Journal of Soft Computing Exploration

Homepage: https://shmpublisher.com/index.php/joscex

JOSCEX

p-ISSN: 2746-7686
e-ISSN: 2746-0991

.

Hybrid support vector machine and random forest for environmental issue sentiment analysis on social media x

Yulius Palumpun[1]*, Moh. Rahmat Irjii Matdoan[2]
[1] Department of Information Systems, Universitas Sains dan Teknologi Jayapura, Indonesia
[2] Department of Informatics Engineering, Universitas Sains dan Teknologi Jayapura, Indonesia

Article Info

Article history:

Received May 20, 2026
Revised June 04, 2026
Accepted June 24, 2026

Keywords:

Sentiment analysis
Hybrid model
Support vector machine
Random forest
Social media
Environmental issues

ABSTRACT

This study investigates the effectiveness of a hybrid machine learning model combining Support Vector Machine (SVM) and Random Forest (RF) for sentiment analysis of environmental issues discussed on social media X. A quantitative experimental research design was employed using textual data related to environmental topics collected through an application programming interface (API)-based data extraction process. Prior to model development, the collected data underwent a series of preprocessing procedures, including text normalization, tokenization, stopword elimination, and stemming. The processed text was then converted into numerical feature vectors using the Term Frequency–Inverse Document Frequency (TF-IDF) technique. To assess classification performance, three models were implemented and compared: Support Vector Machine, Random Forest, and a hybrid SVM–RF ensemble model. Model evaluation was conducted through cross-validation using accuracy, precision, recall, and F1-score as performance indicators. The experimental results revealed that the hybrid model achieved the best overall performance, attaining an accuracy of 89.10%, compared with 85.30% for Random Forest and 82.45% for Support Vector Machine. In addition, the hybrid approach generated higher precision, recall, and F1-score values, demonstrating greater robustness and consistency in sentiment classification. These findings suggest that integrating multiple machine learning algorithms can significantly enhance the analysis of complex and unstructured social media data concerning environmental issues.

Corresponding Author:

Palumpun,
Department of Information Systems,
Universitas Sains dan Teknologi Jayapura,
Padang Bulan, Yabansai, Heram, Abepura, Kota Jayapura, Papua
Email: n2.ardelia@gmail.com
https://doi.org/10.52465/joscex.v7i2.127

1. INTRODUCTION

The development of digital technology over the last decade, accompanied by the rapid growth of social media usage, has reshaped the way people interact, share information, and communicate their views on public issues. Environmental topics, including climate change, pollution, deforestation, renewable energy, and sustainability policies, are frequently discussed across social networking platforms. Among these platforms, X (formerly Twitter) serves as an active online environment where users continuously exchange opinions and reactions related to environmental matters. The large amount of content generated through these interactions provides a valuable source of information for understanding public perspectives. Nevertheless, the unstructured nature of social media data, combined with its volume and diversity, makes conventional analysis difficult to conduct efficiently. As a result, sentiment analysis has become an increasingly important technique for identifying and interpreting emotions, opinions, and attitudes expressed in textual content [1]. When combined with machine learning methods, sentiment analysis offers opportunities to observe public responses in real time, supporting researchers and policymakers in evaluating societal reactions to environmental issues and regulations [2].

Numerous studies have investigated the application of machine learning techniques for sentiment classification. One of the most frequently adopted algorithms is Support Vector Machine (SVM), which is known for its ability to perform effectively in text classification tasks involving high-dimensional feature spaces. The foundational work of Cortes and Vapnik [3] introduced SVM as a supervised learning approach that determines optimal separating boundaries between classes. Subsequent research conducted by Joachims [4] as well as Cristianini and Shawe-Taylor [5] demonstrated that SVM performs particularly well when dealing with sparse textual representations and complex classification problems. Furthermore, Pang et al. [6] reported that SVM often achieves better classification outcomes than several conventional machine learning methods due to its strong generalization capability and ability to handle large feature sets.

In addition to SVM, Random Forest (RF) has emerged as another prominent algorithm for sentiment classification. Introduced by Breiman [7], RF operates through an ensemble strategy that combines the predictions of multiple decision trees to produce more reliable results. This mechanism contributes to improved prediction accuracy while reducing the likelihood of overfitting. Previous studies have shown that ensemble-based learning techniques are effective in processing noisy datasets and maintaining stable performance across different data conditions [8]. Within the domain of social media analysis, Kouloumpis et al. [9] demonstrated that machine learning approaches utilizing contextual and linguistic characteristics can successfully identify sentiment patterns from Twitter data. Moreover, ensemble models have shown considerable advantages when handling nonlinear relationships and diverse feature interactions that commonly occur in user-generated online content [10], [11].

Recent advances in artificial intelligence research have increasingly focused on hybrid learning approaches as a means of enhancing classification performance. Unlike single-model systems, hybrid frameworks combine the strengths of multiple algorithms in order to achieve more balanced and accurate predictions. Such integration can help reduce the limitations associated with individual classifiers while simultaneously improving robustness and predictive capability [11]. Previous findings suggest that combining different learning methods can yield better sentiment classification performance, particularly when datasets contain high levels of noise and complexity [10]. Despite these developments, much of the existing literature has concentrated on areas such as consumer reviews, marketing analytics, customer satisfaction, and political communication. Comparatively limited attention has been given to environmental discussions on social media platforms. This represents an important gap because environmental conversations frequently involve scientific concepts, policy-related debates, advocacy campaigns, and emotionally charged narratives that differ substantially from other forms of online discourse.

Another challenge concerns the characteristics of social media communication itself. Discussions on platform X often contain abbreviations, informal expressions, slang, sarcasm, and multilingual content, all of which can complicate sentiment interpretation. Pak and Paroubek [8] highlighted the value of Twitter as a rich source of sentiment-related data while also emphasizing the linguistic variability present within social media texts. Likewise, Saif et al. [12] argued that semantic understanding and contextual interpretation play essential roles in accurately identifying sentiment expressed online. These issues become even more complex when environmental discussions involve multiple languages, such as English and Indonesian. According to Abbasi et al. [13], multilingual sentiment analysis requires classification models that are capable of accommodating linguistic differences and contextual variations across languages. Consequently, more adaptive and resilient analytical approaches are required to improve sentiment detection in environmental communication on social media.

Although previous research has demonstrated the effectiveness of both SVM and Random Forest in sentiment classification tasks, studies examining the integration of these two methods within environmental sentiment analysis remain relatively limited. Existing investigations frequently rely on standalone classification

.

algorithms, leaving the potential benefits of hybrid approaches insufficiently explored. Therefore, this study introduces a Hybrid SVM-RF model designed to improve the classification of environmental sentiment on platform X. Specifically, the study aims to: (1) examine the individual performance of SVM and RF models, (2) construct a hybrid framework that combines the strengths of both algorithms, and (3) evaluate the effectiveness of the proposed model in producing more accurate and comprehensive sentiment classifications for environmental discussions on social media.

2. METHOD

This study employed a quantitative experimental research design to develop and evaluate a hybrid machine learning model integrating Support Vector Machine (SVM) and Random Forest (RF) for sentiment analysis of environmental discussions on platform X. The methodological framework was based on established principles of text mining, machine learning, and sentiment analysis, which have been widely adopted for extracting meaningful information from large volumes of unstructured textual data [1], [2]. The proposed approach was selected because it enables systematic evaluation of classification algorithms while addressing the complexity of social media content characterized by noise, multilingual expressions, abbreviations, and rapidly evolving topics.

Research Design

This study employed a comparative experimental approach to evaluate the effectiveness of three sentiment classification models: Support Vector Machine (SVM), Random Forest (RF), and a Hybrid SVM-RF model. These models were utilized to classify environmental-related discussions collected from social media platform X.

The purpose of this research design was to investigate whether integrating two established machine learning algorithms could provide superior classification performance compared with using each algorithm independently. To ensure a fair comparison, the SVM and RF models were initially developed and trained using the same preprocessed dataset.

The dataset consisted of 5,000 social media posts concerning environmental issues gathered from platform X between January and March 2025. All records were manually annotated into three sentiment categories positive, neutral, and negative through a supervised labeling process. The dataset included posts written in both English and Indonesian to represent a broader range of public discussions.

SVM was selected due to its well-documented capability in processing high-dimensional textual features and constructing optimal classification boundaries [3]. On the other hand, RF was chosen because its ensemble-based structure enhances prediction consistency, reduces overfitting, and performs effectively when handling noisy data [7]. Previous investigations have also demonstrated the suitability of SVM and RF for text classification and ensemble learning applications [4], [5], [11].

After the standalone models had been established, a Hybrid SVM-RF framework was developed by combining the predictions generated by both classifiers through an ensemble voting mechanism. This integration was designed to capitalize on the strengths of each algorithm, allowing the model to benefit from SVM's decision boundary optimization and RF's resilience to data variability. As a result, the hybrid approach was expected to deliver improved predictive accuracy, greater classification stability, and stronger overall performance.

The methodological workflow consisted of several stages, including data acquisition, text preprocessing, feature extraction using TF-IDF, model training, hybrid model development, validation through cross-validation, and comparative performance analysis. Model effectiveness was assessed using accuracy, precision, recall, and F1-score metrics [12]. Furthermore, cross-validation was applied to evaluate model generalization and minimize the risk of overfitting.

Through this systematic procedure, the study provides a structured framework for evaluating hybrid machine learning approaches in sentiment analysis involving large-scale and unstructured social media datasets.

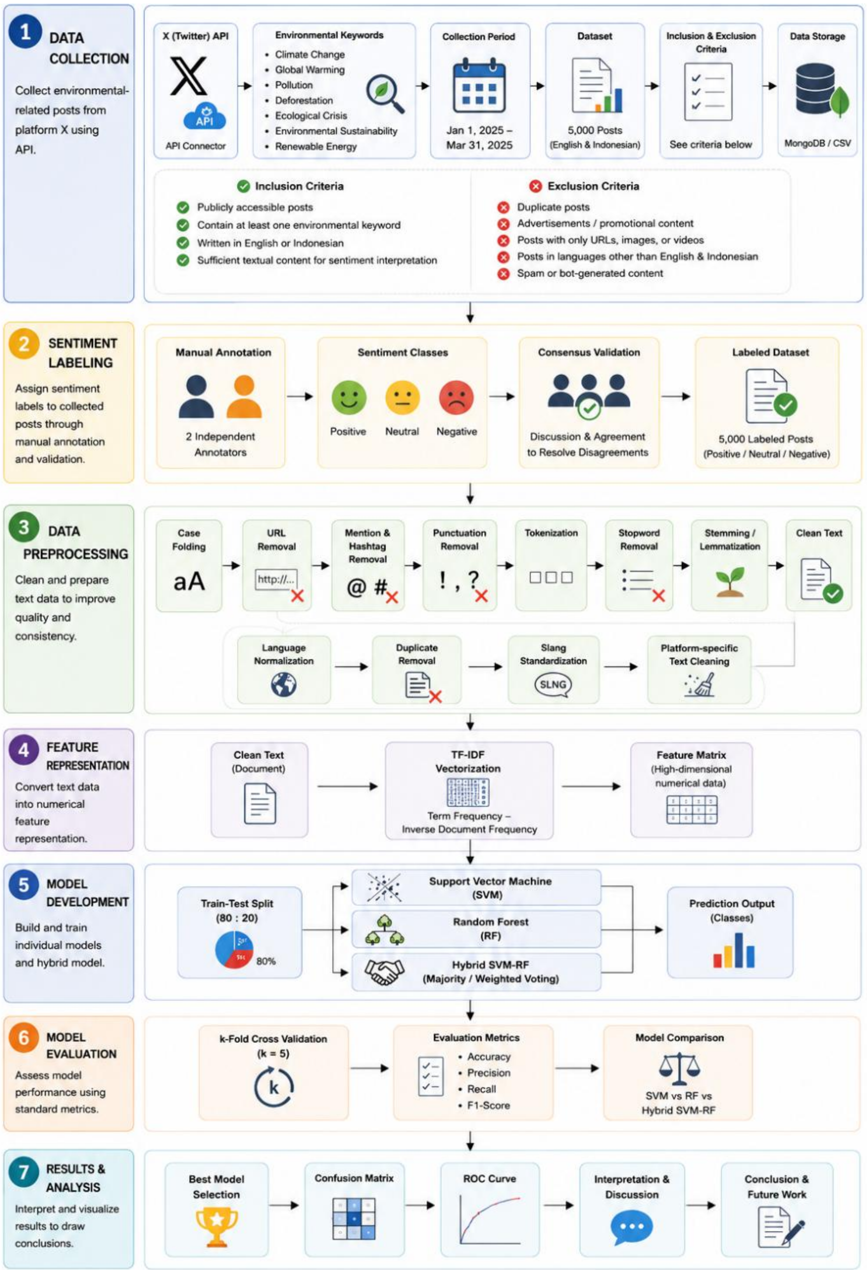

Figure 1. Research design flowchart of hybrid SVM-RF model

Figure 1 illustrates the complete workflow adopted in this research. The process began with the collection of social media data from platform X using API-based retrieval techniques and predefined environmental keywords.

A total of 5,000 posts were gathered using keywords related to environmental issues, including climate change, global warming, pollution, deforestation, ecological crisis, environmental sustainability, and renewable energy. Data collection was conducted over a three-month period, from January to March 2025, to ensure that the dataset reflected a diverse range of environmental discussions while minimizing temporal bias.

Following data acquisition, the collected posts underwent several preprocessing operations, including case normalization, tokenization, stopword elimination, stemming, and noise filtering.

The cleaned text data were subsequently transformed into numerical vectors through the TF-IDF method [2]. The resulting dataset was divided into training and testing subsets before being used to develop the SVM, RF, and Hybrid SVM-RF models. Model performance was then evaluated using cross-validation and standard classification metrics, followed by a comparative analysis to identify the most effective sentiment classification approach.

Data Collection

The dataset used in this study was obtained from platform X through a combination of API integration and web-scraping techniques. To ensure relevance to environmental topics, several keywords were used during the retrieval process, including climate change, global warming, pollution, deforestation, ecological crisis, environmental sustainability, and renewable energy.

A total of 5,000 posts were collected during the observation period from January 1 to March 31, 2025. This timeframe was selected to capture variations in environmental discussions while minimizing the influence of short-lived events and temporary online trends.

The dataset contained posts written in English and Indonesian, allowing the analysis to encompass a broader spectrum of public opinion regarding environmental issues. The inclusion of multilingual data also enabled an assessment of model performance across different linguistic contexts [13].

Additional information, such as timestamps, hashtags, engagement metrics, and contextual attributes, was also recorded to support further analysis. The chosen sampling period was intended to improve data representativeness and reduce potential temporal bias.

To maintain data quality, several inclusion and exclusion criteria were applied. Eligible posts were required to be publicly accessible, contain at least one environmental keyword, be written in English or Indonesian, and provide sufficient textual content for sentiment interpretation. Meanwhile, duplicate records, promotional content, posts containing only URLs or multimedia elements without meaningful text, non-English or non-Indonesian content, and spam-related posts were excluded from the dataset.

From an ethical perspective, only publicly available content was collected, and any information that could potentially identify users was anonymized before analysis.

Sentiment Labeling

Prior to model construction, all collected posts underwent a sentiment annotation process. Each record was assigned to one of three sentiment categories: positive, neutral, or negative.

Positive sentiment represented expressions of support, optimism, or favorable opinions toward environmental issues and sustainability initiatives. Negative sentiment reflected criticism, concern, dissatisfaction, or opposition related to environmental conditions and policies. Neutral sentiment referred to objective or informational statements that did not indicate a clear emotional tendency.

The annotation procedure was conducted manually based on predefined labeling guidelines. Two independent annotators reviewed each post and assigned sentiment labels according to contextual meaning and semantic interpretation. To ensure consistency, both annotators followed identical annotation standards throughout the process. Any disagreements were resolved through discussion until a consensus was reached.

The manually annotated dataset served as the reference standard for training and evaluating the sentiment classification models. Human-validated labels were employed to improve annotation reliability and reduce the potential inaccuracies commonly associated with fully automated labeling approaches [1], [14].

Data Preprocessing

The preprocessing stage is conducted to clean and prepare raw data for analysis. The steps include:

- Case folding: converting all text to lowercase
- Tokenization: breaking text down to individual words
- Stopword removal: eliminating commonly used words with low semantic value
- Stemming/Lemmatization: reducing words to their base forms
- Noise removal: eliminating URLs, symbols, emojis, and other irrelevant characters

According to Feldman and Sanger [2], preprocessing plays a crucial role in determining the effectiveness of text mining applications. This stage is particularly important when analyzing social media data because online content often contains informal language, abbreviations, hashtags, and various non-standard expressions.

Additional procedures were applied to accommodate multilingual content, including language normalization, duplicate elimination, slang standardization, and platform-specific text cleaning. These processes helped improve feature consistency and reduce noise within the dataset.

Feature Representation

Following preprocessing, textual information was converted into numerical representations using the Term Frequency–Inverse Document Frequency (TF-IDF) technique. This method assigns greater weights to terms that carry higher informational value while reducing the influence of frequently occurring words that contribute little to document discrimination [2], [15].

Within the environmental context, TF-IDF emphasized keywords such as sustainability, emissions, conservation, ecological degradation, and climate change. This representation enabled the classification models to focus on meaningful features associated with sentiment expression.

Model Development

This study proposes three main models:

- A classification model using SVM
- A classification model using RF
- A hybrid SVM-RF model

The SVM model classified sentiment by constructing an optimal hyperplane that separates sentiment categories [1]. In contrast, RF generated predictions through an ensemble of decision trees and combined the outputs using a majority voting strategy [7].

The hybrid model integrated predictions from both classifiers through an ensemble voting mechanism. By combining the strengths of SVM and RF, the model was expected to achieve improved robustness and classification accuracy while reducing the limitations associated with individual algorithms [11].

Model Evaluation

The effectiveness of each classification model was assessed using a confusion matrix and four widely used evaluation metrics: accuracy, precision, recall, and F1-score [16].

Accuracy measured the proportion of correctly classified observations, precision reflected the reliability of positive predictions, recall indicated the model's ability to identify relevant instances, and F1-score provided a balanced assessment of precision and recall.

To further evaluate model reliability, k-fold cross-validation was implemented. This procedure enabled performance assessment across multiple data partitions and helped reduce overfitting while improving generalization capability [16].

Research Procedure

The study follows a structured sequence of steps:

a. Data collection from platform X
b. Data preprocessing and cleaning
c. Feature extraction using TF-IDF
d. Training standalone SVM and RF models
e. Developing the Hybrid SVM-RF model
f. Performance assessment and comparative analysis across models
g. Result analysis and interpretation

Throughout these stages, optimization and weighting strategies are applied to improve predictive performance. This structured workflow ensures methodological rigor while enabling flexible analysis of environmental sentiment in social media contexts.

3. RESULTS AND DISCUSSIONS

This section presents both quantitative analysis and qualitative interpretation used to address the research question, namely whether a hybrid SVM-RF model is capable of supporting sentiment analysis on platform X in the context of environmental issues. The structure of this section begins with data distribution analysis, followed by model evaluation results, and concludes with a discussion of the key findings obtained.

Sentiment Data Distribution

Prior to the modeling stage, an exploratory data analysis was conducted to examine the basic characteristics and distribution patterns of the dataset. As expected from a sequentially collected Twitter dataset, the results revealed a noticeable imbalance among sentiment classes.

Table 1. Sentiment data distribution

Sentiment Class	Number of Data	Percentage
Positive	2,450	49%
Neutral	1,850	37%
Negative	700	14%

The data presented in Table 1 indicate that positive sentiment dominates the dataset. This suggests that users tend to express more optimistic viewpoints when discussing environmental topics. On the other hand, the negative sentiment class appears significantly underrepresented compared to the others. Although smaller in proportion, this category can be more linguistically complex, as it may include indirect criticism, sarcasm, or implicit negative expressions. Such characteristics can introduce ambiguity and potentially lead to misclassification during the modeling process.

Model Performance Evaluation

To assess the effectiveness of the proposed approach, the performance of SVM, RF, and Hybrid SVM-RF models was compared using accuracy, precision, recall, and F1-score metrics [16].

Table 2. Performance comparison of sentiment classification models

Model	Accuracy	Precision	Recall	F1-Score
SVM	82,45%	80,12%	78,90%	79,50%
Random Forest	85,30%	83,75%	82,40%	83,07%
Hybrid SVM-RF	89,10%	87,85%	86,90%	87,37%

As shown in Table 2, a large performance gap can be observed between the single model and hybrid models for sentiment classification. The SVM model was able to correctly classify 82.45% of the samples, with a precision is 80.12%, recall is 78.90%, and F1-score is 79.50%. On the other hand, Random Forest model performed better than SVM and could achieve an accuracy of 85.30%, precision is 83.75%, recall is 82.40% and F1-Score is 83.07%.

Of all models, the Hybrid SVM-RF model performed best across all metrics. The accuracy is at least 3-7% better than both of the single models, suggesting that this combination to pair algorithms allows for a more ideal model in predicting sentiment data.

Additionally, in the hybrid model precision and recall values increase leading to this model more accurate rather, presenting even with lowered classification error both sides false positive and negative for each of sentiment. Finally, these results confirm that the hybrid method has a true improvement to increase predicative power compared with single algorithms.

Hybrid Model Workflow

In Figure 2 Workflow of hybrid model used in this study to assess the mechanism The diagram explains the full process stages — from data input to final prediction generation.

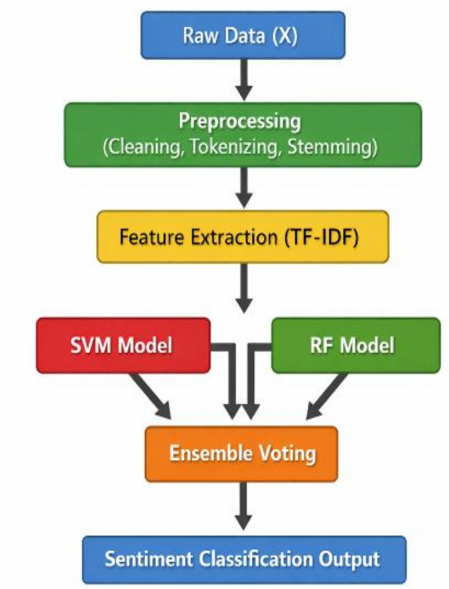

Figure 2. Hybrid SVM-RF model workflow

The proposed framework for hybrid-model-driven sentiment classification is descriptively defined in Figure1. At the level of features, first is raw data (X) which are means unprocessed data, then goes to preprocessing stage includes text cleaning, text tokenizing and stemming. Here is a stage without noise since the data will be prepared for further processing.

The feature extraction uses a technique called TF-IDF, which is used for processing of cleaned data in this part where only the numerical representation of text is built; here we can call them features, so that machine learning algorithms can learn from these features.

Information from during grain "feature perspective" which is obtained using three dynamic gating networks in a hybrid model, and two matrix multiplications where features are extracted constitute branched layers. In core stage we put these exploded features to the Support Vector Machine (SVM) or Random Forest (RF). Each of the predictions come from separate models and possess their own traits.

Voting in ensemble decision making based on majority or average of predictions from both models. The idea behind this is to make use of the benefits of all the algorithms and yield better predictions. Finally, a sentiment classification output that is literally the prediction result of an entire procedure. In conclusion, this workflow demonstrates how to leverage the strengths of two different models into a single model through a hybrid approach to push beyond baseline metrics.

Confusion Matrix Analysis

Table 3 illustrates the results confusion matrix, where we provide a detailed overview of how well our model is performing with respect to classifying each sentiment class in terms of actual vs predicted values.

Table 3. Hybrid model confusion matrix

Actual / Predicted	Positive	Neutral	Negative
Positive	2300	120	30
Neutral	140	1650	60
Negative	50	90	560

From Table 3, the performance of different classifiers for the three classes (positive, neural and negative) used to classify sentiment are provided.

In the case of predicting positive class, as model was able to classify 2300 data points correctement. On the other hand, 120 positive data points were incorrectly classified as neutral and 30 were negated. This means that the model is very good at identifying positive sentiment with a low error rate. For neutral, the model made 1,650 correct predictions. 140 neutral data points predicted as Positive, and 60 Neutral predicted into Negative. That means as the neutral class is more ambiguous than positive, it has a higher in-level confusion for our model.

On the other hand, for negative class it correctly classified 560 data points. 50 false positives the negative data point predicted as positive, and 90 neutral errors. This looks good on the performance for negative class but still is a little tricky when it comes to disambiguating between neutral and negative sentiment classes. Overall, the model was best for the positive class followed by neutral and relatively poor on negative. The top one is the class with no characteristics, then misclassification occurrences between neutral and positive classes.

Model Performance Visualization

To complement the numerical evaluation results, a graphical visualization was employed to provide a clearer comparison of the performance achieved by the Support Vector Machine (SVM), Random Forest (RF), and Hybrid SVM-RF models. This visualization facilitates a more intuitive interpretation of classification outcomes and helps illustrate the differences in predictive performance among the evaluated approaches.

.

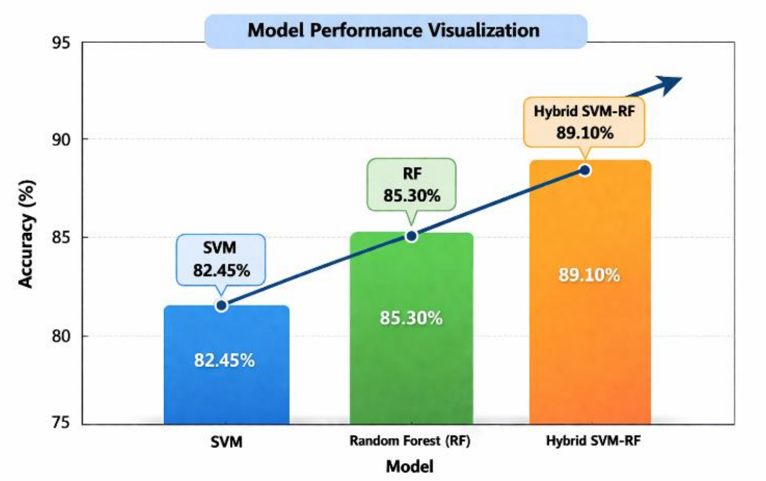

Figure 3. Model accuracy comparison graph

As shown in Figure 3, the classification performance exhibits a gradual improvement from SVM to RF and subsequently to the Hybrid SVM-RF model. Although SVM produced the lowest accuracy among the three models, it still served as a reliable baseline classifier for sentiment analysis. Previous studies have demonstrated that SVM performs effectively in high-dimensional text classification tasks because of its ability to construct optimal decision boundaries and maintain strong generalization capabilities [4], [5]. However, the presence of noisy, heterogeneous, and context-dependent social media data may reduce its classification effectiveness when applied independently.

The RF model achieved higher accuracy than SVM. This improvement can be attributed to its ensemble-learning architecture, which combines multiple decision trees to capture diverse patterns within the data and reduce the risk of overfitting [7]. Ensemble-based approaches have also been reported to enhance classification robustness and stability, particularly when dealing with complex and noisy datasets [11].

Among all evaluated models, the Hybrid SVM-RF framework achieved the highest classification accuracy. The observed improvement suggests that combining the strengths of SVM and RF can enhance predictive performance. The hybrid model benefits from the capability of SVM to identify optimal classification boundaries [4], [5], while simultaneously leveraging the robustness of RF in handling nonlinear relationships and heterogeneous data distributions [7], [11].

Furthermore, the graphical results are consistent with the evaluation metrics presented in Table 2. In addition to achieving the highest accuracy, the Hybrid SVM-RF model also recorded superior precision, recall, and F1-score values. These findings are consistent with previous studies indicating that ensemble and hybrid learning approaches can improve classification performance by integrating complementary learning mechanisms while reducing the limitations of individual classifiers [10], [11].

Impact of Preprocessing and Feature Representation

Results show that preprocessing stage contributes the most significant impact to achieve better quality of data before performing classification based on Support Vector Machine and Random Forest methods. Data retrieved from platform X are intrinsically unstructured and noisy, filled with abbreviations or acronyms, emojis (directly denoted by its unicode), URLs caused a human-readable text synthesis like "has anyone tried the new https:/example.com/" in different variations of informal language. These effects play into model accuracy, especially if uncleaned data is used.

Particularly text preprocessing functions to help standardize different forms of written words, such as converting all the letters into lowercase and getting rid of mess across variations in writing. This step is crucial to avoid the duplication of features because of formatting disrepancies. Then, the tokenization divides the text into basic word units to prepare it for analysis. Tokenize appropriately, without this the model will fail to easily pick up relevant linguistic correlations.

Removal of stopword is also an important step in enhancing the data representation. Some common words such as “and”, “the” or ”is” occur repeatedly in a text but are not really discriminative for sentiment classification. These words can, in theory and practice, be ignored so by removing them the model works much better focusing on actual terms. Furthermore, stemming or lemmatization can reduce words into their root form avoiding to increase feature dimensions due to affine spaces of some words.

After preprocessing, feature representation needs to be considered next. This study applies the TF-IDF (Term Frequency–Inverse Document Frequency) method which is well known to be effective in text mining settings. This technique gives larger importance to words that frequently appear in a particular document but less often occur across the entire corpus. Thus, TF-IDF emphasizes only the important terms that really help in differentiation of sentiment.

For example, in the context of environmental issues words like “climate”, “pollution” and “sustainability” get more weight due to high co-occurrences on related topics. In addition, these terms are, not only topic indicators but also points of sentiment polarity. For example, the word pollution has a negative sentiment and sustainability occurs in positive contexts more. That is, good feature representation helps in topic detection but also facilitates the accurate probability of sentiment classification.

Further, with feature representation using TF-IDF, can fit the model effectively. Faster training and overfitting avoidance by limiting feature dimension. This becomes especially important in regards to the models being deployed at extremely scale and real-time.

Discussion and Interpretation

The findings indicate that the Hybrid SVM-RF approach delivered the strongest performance among the three classification models evaluated in this study. As reported in Table 2, the hybrid model attained an accuracy of 89.10%, exceeding the results obtained by Random Forest (85.30%) and SVM (82.45%). A similar trend can be observed in precision, recall, and F1-score values, suggesting that the superiority of the hybrid model was consistently reflected across multiple performance indicators rather than being confined to a single metric.

These results imply that integrating SVM and RF allows the classification system to benefit from complementary learning mechanisms. Instead of depending on a single predictive strategy, the ensemble model combines the strengths of both algorithms in identifying sentiment patterns contained within environmental discussions. The observed increase in accuracy, amounting to 6.65% relative to SVM and 3.80% relative to RF, demonstrates that the hybrid framework contributed meaningfully to improving sentiment classification outcomes in multilingual social media data.

Further evidence can be obtained from the confusion matrix presented in Table 3. The positive sentiment category achieved the highest classification accuracy, with 2,300 correctly identified instances from a total of 2,450 positive posts. This outcome corresponds with the sentiment distribution shown in Table 1, where positive sentiment constituted the largest share of the dataset at 49%. The greater availability of positive training examples may have enabled the model to learn more representative linguistic patterns associated with favorable environmental opinions.

The neutral category also produced satisfactory results, recording 1,650 correctly classified observations. Nevertheless, a noticeable degree of overlap remained between positive and neutral sentiments. Specifically, 140 neutral posts were predicted as positive, while 120 positive posts were classified as neutral. This tendency suggests that environmental conversations often contain factual information accompanied by subtle evaluative language, making the distinction between these two sentiment categories less explicit.

Among the three sentiment classes, negative sentiment proved to be the most difficult to identify accurately. Although 560 negative posts were classified correctly, the number of classification errors remained relatively higher than those observed in the positive category. A substantial portion of the errors involved negative posts being assigned to the neutral class, with 90 such cases recorded in the confusion matrix. One possible explanation for this result is the limited representation of negative sentiment in the dataset. As shown in Table 1, negative posts accounted for only 14% of the total observations, providing fewer examples from which the model could learn distinctive negative sentiment patterns.

In addition to class imbalance, the linguistic characteristics of negative environmental discourse may also contribute to classification challenges. Expressions of dissatisfaction or criticism are often conveyed indirectly through nuanced wording, contextual implications, uncertainty, or implicit concerns rather than through clearly negative terms. Such language structures can obscure sentiment polarity and increase the likelihood of confusion between negative and neutral categories. Therefore, the misclassification observed in this study appears to stem not only from unequal class distribution but also from the inherent complexity of environmental communication on social media platforms.

The evaluation results further indicate a relatively stable performance of the hybrid model. Improvements were observed simultaneously across accuracy, precision, recall, and F1-score, suggesting that the enhancement was balanced rather than concentrated on a specific metric or sentiment class. In addition, the implementation of k-fold cross-validation reduced dependence on a single train-test partition and provided a broader assessment of model performance across different data subsets. The consistency of these results supports the reliability of the proposed approach when applied to unseen data.

.

Consequently, statements regarding robustness and resistance to overfitting should be interpreted in light of the empirical evidence generated during model evaluation. The combination of higher classification scores and stable validation results indicates that the hybrid model achieved stronger generalization capability than either standalone classifier within the environmental sentiment dataset examined in this study.

Overall, the results suggest that the integration of SVM and RF offers practical advantages for sentiment analysis involving multilingual environmental discussions on platform X. Beyond improving overall predictive accuracy, the hybrid framework reduced classification errors across sentiment categories and produced more balanced performance metrics. Nevertheless, the classification of minority negative sentiment remains a challenge. Future studies may address this limitation by incorporating resampling techniques, class-balancing strategies, more sophisticated feature representations, or deep learning architectures specifically designed to handle imbalanced sentiment datasets.

Comparison with Previous Studies

The results obtained in this study align with a growing body of literature suggesting that ensemble learning methods frequently outperform individual machine learning algorithms in sentiment classification applications. Although both SVM and Random Forest achieved satisfactory results when applied independently, the Hybrid SVM-RF model consistently produced the strongest outcomes across all evaluation measures. The model recorded an accuracy of 89.10% and an F1-score of 87.37%, accompanied by higher precision and recall values. These findings indicate that integrating multiple classifiers can enhance predictive capability when processing noisy, heterogeneous, and unstructured social media content.

In comparison with many previous studies that employed a single classification technique, the hybrid framework proposed in this research demonstrated a greater capacity to identify varied sentiment characteristics within environmental conversations. The improvement observed across all performance indicators suggests that the combination of SVM and RF enabled the model to utilize complementary classification strengths, resulting in more reliable sentiment predictions.

A further contribution of this study is associated with the research context itself. Whereas much of the existing sentiment analysis literature concentrates on product reviews, customer feedback, or commercial services, this investigation focused on environmental discussions occurring on platform X. Such discussions are often composed of scientific terminology, policy-related narratives, social concerns, and personal viewpoints that appear simultaneously within the same conversation. As a result, determining sentiment polarity becomes more challenging because opinions are not always expressed in a direct or explicit manner.

This complexity is also reflected in the confusion matrix results. While the hybrid model showed strong performance in recognizing positive and neutral sentiments, the negative category remained comparatively more difficult to classify. The presence of indirect criticism, expressions of concern, and uncertainty within environmental discourse may contribute to overlaps between negative and neutral sentiment categories. Similar patterns have been documented in studies examining socially sensitive or issue-driven discussions, where sentiment boundaries tend to be less distinct than those found in conventional review datasets.

Taken together, the findings demonstrate that the Hybrid SVM-RF model provides a suitable solution for sentiment analysis involving multilingual environmental conversations on social media. Its ability to achieve higher classification performance while accommodating diverse linguistic expressions highlights its applicability to domains characterized by complex, context-dependent, and highly dynamic communication patterns.

Research Implications and Contributions

The findings of this study provide several practical implications for real-world applications. The proposed model can be utilized to monitor shifts in public opinion regarding environmental issues over time, and when integrated with a real-time processing system, it is capable of identifying sentiment trends dynamically. More importantly, the insights generated by the model may support evidence-based policy formulation by offering a more accurate representation of public perceptions.

In addition, the system contributes to the development of advanced social media analytics tools that are more adaptive and responsive to user needs. With these capabilities, the hybrid framework can serve as a useful instrument for governments, environmental organizations, and researchers in examining the relationship between public discourse and environmental concerns in a more efficient manner.

From a methodological perspective, this research expands the application of machine learning by investigating hybrid strategies for sentiment classification. The results indicate that combining multiple algorithms yields better performance compared to using individual models separately. Furthermore, this work opens opportunities for future research, particularly in integrating hybrid machine learning approaches with deep learning techniques to enhance semantic understanding and contextual analysis.

Limitations and Model Robustness

Although the proposed approach demonstrates strong performance, several limitations should be acknowledged. One key challenge lies in the nature of social media data itself, which is often noisy, informal, and ambiguous. Elements such as slang usage, inconsistent grammar, and emojis can introduce uncertainty and potentially affect classification accuracy.

Nevertheless, the hybrid model proposed in this study shows greater stability compared to standalone models. This is reflected in the relatively consistent F1-score, suggesting a balanced performance between precision and recall. Even so, the model's effectiveness remains influenced by the quality and characteristics of the dataset, which can act both as a strength and a limitation depending on the data condition.

Final Synthesis

Overall, the results confirm that the hybrid combination of Support Vector Machine (SVM) and Random Forest (RF) is an effective approach for sentiment analysis of environmental discussions on platform X. This integration produces a model that is more accurate, stable, and adaptable in handling the complexity of social media text data.

Both empirical results and evaluation metrics support the practical and theoretical relevance of this approach. On the practical side, the model has potential for direct application in public opinion monitoring systems. Theoretically, this study contributes to the growing body of literature on hybrid machine learning methods for sentiment analysis and provides a foundation for future research in more advanced analytical frameworks.

4. CONCLUSION

This research developed a hybrid sentiment classification model that combines Support Vector Machine (SVM) and Random Forest (RF) to analyze environmental discussions on platform X. The findings reveal that the proposed Hybrid SVM-RF approach consistently achieved better classification performance than the individual SVM and RF models. Improvements were observed across all evaluation measures, including accuracy, precision, recall, and F1-score. These results suggest that integrating the strengths of different machine learning algorithms can provide a more effective solution for processing complex and highly variable social media data.

The study also demonstrates the importance of the preprocessing stage and TF-IDF feature extraction in improving classification quality. By reducing noise and transforming textual information into meaningful numerical representations, the proposed framework was able to produce more accurate and consistent sentiment predictions. Furthermore, the hybrid model showed a stronger ability to manage multilingual environmental discussions that often contain diverse linguistic structures, informal expressions, and contextual variations.

From an application perspective, the developed model may be useful for monitoring public attitudes toward environmental issues, evaluating community responses to environmental policies, and supporting social media–based decision-making processes. The insights generated from sentiment analysis can help policymakers, environmental institutions, and researchers gain a deeper understanding of public concerns and perceptions regarding sustainability-related topics. In addition, this study enriches existing research on sentiment analysis by providing empirical evidence of the advantages offered by hybrid machine learning techniques.

Nevertheless, several limitations should be considered when interpreting the findings. The dataset was collected from a single social media platform within a specific observation period, which may restrict the broader applicability of the results. Future studies are encouraged to utilize larger and more diverse datasets, incorporate data from multiple social media platforms, and investigate the integration of hybrid machine learning frameworks with deep learning approaches. Such developments may further enhance the model's ability to capture semantic meaning, contextual relationships, and evolving patterns of online discourse.

REFERENCES

[1] C. Liu and others, "Peroxymonosulfate Activation through 2D/2D Z-Scheme CoAl-LDH/BiOBr Photocatalyst under Visible Light for Ciprofloxacin Degradation," *J. Hazard. Mater.*, vol. 420, p. 126613, 2021.

[2] R. Feldman and J. Sanger, *The Text Mining Handbook: Advanced Approaches in Analyzing Unstructured Data*. Cambridge, UK: Cambridge University Press, 2007.

[3] C. Cortes and V. Vapnik, "Support-Vector Networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, Sep. 1995.

[4] T. Joachims, "Text Categorization with Support Vector Machines: Learning with Many Relevant Features," in *Machine Learning: ECML-98*, 1998, vol. 1398, pp. 137–142.

.

[5] N. Cristianini and J. Shawe-Taylor, *An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods*. Cambridge, UK: Cambridge University Press, 2000.

[6] B. Pang, L. Lee, and S. Vaithyanathan, “Thumbs Up? Sentiment Classification Using Machine Learning Techniques,” in *Proceedings of the 2002 Conference on Empirical Methods in Natural Language Processing (EMNLP 2002)*, 2002, pp. 79–86.

[7] L. Breiman, “Random forests,” *Mach. Learn.*, vol. 45, pp. 5–32, 2001.

[8] A. Pak and P. Paroubek, “Twitter as a Corpus for Sentiment Analysis and Opinion Mining,” in *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC 2010)*, 2010, pp. 1320–1326.

[9] E. Kouloumpis, T. Wilson, and J. Moore, “Twitter Sentiment Analysis: The Good the Bad and the OMG!,” in *Proceedings of the International AAAI Conference on Web and Social Media*, 2011, vol. 5, no. 1, pp. 538–541.

[10] N. Ofek, S. Poria, L. Rokach, E. Cambria, A. Hussain, and A. Shabtai, “Unsupervised Commonsense Knowledge Enrichment for Domain-Specific Sentiment Analysis,” *Cognit. Comput.*, vol. 8, no. 3, pp. 467–477, Jun. 2016.

[11] L. Rokach, “Ensemble-Based Classifiers,” *Artif. Intell. Rev.*, vol. 33, no. 1--2, pp. 1–39, Feb. 2010.

[12] H. Saif, Y. He, and H. Alani, “Semantic Sentiment Analysis of Twitter,” in *The Semantic Web -- ISWC 2012*, 2012, vol. 7649, pp. 508–524.

[13] A. Abbasi, H. Chen, and A. Salem, “Sentiment Analysis in Multiple Languages: Feature Selection for Opinion Classification in Web Forums,” *ACM Trans. Inf. Syst.*, vol. 26, no. 3, pp. 1–34, Jun. 2008.

[14] A. Go, R. Bhayani, and L. Huang, “Twitter Sentiment Classification Using Distant Supervision,” 2009.

[15] M. Hu and B. Liu, “Mining and Summarizing Customer Reviews,” in *Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2004, pp. 168–177.

[16] M. Sokolova and G. Lapalme, “A Systematic Analysis of Performance Measures for Classification Tasks,” *Inf. Process. Manag.*, vol. 45, no. 4, pp. 427–437, 2009.