.org/10.1016/j.jksuci.2024.102068

[32] K. Dashtipour, M. Gogate, A. Adeel, H. Larijani, and A. Hussain, "Sentiment analysis of persian movie reviews using deep learning," *Entropy,* vol. 23, no. 5, p. 596, 2021. https://doi.org/10.3390/e23050596

[33] U. Mahadevaswamy and P. Swathi, "Sentiment analysis using bidirectional LSTM network," *Procedia Computer Science,* vol. 218, pp. 45-56, 2023. https://doi.org/10.1016/j.procs.2022.12.400

[34] C. Pavlatos, E. Makris, G. Fotis, V. Vita, and V. Mladenov, "Enhancing electrical load prediction using a bidirectional LSTM neural network," *Electronics,* vol. 12, no. 22, p. 4652, 2023. https://doi.org/10.3390/electronics12224652

[35] A. Abujaber, A. Fadlalla, D. Gammoh, H. Abdelrahman, M. Mollazehi, and A. El-Menyar, "Prediction of in-hospital mortality in patients with post traumatic brain injury using National Trauma Registry and Machine Learning Approach," *Scandinavian Journal of Trauma, Resuscitation and Emergency Medicine,* vol. 28, p. 44, 2020. https://doi.org/10.1186/s13049-020-00738-5

[36] F. Martínez-Plumed *et al.*, "CRISP-DM twenty years later: From data mining processes to data science trajectories," *IEEE Transactions on Knowledge and Data Engineering,* vol. 33, no. 8, pp. 3048-3061, 2019. https://doi.org/10.1109/TKDE.2019.2962680