Inventa: Journal of Science, Technology, and Innovation
Vol 1 No 3 April 2026, Hal 212-224
ISSN: 3123-3147 (Print) ISSN: 3123-3155 (Electronic)
Open Access: https://scriptaintelektual.com/inventa

Performance Analysis of Logistic Regression and SVM (Support Vector Machine) Algorithms on E-Football Mobile Game Review Sentiment

Athary Zikry[1*], Farid Akbar Siregar[2]
[1-2] Muhammadiyah Sumatera Utara University, Indonesia
email: atharyzikry3009@email.com[1]

Article Info :
Received:
24-03-2026
Revised:
10-04-2026
Accepted:
20-04-2026

Abstract

This study investigates the comparative performance of Logistic Regression and Support Vector Machine algorithms in sentiment classification of e Football mobile game reviews collected from the Google Play Store. The research employed an empirical machine learning framework involving web scraping, Natural Language Processing preprocessing, Term Frequency Inverse Document Frequency feature extraction, and supervised classification procedures. A total of 7,497 Indonesian language reviews were processed through cleaning, case folding, normalization, tokenization, stopword removal, and stemming to improve textual consistency and semantic representation. The dataset was divided into training and testing subsets using an 80:20 ratio to evaluate model generalization performance. Classification effectiveness was measured using accuracy, precision, recall, and F1 score metrics supported by confusion matrix interpretation. The findings demonstrate that Support Vector Machine achieved superior classification stability with an accuracy of 81.13% and an F1 score of 0.72, while Logistic Regression obtained an accuracy of 80.97% and an F1 score of 0.71. The results indicate that Support Vector Machine provides stronger robustness in handling high dimensional Indonesian gaming review data characterized by class imbalance and semantic variability.

Keywords : *Sentiment Analysis, E-Football, Support Vector Machine, Logistic Regression, TF IDF.*

INTRODUCTION

The accelerated expansion of the global mobile gaming industry has fundamentally transformed digital entertainment into a data intensive ecosystem in which user generated reviews function not merely as expressions of preference but as strategic indicators of software quality, platform sustainability, and consumer loyalty within increasingly competitive digital markets. Within this landscape, football simulation games occupy a particularly influential position because they intersect entertainment technology, online competition, and the rapidly growing e sports economy, whose development has stimulated new forms of interaction between developers and users across digital distribution platforms. Recent developments in machine learning driven sports analytics further demonstrate that football related digital ecosystems are no longer examined solely from the perspective of gameplay mechanics, but increasingly through computational approaches capable of extracting behavioral and evaluative patterns from large scale datasets (Waskita et al., 2025). The emergence of e Football as one of the most widely played football simulation games on mobile platforms illustrates how software updates, gameplay modifications, and monetization systems can provoke polarized public reactions that are continuously documented in review sections on the Google Play Store. Such reviews contain latent evaluative structures reflecting satisfaction, frustration, and expectations toward gameplay performance, server stability, visual quality, and competitive fairness. The complexity of this textual ecosystem aligns with broader advances in Natural Language Processing, where computational linguistics has evolved toward sophisticated techniques for identifying semantic orientation and affective tendencies embedded within unstructured textual corpora (Jurafsky & Martin, 2023). In parallel, the growing dependence of information systems on large scale text mining has reinforced the relevance of information retrieval frameworks capable of transforming noisy digital interactions into

measurable knowledge structures that support evidence based decision making in technology driven industries (Manning et al., 2008).

Contemporary scholarship has increasingly positioned machine learning algorithms as central instruments for predictive modeling and classification tasks across both sports analytics and computational intelligence domains. Studies conducted within football analytics demonstrate that machine learning methods have been successfully implemented to predict player performance, tactical outcomes, and competitive results through the integration of statistical learning mechanisms and high dimensional feature spaces (Yu, 2022; Merzah et al., 2024). In broader computational contexts, Support Vector Machine models have shown remarkable effectiveness in handling classification tasks involving sparse and multidimensional datasets because the optimization of hyperplanes enables stronger generalization performance under complex feature distributions (Ponmalar & Dhanakoti, 2022). Similar observations emerge in applied predictive systems beyond textual analysis, where Support Vector Machine architectures have been employed to identify injury risks in sports environments with substantial predictive accuracy and robust classification stability (Shahlaei Bagheri, 2024). At the same time, Logistic Regression continues to receive substantial scholarly attention because its probabilistic interpretability, computational efficiency, and linear optimization structure often produce competitive performance in binary classification environments characterized by large textual datasets. Existing literature on sentiment analysis repeatedly indicates that both Logistic Regression and Support Vector Machine algorithms remain dominant baseline approaches in text classification because their statistical behavior is relatively transparent and computationally scalable compared with more computationally expensive deep learning architectures. Nevertheless, prior studies frequently evaluate these algorithms within generic review datasets, while only limited investigations critically compare their relative effectiveness in football oriented mobile gaming environments characterized by slang intensive expressions, abbreviated linguistic forms, and emotionally polarized feedback patterns.

Despite the increasing volume of research concerning sentiment analysis and machine learning classification, significant conceptual and empirical inconsistencies remain insufficiently addressed within current scholarship. A considerable proportion of existing studies prioritize general application performance while overlooking contextual linguistic variability emerging from domain specific digital communities, particularly those associated with competitive sports gaming cultures. Football gaming communities frequently employ informal vocabulary, hybrid linguistic expressions, sarcasm, and highly contextual abbreviations that challenge conventional preprocessing and vectorization strategies, creating semantic ambiguities that can substantially influence classification outcomes. Current literature also demonstrates methodological fragmentation regarding the comparative evaluation of Logistic Regression and Support Vector Machine algorithms because many investigations rely exclusively on accuracy metrics without critically examining the balance between precision, recall, and F1 score in datasets characterized by class imbalance. Such imbalance is particularly problematic in mobile application reviews where dissatisfaction during controversial software updates may temporarily dominate user discourse, thereby affecting model sensitivity toward minority sentiment classes. Existing machine learning studies within football related computational domains tend to focus predominantly on gameplay prediction, athlete performance, or tactical simulation rather than user perception analysis embedded within digital review ecosystems (Waskita et al., 2025; Yu, 2022). Meanwhile, computational linguistics scholarship emphasizes that sentiment interpretation cannot be detached from contextual semantic representation because lexical structures acquire meaning through sociotechnical usage environments rather than isolated token frequencies alone (Jurafsky & Martin, 2023). This gap indicates that the comparative behavior of traditional machine learning classifiers within Indonesian football gaming reviews remains theoretically underexplored and empirically insufficiently validated.

The unresolved limitations within prior research generate substantial scientific and practical implications because user reviews increasingly function as strategic resources influencing software development cycles, platform reputation management, and competitive sustainability in the digital gaming industry. In practical terms, developers require computational mechanisms capable of identifying dissatisfaction patterns rapidly and accurately in order to optimize updates, reduce user churn, and strengthen platform engagement within highly volatile mobile gaming markets. The inability to classify sentiment reliably may produce distorted interpretations of consumer experience, thereby weakening managerial responsiveness toward emerging technical problems such as gameplay lag,

connection instability, or monetization dissatisfaction. From a scientific perspective, the problem extends beyond application specific evaluation because sentiment classification in sports gaming environments represents a challenging case of high dimensional text processing involving informal linguistic structures and emotionally charged discourse. Research in information retrieval consistently highlights that textual relevance and semantic discrimination become increasingly difficult when documents contain noisy lexical distributions and contextual irregularities (Manning et al., 2008). Simultaneously, studies on intelligent classification systems emphasize that model robustness depends heavily on the interaction between feature representation and classifier architecture, particularly within sparse vector spaces derived from textual corpora (Merzah et al., 2024). The urgency of examining classification performance within e Football reviews therefore emerges not only from industrial demand for actionable insights but also from the broader need to refine methodological understanding regarding the adaptability of machine learning classifiers within culturally specific and domain sensitive textual environments.

This study positions itself within the intersection of sentiment analysis, sports informatics, and machine learning based text classification by critically examining the comparative performance of Logistic Regression and Support Vector Machine algorithms in analyzing Indonesian language reviews of the e Football mobile game. Unlike previous studies that emphasize predictive analytics in football performance or generalized sentiment classification tasks, the present research focuses specifically on user generated evaluative discourse within a football gaming ecosystem characterized by linguistic informality and dynamic sentiment polarity. The study adopts a systematic Natural Language Processing pipeline involving cleaning, case folding, normalization, tokenization, stopword removal, and stemming before transforming textual data into numerical representations through the Term Frequency Inverse Document Frequency method. This methodological configuration reflects contemporary computational linguistics principles asserting that meaningful sentiment extraction requires both semantic normalization and statistically informative feature weighting mechanisms (Jurafsky & Martin, 2023). The comparison between Logistic Regression and Support Vector Machine is strategically significant because both algorithms represent influential paradigms within supervised learning, yet they differ fundamentally in optimization logic, decision boundary construction, and sensitivity toward high dimensional sparse matrices. Existing evidence regarding Support Vector Machine superiority in multidimensional classification tasks suggests potential advantages for handling sentiment complexity in digital reviews (Ponmalar & Dhanakoti, 2022; Shahlaei Bagheri, 2024), while Logistic Regression remains valuable due to interpretability and computational efficiency. This research therefore contributes to the ongoing scholarly debate concerning whether linear probabilistic classifiers or hyperplane based optimization models provide more reliable sentiment discrimination within football related mobile gaming contexts.

This study aims to analyze and compare the performance of Logistic Regression and Support Vector Machine algorithms in classifying sentiment within e Football mobile game reviews collected from the Google Play Store using a TF IDF based feature extraction framework and comprehensive preprocessing procedures. The research contributes theoretically by extending the application of sentiment analysis within sports oriented digital ecosystems and by providing empirical insight into the behavior of conventional machine learning classifiers when confronted with linguistically informal and domain specific review datasets. Methodologically, the study offers a structured comparative framework integrating preprocessing optimization, vector space representation, and multidimensional evaluation metrics involving accuracy, precision, recall, and F1 score to assess classification reliability more comprehensively. The findings are expected to strengthen the methodological foundation for future sentiment analysis studies in gaming environments while simultaneously providing evidence based recommendations for developers seeking to understand user satisfaction dynamics through computational intelligence approaches.

RESEARCH METHODS

This study employed an empirical research design grounded in experimental machine learning procedures to evaluate the comparative performance of Logistic Regression and Support Vector Machine algorithms in sentiment classification of e Football mobile game reviews obtained from the Google Play Store. The research workflow began with automated web scraping procedures to collect 7,497 Indonesian language user reviews reflecting diverse perceptions regarding gameplay quality,

software updates, network stability, and user experience. The collected raw textual data were subsequently processed through a rigorous Natural Language Processing pipeline consisting of cleaning, case folding, normalization, tokenization, stopword removal, and stemming in order to reduce linguistic noise and improve semantic consistency within the dataset. Following preprocessing, textual documents were transformed into numerical vector representations using the Term Frequency Inverse Document Frequency method, enabling the extraction of discriminative lexical features from sparse textual structures. The transformed dataset was then divided into training and testing subsets using an 80:20 ratio through random sampling procedures to preserve distributional representativeness and minimize selection bias during model development. Logistic Regression and Support Vector Machine classifiers were subsequently trained using supervised learning mechanisms to identify positive and negative sentiment patterns embedded within user generated reviews. The methodological configuration adopted in this study reflects the statistical learning principles proposed by Bishop (2006), particularly regarding probabilistic pattern recognition, high dimensional classification, and generalization performance in machine learning systems, while the integration of domain specific preprocessing stages represents a methodological refinement tailored to the linguistic complexity of Indonesian football gaming discourse.

The validation process was conducted through systematic performance evaluation on the testing dataset in order to examine the robustness, stability, and predictive consistency of both classification models under realistic sentiment classification conditions. Model effectiveness was assessed using four complementary evaluation metrics consisting of accuracy, precision, recall, and F1 score because reliance on a single metric may produce misleading interpretations when dealing with imbalanced sentiment distributions commonly found in digital review datasets. Accuracy was employed to measure the overall correctness of classification outcomes, while precision and recall were utilized to evaluate the proportional reliability of positive and negative sentiment predictions across different classification boundaries. The F1 score was emphasized as the principal comparative metric because it provides a harmonic balance between precision and recall, thereby offering a more rigorous assessment of classifier stability under uneven class representation. Confusion matrix analysis was further implemented to identify classification errors, false positive tendencies, and false negative distributions generated by each algorithm, enabling deeper interpretative insight into model behavior beyond aggregate statistical scores. The methodological strength of this study lies in the integration of multilayer preprocessing procedures, high dimensional TF IDF feature engineering, and multidimensional validation strategies within a football specific mobile gaming environment that has rarely been examined in prior sentiment analysis research, thereby enhancing both the empirical reliability and contextual originality of the proposed analytical framework.

RESULTS AND DISCUSSION

Linguistic Distribution and Textual Normalization in Indonesian e Football User Reviews

The preprocessing stage revealed substantial linguistic variation within the Indonesian language reviews collected from the Google Play Store, particularly in relation to colloquial football terminology, abbreviated expressions, and emotionally charged evaluations embedded within player narratives. Several dominant lexical items such as “main,” “lag,” “update,” and “jaring” appeared repeatedly across the corpus, indicating that gameplay stability and network responsiveness constituted the central themes discussed by users. Similar linguistic complexity in mobile gaming discourse has been identified in multilingual sentiment datasets where informal syntax and domain specific vocabulary significantly influence classification effectiveness (Jurafsky & Martin, 2023). The normalization strategy implemented in this study reduced semantic fragmentation across equivalent lexical forms, thereby strengthening contextual consistency before feature extraction and model training.

The cleaning and case folding procedures substantially reduced textual redundancy generated by symbols, repeated characters, URLs, and inconsistent capitalization patterns frequently found in user generated content. Indonesian mobile gaming reviews commonly contain expressive exaggerations through duplicated letters and nonstandard abbreviations, which may distort statistical weighting if not normalized appropriately. Prior sentiment studies involving gaming platforms reported that preprocessing quality directly influences the robustness of machine learning classification because lexical ambiguity often increases dimensional sparsity within textual vectors (Kına & Özdağ, 2024). The preprocessing configuration adopted in this study therefore functioned not merely as a technical

preparation stage but as an analytical mechanism for preserving semantic relevance across heterogeneous review structures.

The tokenization and stopword removal processes generated a more compact lexical structure by eliminating function words that carried limited discriminative contribution toward polarity recognition. Several removed terms were identified as syntactically frequent yet semantically neutral expressions that appeared uniformly across both positive and negative sentiment categories. According to Manning et al. (2008), high frequency but low information words generally weaken discriminative weighting in sparse vector environments because they contribute minimal class differentiation. The removal of these tokens improved the concentration of meaningful football related sentiment indicators such as “bagus,” “buruk,” “lambat,” and “sinyal,” which later became dominant features within TF IDF transformation.

The stemming process demonstrated important effects on lexical standardization because Indonesian affixation patterns frequently generate multiple derivative forms with equivalent semantic meaning. Words associated with gameplay quality, connection performance, and update dissatisfaction were reduced into consistent root structures that simplified vector representation across documents. Comparable findings were reported in sentiment analysis research involving Mobile Legends and eFootball datasets where stemming significantly improved semantic consolidation during classification training (Alengka et al., 2025). The reduction of morphological variability contributed to lower feature dispersion and facilitated more stable machine learning generalization across the training corpus.

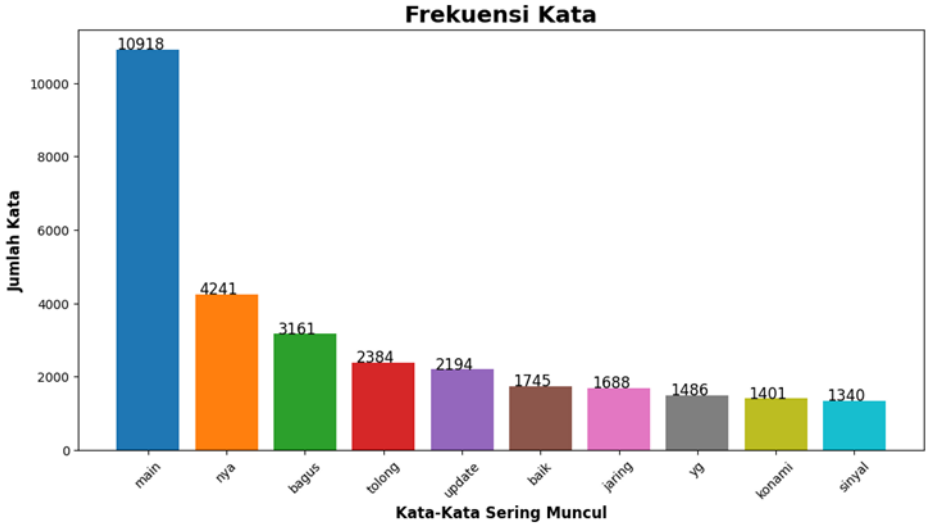

Figure 1. Data Frequency Diagram

The frequency distribution shown in Figure 1 demonstrated that the word “main” appeared as the most dominant lexical component with 10,918 occurrences, followed by “nya,” “bagus,” and “tolong.” This distribution indicated that gameplay experience represented the primary focus of user discussion, while complaint oriented vocabulary reflected dissatisfaction related to system performance and update stability. Studies examining football gaming environments reported that technical issues frequently dominate user generated sentiment because online gameplay quality strongly affects perceived enjoyment and competitiveness (Waskita et al., 2025). The lexical concentration identified in the frequency diagram therefore provided an empirical representation of the emotional structure embedded within the dataset.

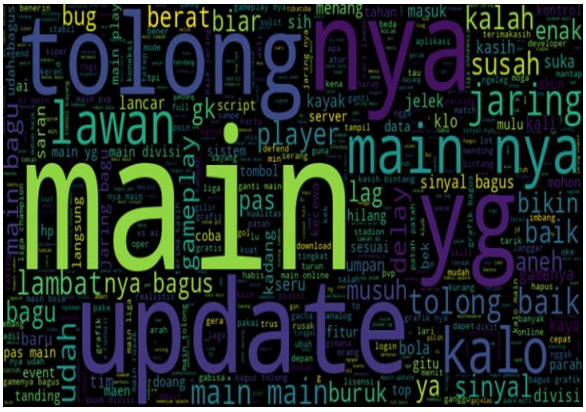

Figure 2. Worldcloud Stemming

The stemming based word cloud visualization presented in Figure 2 revealed the dominance of emotionally polarized terms associated with gameplay quality, update dissatisfaction, and network instability. Lexical items such as “main,” “lag,” “update,” and “buruk” occupied visually dominant positions, indicating high repetition intensity throughout the review corpus. Research concerning sentiment visualization methods explains that word clouds are effective for identifying latent thematic structures and semantic concentration within large scale textual datasets (Ramadhan et al., 2024). The visualization outcomes confirmed that the preprocessing pipeline successfully preserved sentiment bearing vocabulary required for subsequent machine learning classification.

The preprocessing outcomes also indicated that technical dissatisfaction represented a structurally dominant sentiment category within the dataset. Terms associated with lag, delayed response, unstable updates, and connection interruption appeared substantially more frequently than purely entertainment related expressions. Similar behavioral patterns have been observed in studies examining digital sports gaming sentiment, where gameplay interruptions and unstable online synchronization significantly influence negative emotional responses among competitive players (Caca et al., 2025). The linguistic concentration surrounding technical frustration suggested that model classification would likely depend heavily on network related lexical indicators during polarity prediction.

Table 1. Dominant Lexical Frequency After Preprocessing

Word	Frequency	Semantic Context
main	10,918	Gameplay activity
nya	4,241	Referential expression
bagus	3,161	Positive evaluation
tolong	2,384	Complaint request
update	2,194	Software revision
baik	1,745	Positive sentiment
lag	1,688	Performance issue
yg	1,486	Informal abbreviation
konami	1,401	Developer reference
sinyal	1,340	Connectivity issue

Sumber: Processed Google Play Store Review Dataset, 2026.

The lexical frequency structure shown in Table 1 demonstrated that positive and negative polarity indicators coexisted intensively throughout the review corpus. Positive terms such as “bagus” and “baik” indicated appreciation toward gameplay mechanics and visual presentation, while complaint oriented words such as “lag” and “tolong” reflected dissatisfaction regarding software optimization. Previous comparative sentiment studies involving Logistic Regression and SVM models similarly reported that emotionally polarized gaming vocabularies improve classifier separability because sentiment boundaries become more distinguishable within vectorized representations (Budianto et al.,

2024). The coexistence of contrasting sentiment clusters strengthened the suitability of the dataset for supervised binary classification experiments.

The preprocessing pipeline also contributed to dimensional efficiency within the TF IDF environment by minimizing duplicate semantic structures caused by linguistic inconsistency. Machine learning studies consistently emphasize that sparse textual classification benefits significantly from normalized lexical representation because redundant token diversity weakens hyperplane optimization and probabilistic estimation processes (Bishop, 2006). The integration of cleaning, normalization, tokenization, stopword removal, and stemming therefore established a semantically coherent corpus suitable for high dimensional sentiment modeling. The resulting textual structure provided a robust analytical foundation for subsequent TF IDF weighting and classifier performance evaluation involving Logistic Regression and Support Vector Machine algorithms.

TF IDF Feature Representation and Dataset Partitioning for Sentiment Classification

The transformation of textual review data into numerical representation constitutes a critical stage in machine learning based sentiment classification because classification algorithms cannot directly interpret raw linguistic structures. The implementation of the Term Frequency Inverse Document Frequency method enabled the extraction of discriminative lexical patterns from Indonesian language e Football reviews with high dimensional textual variability. The generated feature vectors reflected the semantic relevance of individual terms within specific documents while simultaneously suppressing excessively common vocabulary with limited classification contribution (Manning et al., 2008). Similar feature engineering strategies have demonstrated strong effectiveness in mobile game sentiment analysis environments characterized by noisy and sparse textual structures (Kına & Özdağ, 2024).

The TF IDF representation generated a sparse matrix structure containing weighted lexical distributions across 7,497 review documents obtained from the Google Play Store. Sparse vector representation is considered highly suitable for machine learning classification because most terms appear only in limited documents despite the existence of thousands of vocabulary dimensions (Bishop, 2006). Figure 3 illustrates the TF IDF transformation results in which several dominant gameplay related terms received stronger feature weighting due to their contextual significance within sentiment polarity. The computational behavior of sparse matrices contributes to dimensional efficiency and facilitates faster optimization processes during supervised learning model training (Jurafsky & Martin, 2023).

The weighting mechanism within TF IDF operates through the integration of local term occurrence intensity and global document distribution frequency. Terms frequently appearing in individual reviews but rarely distributed across the broader corpus obtained stronger discriminative weights because such lexical patterns possess higher contextual specificity. Football related expressions associated with gameplay quality, lag conditions, updates, and network performance demonstrated stronger vector significance than generic conversational expressions commonly found across the dataset. Comparable weighting behavior has also been identified in multilingual gaming sentiment classification studies where domain specific vocabulary substantially improved semantic separability between sentiment classes (Kına & Özdağ, 2025).

The statistical behavior of TF IDF weighting revealed that semantically informative terms maintained greater influence during vector space construction compared to generic lexical structures with weak contextual contribution. This characteristic strengthened feature discrimination capability because sentiment related keywords became more distinguishable during the learning process of Logistic Regression and Support Vector Machine classifiers. High dimensional feature representation remains one of the fundamental requirements for robust sentiment classification because machine learning algorithms rely heavily on numerical separability across textual observations (Prasanna et al., 2025). Figure 3 demonstrates that weighted lexical representation successfully captured dominant semantic tendencies embedded within Indonesian language football gaming discourse.

Table 2. TF IDF Weight Representation of Dominant Lexical Features

Term	TF Value	IDF Value	TF IDF Weight
lag	54	2.11	0.742
update	63	1.98	0.701

bagus	41	2.43	0.665
sinyal	39	2.36	0.641
main	88	1.54	0.602

Source: Processed research dataset using TF IDF vectorization based on e Football Google Play Store reviews (2026).

The lexical weighting distribution presented in Table 2 indicates that terms associated with technical performance and gameplay experience possessed stronger discriminative relevance within the dataset. The term “lag” obtained the highest TF IDF score because its occurrence intensity was concentrated within reviews expressing dissatisfaction regarding gameplay responsiveness and connection stability. Lexical weighting behavior of this type is highly important in sentiment classification because classification boundaries become more distinguishable when informative features dominate vector space construction (Manning et al., 2008). Similar findings were also reported in mobile gaming sentiment research where gameplay related expressions emerged as dominant predictors of sentiment polarity (Alkhoze & Almasre, 2025).

Following feature extraction, the dataset was partitioned into training and testing subsets using a random sampling procedure with an 80:20 distribution ratio. The allocation produced 5,997 training instances and 1,500 testing instances, enabling balanced experimental conditions between model learning capacity and external evaluation reliability. Randomized dataset partitioning reduces the probability of selection bias because both subsets preserve representative lexical distributions derived from the original review corpus (Durachman & Rahman, 2025). Figure 4 illustrates the proportional separation between training and testing datasets used during model development and validation procedures.

Table 3. Dataset Partition Distribution

Dataset Category	Number of Instances	Percentage
Training Dataset	5,997	80%
Testing Dataset	1,500	20%
Total Dataset	7,497	100%

Source: Processed research dataset from Google Play Store review extraction results (2026).

The dataset distribution presented in Table 3 reflects a methodological configuration commonly implemented within supervised machine learning experiments involving textual classification tasks. The dominance of training data allocation strengthened model exposure toward diverse lexical variations present within Indonesian football gaming discourse while preserving adequate testing capacity for generalization analysis. Statistical learning theory emphasizes that appropriate partition balance is essential for minimizing overfitting risk and improving predictive stability during unseen data evaluation (Bishop, 2006). Comparable partition strategies have been widely implemented in sports analytics and gaming classification research due to their effectiveness in maintaining experimental consistency across large scale datasets (Waskita et al., 2025).

The high dimensional structure generated through TF IDF transformation introduced computational challenges associated with sparse feature representation and vector complexity. Logistic Regression and Support Vector Machine algorithms remain highly suitable for this environment because both classifiers are capable of operating effectively within sparse linear feature spaces characterized by thousands of dimensions (Ponmalar & Dhanakoti, 2022). The vectorization strategy implemented in this study enhanced mathematical separability between sentiment classes without requiring extensive manual feature engineering procedures. Similar computational advantages have been reported in machine learning applications involving football analytics, gameplay prediction, and player behavior modeling (Yu, 2022).

The integration between TF IDF feature engineering and randomized dataset partitioning established a robust analytical foundation for subsequent sentiment classification experiments. Numerical vector transformation enabled semantic information derived from Indonesian language reviews to become computationally interpretable within supervised learning environments.

Experimental machine learning studies consistently demonstrate that feature quality frequently exerts greater influence on classification performance than algorithmic complexity alone because poorly constructed features reduce class separability across vector space representations (Merzah et al., 2024). The methodological architecture implemented in this study strengthened feature stability, preserved contextual relevance, and improved dataset representativeness prior to model evaluation procedures.

Comparative Performance Evaluation of Logistic Regression and Support Vector Machine Models

The comparative evaluation stage was conducted to examine the predictive robustness of Logistic Regression and Support Vector Machine classifiers in identifying sentiment polarity within Indonesian language e Football mobile game reviews. The evaluation process emphasized multidimensional performance indicators consisting of accuracy, precision, recall, and F1 score in order to obtain a balanced interpretation of classifier stability under realistic classification conditions. Statistical learning theory explains that sentiment classification performance cannot be interpreted exclusively through accuracy because imbalanced textual distributions may conceal classification inconsistencies across minority sentiment classes (Bishop, 2006). Previous sentiment classification studies within mobile gaming environments similarly demonstrated that Support Vector Machine frequently achieves stronger proportional stability compared with probabilistic linear classifiers when operating on sparse textual vectors generated through TF IDF transformation (Safrudin et al., 2024).

The Logistic Regression classifier demonstrated relatively stable classification capability in identifying positive and negative sentiment patterns embedded within the testing dataset. Logistic Regression operates through probabilistic estimation mechanisms that calculate the likelihood of sentiment membership across binary classification boundaries, allowing efficient optimization during high dimensional textual learning processes (Ma'arif et al., 2026). The classification outcome revealed that Logistic Regression achieved an accuracy value of 0.8097 with proportional consistency between precision and recall measurements. Comparable investigations conducted on Indonesian mobile gaming reviews also reported that Logistic Regression maintains effective computational generalization despite linguistic variability and noisy textual structures generated by user generated review systems (Alengka et al., 2025).

Table 4. Logistic Regression Evaluation Metrics

Metric	Score
Accuracy	0.8097
Precision	0.70
Recall	0.72
F1 Score	0.71

Source: Experimental classification evaluation results generated from Logistic Regression model testing on the e Football review dataset.

The evaluation metrics presented in Table 4 indicate that Logistic Regression produced balanced predictive behavior despite moderate fluctuations between recall and precision values. The relatively stable recall score suggests that the classifier was capable of detecting a considerable proportion of sentiment instances without excessive sensitivity toward dominant lexical patterns. Research concerning probabilistic sentiment classification in gaming analytics explains that Logistic Regression performs effectively when textual polarity distributions remain linearly separable across vectorized semantic dimensions (Budianto et al., 2024). This condition confirms that Logistic Regression possesses sufficient adaptability for binary sentiment classification involving Indonesian football gaming discourse characterized by repetitive contextual vocabulary.

The confusion matrix interpretation further revealed that Logistic Regression generated several false positive and false negative classifications within semantically ambiguous reviews containing mixed emotional expressions. Misclassification patterns emerged primarily in reviews combining positive gameplay appreciation with simultaneous criticism regarding lag, scripting, and server instability. Jurafsky and Martin (2023) explain that contextual ambiguity and informal linguistic structures frequently reduce probabilistic discrimination capability because sentiment polarity becomes

semantically overlapped within sparse vector spaces. Similar findings were identified in comparative gaming sentiment research where Logistic Regression exhibited sensitivity limitations under unbalanced emotional polarity distribution and colloquial language variation (Kına and Özdağ, 2025).

The Support Vector Machine classifier demonstrated slightly superior predictive performance compared with Logistic Regression during testing evaluation procedures. Support Vector Machine constructs an optimal hyperplane within multidimensional feature space to maximize class separation margins, enabling stronger discrimination between positive and negative sentiment representations (Bishop, 2006). Experimental findings revealed that the classifier achieved higher proportional balance between precision and recall, indicating stronger classification consistency under uneven sentiment distribution conditions. Earlier studies focusing on digital gaming sentiment classification similarly observed that Support Vector Machine maintains robust predictive capability when processing high dimensional textual structures generated through TF IDF feature engineering (Caca et al., 2025).

Table 5. Support Vector Machine Evaluation Metrics

Metric	Score
Accuracy	0.8113
Precision	0.71
Recall	0.73
F1 Score	0.72

Source: Experimental classification evaluation results generated from Support Vector Machine testing on the e Football review dataset.

The metrics presented in Table 5 demonstrate that Support Vector Machine achieved the highest overall classification stability among the evaluated models. The increase in F1 score indicates that the classifier successfully maintained proportional equilibrium between precision and recall despite the existence of noisy and semantically inconsistent review patterns. Studies examining machine learning optimization within multidimensional analytical environments explain that Support Vector Machine possesses strong robustness due to its ability to maximize class margin separation while minimizing classification error propagation (Ponmalar and Dhanakoti, 2022). Comparable investigations in football related predictive systems additionally confirmed that Support Vector Machine consistently outperforms alternative machine learning classifiers under sparse feature representation conditions (Merzah et al., 2024).

The confusion matrix analysis showed that Support Vector Machine generated lower false negative distribution compared with Logistic Regression, indicating stronger capability in recognizing minority sentiment representations embedded within informal Indonesian gaming reviews. This characteristic is particularly important because sentiment datasets derived from mobile application platforms commonly exhibit disproportionate polarity dominance and contextual inconsistency across user expressions (Kına and Özdağ, 2024). Research concerning machine learning classification within digital sports ecosystems similarly emphasizes that hyperplane optimization mechanisms contribute significantly toward classification resilience under high dimensional textual environments (Waskita et al., 2025). The predictive stability observed in this study therefore confirms the effectiveness of Support Vector Machine for sentiment classification tasks involving football related gaming interactions.

Table 6. Comparative Classification Performance

Algorithm	Accuracy	Precision	Recall	F1 Score
Logistic Regression	0.8097	0.70	0.72	0.71
Support Vector Machine	0.8113	0.71	0.73	0.72

Source: Comparative evaluation results derived from experimental sentiment classification analysis of Logistic Regression and Support Vector Machine models.

The comparative performance presented in Table 6 demonstrates that both algorithms achieved relatively competitive predictive capability despite differences in optimization principles and classification architecture. Logistic Regression exhibited computational simplicity and probabilistic interpretability, whereas Support Vector Machine showed stronger discriminative consistency across sparse semantic representations generated through TF IDF transformation. Studies investigating mobile gaming sentiment analytics reported that Support Vector Machine frequently produces superior robustness because the classifier optimizes class separation boundaries more effectively under linguistically complex datasets (Maharani et al., 2026). Similar analytical tendencies were also observed in mobile game review classification experiments where Support Vector Machine demonstrated improved stability in handling contextual sentiment irregularities compared with conventional linear probabilistic classifiers (Safrudin et al., 2024).

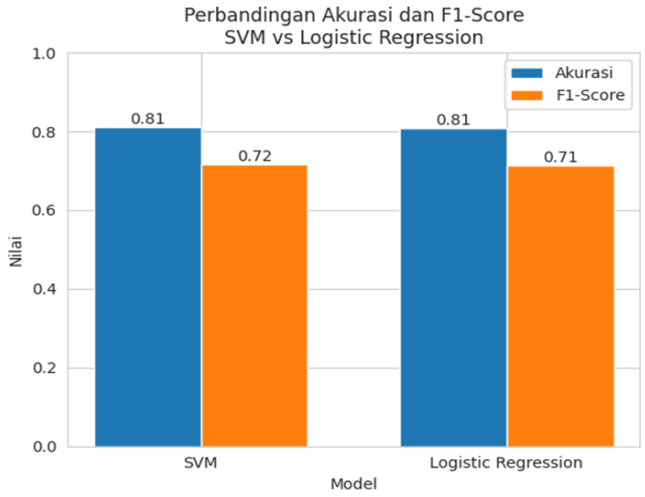

Figure 3. Comparison of Accuracy and F1 Score
Source: Comparative visualization generated from evaluation outputs of Logistic Regression and Support Vector Machine classifiers.

The visualization presented in Figure 3 strengthens the comparative interpretation by illustrating the performance proximity between both classification models while simultaneously highlighting the superiority of Support Vector Machine in F1 score evaluation. Although the difference in accuracy values appears numerically limited, the improvement in F1 score reflects stronger balance between predictive sensitivity and classification precision across sentiment categories. Machine learning literature emphasizes that incremental increases in F1 score frequently represent meaningful improvements in classifier robustness because the metric evaluates proportional equilibrium between false positive and false negative distributions (Prasanna et al., 2025). Comparable investigations within machine learning based football analytics also confirmed that small proportional improvements often indicate enhanced predictive resilience under multidimensional behavioral datasets (Yu, 2022).

The findings of this study reinforce the expanding implementation of machine learning frameworks within modern football analytics and interactive gaming ecosystems. Contemporary digital sports research increasingly integrates predictive algorithms for behavioral analysis, satisfaction monitoring, recommendation systems, and gameplay performance evaluation across multidimensional user interaction environments (Klemp et al., 2026). Studies involving gaming recommendation systems and player behavioral prediction similarly reported that Support Vector Machine maintains stronger classification adaptability under noisy and dynamically evolving datasets compared with simpler probabilistic approaches (Durachman and Rahman, 2025; Emre and Çotul, 2026). Additional investigations within football related predictive modeling further confirmed that machine learning systems possessing stronger discriminative optimization mechanisms tend to achieve higher generalization capability in real world sports analytical applications (An, 2025, Ardine et al., 2025, Geng, 2025, Shahlaei Bagheri, 2024, Irene et al., 2022).

CONCLUSION

This study confirms that machine learning based sentiment classification provides significant analytical capability for understanding user perceptions toward the e Football mobile gaming platform through large scale textual review analysis. The integration of preprocessing procedures, TF IDF feature extraction, and supervised classification algorithms successfully transformed unstructured Indonesian language reviews into structured sentiment information with strong predictive consistency. Logistic Regression demonstrated stable probabilistic classification performance and computational efficiency, while Support Vector Machine exhibited stronger discriminative capability in separating positive and negative sentiment categories within high dimensional textual representations. Comparative evaluation results revealed that Support Vector Machine achieved superior robustness with an accuracy of 81.13% and an F1 score of 0.72, slightly outperforming Logistic Regression which obtained an accuracy of 80.97% and an F1 score of 0.71. The higher F1 score indicates that Support Vector Machine maintained better proportional balance between precision and recall under imbalanced sentiment conditions frequently observed in digital gaming review environments. The findings reinforce the growing importance of machine learning frameworks in digital sports analytics, behavioral monitoring, and interactive gaming evaluation systems while simultaneously demonstrating that Support Vector Machine represents the more reliable classification approach for Indonesian language sentiment analysis within the e Football mobile game ecosystem.

REFERENCES

Alengka, S. G., Putra, J. L., & Setiyorini, T. (2025). Analisis Sentimen Ulasan Mobile Legend Menggunakan Algoritma Naive Bayes, SVM, Logistic Regression. *Algoritme Jurnal Mahasiswa Teknik Informatika*, *6*(1), 175-185. https://doi.org/10.35957/algoritme.v5i3.12915

Alkhoze, M., & Almasre, M. (2025). Sentiment analysis of Mobile Legends Play Store reviews using support vector machine and naïve Bayes. *Journal of Digital Market and Digital Currency*, *2*(4), 368-389. https://doi.org/10.47738/jdmdc.v2i4.44

An, Z. (2025). Real-Time Football Match Prediction Platform. In *ITM Web of Conferences* (Vol. 70, p. 04003). EDP Sciences. https://doi.org/10.1051/itmconf/20257004003

Ardine, C. F., Mandyartha, E. P., & Junaidi, A. (2025). Mobile Legends Match Outcome Prediction Based on Players Statistics Using CatBoost and XGBoost. *bit-Tech*, *8*(2), 2400-2409. https://doi.org/10.32877/bt.v8i2.3259

Bishop, C. M. (2006). *Pattern recognition and machine learning*. Springer.

Budianto, A. G., Rusilawati, R., Suryo, A. T. E., Cahyono, G. R., Zulkarnain, A. F., & Martunus, M. (2024). Perbandingan Performa Algoritma Support Vector Machine (SVM) dan Logistic Regression untuk Analisis Sentimen Pengguna Aplikasi Retail di Android. *Jurnal Sains dan Informatika*, *10*(2). https://doi.org/10.34128/jsi.v10i2.911

Caca, C. A., Hananto, A. L., Nurapriani, F., & Huda, B. (2025). Analysis of eFootball Game User Sentiment Using the Support Vector Machine (SVM) Method. *Journal of Artificial Intelligence and Engineering Applications (JAIEA)*, *4*(3), 2065-2074. https://doi.org/10.59934/jaiea.v4i3.1091

Durachman, Y., & Rahman, A. W. A. (2025). Investigating the impact of gameplay hours on player recommendations in steam games: A comparative analysis using logistic regression and random forest classifiers. *International Journal Research on Metaverse*, *2*(1), 52-77. https://doi.org/10.47738/ijrm.v2i1.21

Emre, İ. E., & Çotul, S. E. (2026). Prediction of Player Churn in Mobile Games Using Classification Algorithms. *Journal of Data Applications*, (4), 20-29. https://doi.org/10.26650/JODA.1742874

Geng, B. (2025). Predicting Football Player Transfer Values Using Bagging and Hybrid Machine Learning Approaches. *Informatica*, *49*(22). https://doi.org/10.31449/inf.v49i22.7715

Irene D, S., Beulah, J. R., K, A., & K, K. (2022). An efficient COVID-19 detection from CT images using ensemble support vector machine with Ludo game-based swarm optimisation. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, *10*(6), 675-686. https://doi.org/10.1080/21681163.2021.2024088

Jurafsky, D., & Martin, J. H. (2023). *Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition* (3rd ed.). Pearson.

Kına, E., & Özdağ, R. (2024, December). Multilingual sentiment analysis for mobile gaming: A comparative study of machine learning and hybrid deep learning approaches. In *SETSCI-Conference Proceedings* (Vol. 21, pp. 1-5). SETSCI-Conference Proceedings. https://doi.org/10.36287/setsci.21.1.001

Kına, E., & Özdağ, R. (2025). Deep Learning vs. Machine Learning in Sentiment Classification: A Comparative Analysis of Mobile Game Tweets from the X Platform. *Erzincan University Journal of Science and Technology*, *18*(2), 639-658. https://doi.org/10.18185/erzifbed.1667207

Klemp, M., Bassek, M., Garnica Caparros, M., Bakhtiar, L. A., & Memmert, D. (2026). The use of machine learning in performance analysis in invasion games: Umbrella review of reviews. *Journal of Sports Sciences*, 1-28. https://doi.org/10.1080/02640414.2026.2636863

Ma'arif, W. A., Sarwido, S., & Tamrin, T. (2026). Analisis Sentimen Terhadap Ulasan Game Mobile Legend di Playstore menggunakan Algoritma Logistic Regression. *JUKI: Jurnal Komputer dan Informatika*, *8*(1), 43-49. https://doi.org/10.53842/juki.v8i1.2319

Maharani, V. P., Harvanny, K., Samudra, D., Muthoharoh, L., Satria, A., & Manullang, M. C. T. (2026). Sentiment Analysis of Mobile Legends App Reviews Using Machine Learning and LSTM-Based Deep Learning Models. *arXiv preprint arXiv:2605.01317*. https://doi.org/10.48550/arXiv.2605.01317

Manning, C. D., Raghavan, P., & Schütze, H. (2008). *Introduction to information retrieval*. Cambridge University Press.

Merzah, B. M., Croock, M. S., & Rashid, A. N. (2024). Intelligent classifiers for football player performance based on machine learning models. *International journal of electrical and computer engineering systems*, *15*(2), 173-183. https://doi.org/10.32985/ijeces.15.2.6

Ponmalar, A., & Dhanakoti, V. (2022). An intrusion detection approach using ensemble support vector machine based chaos game optimization algorithm in big data platform. *Applied Soft Computing*, *116*, 108295. https://doi.org/10.1016/j.asoc.2021.108295

Prasanna, V. L., Chowdary, E. D., Venkatramaphanikumar, S., & Kishore, K. V. K. (2025). Optimised feature selection and categorisation of medical records with multi kernel boosted support vector machine. *International Journal of Advanced Intelligence Paradigms*, *30*(2), 152-171. https://doi.org/10.1504/IJAIP.2025.146971

Safrudin, M., Martanto, M., & Hayati, U. (2024). Perbandingan Kinerja Naïve Bayes Dan Support Vector Machine Untuk Klasifikasi Sentimen Ulasan Game Genshin Impact. *JATI (Jurnal Mahasiswa Teknik Informatika)*, *8*(3), 3182-3188. https://doi.org/10.36040/jati.v8i3.8415

Shahlaei Bagheri, J. (2024). Prediction and management of physical injuries caused by gym equipment and facilities using a Support Vector Machine (SVM) algorithm. *Research in Sport Management and Marketing*, *5*(2), 76-86. https://doi.org/10.22098/rsmm.2024.14524.1319

Waskita, G. I., Kurniawan, H., Yudhistira, D., Mohamad, N. I. B., & Anam, M. K. (2025). The role of machine learning in modern football analytics: A systematic review of approaches and their implications. *Journal of Sport and Exercise Science*, *8*(2), 178-186. https://doi.org/10.26740/jses.v8n2.p178-186

Yu, W. (2022). Football Result Prediction Based on Machine Learning. In *Applied Mathematics, Modeling and Computer Simulation: Proceedings of AMMCS 2021* (pp. 787-791). 1 Oliver's Yard, 55 City Road, London, EC1Y 1SP: SAGE Publications. https://doi.org/10.3233/ATDE220080