Anusandhanvallari
Vol 2024, No.1
September 2024
ISSN 2229-3388

Enhanced Social Media Opinion Mining Using Feature Optimization and Ensemble Machine Learning

Tenkale Pallavi S[1]., Dr. S. Jagannatha[2]

[1]Research Scholar, Department of Computer Applications, M S Ramaiah Institute of Technology, Affiliated to Visvesvaraya Technological University,Belagavi 560 054

[2]Associate Professor, Department of ComputerApplications, M S Ramaiah Institute of Technology, Affiliated to Visvesvaraya Technologial University, Belagavi 560 054

Abstract: With the increasing amount of user generated content (UGC) on social media, accurate and efficient opinion mining has become more and more crucial. In this research, an improved framework for sentiment classification with feature optimization and ensemble machine learning technique is explored. Natural language processing (NLP) methods are used to process textual data and represent data as Term Frequency – Inverse Document Frequency (TF-IDF). Redundant attributes were thought to be eliminated by feature-selection methods such as Information Gain, Chi-square or optimization using the genetic algorithm. Individual performance of Support Vector Machine, Random Forest, Logistic Regression and Naive Bavies are examined as well as their performance in ensemble. Evaluation models include accuracy, precision, recall, F1-score and hypothesis testing. The suggested framework strives to enhance the reliability of sentiment classification, construction of lower dimensionality feature space, and the robustness of social media sentiment classification for practical use.

Keywords: Opinion Mining, Social Media, Feature Optimization, Ensemble Machine Learning, Sentiment Analysis, TF-IDF, Feature Selection, Classification Accuracy.

1. Introduction

Social media platforms have exploded in number, and consequently have changed the way people post thoughts, share information, and describe their own experiences about products, services, public policies, and social issues. Social media like X, Facebook, Instagram, and online forums can serve as horrific stores of consumer-generated content that offers insight into trends, consumer attitudes and other social media. The content of social media is greatly context-dependent and unstructured, however, making it difficult to systematically interpret. This problem is met by Opinion Mining (sentiment analysis), which uses Natural Language Processing (NLP), Text Mining, and machine learning to recognize and categorize opinions following the sentiment orientation they contain. Pang and Lee (2008) provide an important foundation for the present research, by reviewing the leading computational techniques and problems of opinion-oriented information retrieval. After this, Ravi and Ravi, (2015), focused their work on both major tasks, methodologies and applications of opinion mining and brought into light the increasing relevance of computational approach that entices meaningful information from textual data collected online.

Although there have been substantial improvements in automated sentiment classification in recent years, traditional machine learning methods still face a number of problems in the analysis of social media text. The sentiment polarity is difficult to identify for: Informal expressions, Spelling variations, Abbreviations, Emoji, Sarcasm, Negation (yes / no), Changing vocabulary. In addition, high-dimensional textual representations, in particular the ones derived using bag-of-words or term frequency-inverse document frequency (TF-IDF) methods, could carry a high number of keyword that are not relevant and redundant or only informatively weak. These features can make the computations more complicated, make the models less interpretable, and negatively impact classification accuracy. As mentioned in the introduction by Xia et al (2011), the importance of ensemble

techniques for sentiment classification was shown by combining various sets of features and classifiers. Likewise, ensemble learning research has identified the benefits of working with a collection of complementary classifiers that may yield more predictable results than using any individual learning algorithm.

Optimization of features provides a systematic way of solving these problems by selecting features, which is most representative of the actual data and throwing out superfluous or redundant features. There are numerous methods for reducing the dimensionality of features, including information gain, chi-squares, correlation based feature selection, genetic algorithm features and wrapper based optimization, that can be used to help optimize calls to subsequent classifiers. Jain and Jain (2022) explored hybrid feature selection and hybrid-cognitive ensemble classification in which statistical feature-selection techniques were combined with optimisation by genetic algorithm. Their work confirms the impact of the need for integration of feature reduction and classifier selection, instead of considering the 2 processes separately. Likewise, Yenkikar et al. (2022) introduced a semantic relational machine learning model based on cascaded feature selection and heterogeneous classifier ensemble to perform sentiment analysis in a few Twitter data sets.

Taking this background, the present study is entitled Enhanced Social Media Opinion Mining Using Feature Optimization and Ensemble Machine Learning, which aims to create an integrated framework for the text-based opinion mining problem by optimizing the textual features for text representation and performing ensemble-based classification task of sentiment. The philosophy is that features carefully selected and processed with complementary machine learning algorithms can enhance the discrimination of the positive, negative and neutral opinions and manage computation costs. The study also focuses on the performance evaluation that is systematic based on accuracy, precision, recall, F1 score, and confusion matrix analysis. The research aims to advance the field of efficient, robust and generalizable opinion-mining systems for social media analytics by bringing feature optimization and ensemble learning into a common framework.

2. Literature Review

Sentiment analysis studies have gradually progressed from traditional text classification to the integrated approach which combines feature engineering, optimization, and ensemble learning. Pang and Lee (2008) conducted a thorough survey of the methods in opinion mining and paid particular attention to the challenges of recognizing individuals' opinion expressions, which are subjective, and correctly identifying the sentiment polarity. Their work laid the conceptual groundwork for later research on sentiment classification using machine-learning. Building on this treatment, Ravi and Ravi (2015) analyzed sentiment analysis tasks, computational methods, and real-world examples, and found that current methodological problems include the representation of text, interpretation of the given context, and the effectiveness of classification.

Xia et al., (2011) discuss their experiments on combining alternative feature representations and learning algorithms. They used part-of-speech and word-relation based features, comparing their performance with the Naive Bayes, maximum entropy and support vector machine classifiers. The investigation showed that the choice and combination of both the features and classifiers contribute to the ensemble performance. An Ensemble Classification System for Twitter Sentiment Analysis (2018) in parallel explored the use of traditional classifiers alongside and found better performance than using individual classifiers or majority voting. The results of this study indicate that classifiers used in ensemble construction should be capable of providing certain predictors that enhance the performance of the classifiers in the ensemble.

Another significant line of research regarding sentiment analysis is related to feature optimization. To maximize the efficiency in social media sentiment classification, Sharma and Jain (2020) studied a hybrid ensemble learning approach, which employs information gain and chi-square feature selection methods and an ensemble of AdaBoost, support vector machine and logistic regression. Jain and Jain (2022) further conducted a study on a hybrid approach that integrates information gain, chi square, gini index as well as genetic algorithm

optimization for feature selection. Combined, these studies highlight the possibility of the benefits of dimensionally reducing the dimension of a document set while simultaneously retaining sentiment-relevant information, which could further enhance the capabilities of machine learning classifiers.

Ghosh and Sanyal (2018) compared the performance of three feature selection algorithms including information gain, chi-square and gini-index on three feature representations, unigram, bigram and combined features. They concluded that the use of composite unigram–bigram features would increase the classification efficiency and in the experiments they carried out, information gain in combination with support vector machines achieved comparable results. This highlights the importance of testing the different feature-selection mechanisms in practice, as they may perform differently on various datasets and classifiers.

Later work has focused not only on advanced learning architectures, but also on some of the methods of optimisation. Yenkikar, B., et al. (2022) developed three improvements to the model – a cascaded feature selection, word embeddings, and sentiment lexicon – in combination with a heterogeneous classifier ensemble. They showed that it is possible to select informative features and at the same time use complementary classifiers. Ensemble-based modelling has been extended by Kora and Mohammed (2023) who study the relationships between learning stages of multiple models using a meta-ensemble deep learning approach for sentiment analysis. These studies suggest that the effectiveness of an improvement in performance is not just boosted because of an increase in model complexity, but also due to the synthesis of appropriate combinations of representations, feature selection schemes and learning algorithms.

Sentiment classication in other social media settings has also been very much researched. Habernal et al. (2015) analyzed supervised sentiment detection for Czech social media, testing the effectiveness of preprocessing techniques, feature selection methods and classification algorithms. They create an illustration of feature engineering according to the linguistic properties of the target corpus. Reading and comparing sentiment-analysis web services, Serrano-Guerrero et al. (2015) highlighted the value of assessing the capabilities and limitations of different sentiment-analysis systems. Sun et al. (2015) offered a sentiment aware social media recommendation framework employing ensemble learning for categorizing the affective text, thus demonstrating the usefulness of sentiment classification not just for polarity detection.

A Tweet Sentiment Classification Approach Using an Ensemble Classifier (2024), a recent paper authored by a group of researchers, used a random forest (RF), decision tree (DT), and support vector machine (SVM) classifiers, dimensionality reduction-based strategies, and wrapper based feature selection methods. The authors claimed 93.42% accuracy for their proposed method in their dataset. While this is an indication of the potential of using feature optimization in combination with ensemble learning, the performance for classification should be also evaluated in the context of the dataset composition, class distribution, method validation, and baseline selection.

Literature in general shows that feature optimization and ensemble classification are viable complementary approaches to opinion mining on the social media. The effectiveness of their methods is sensitive to feature representation, feature selection, type of classifiers, as well as the properties of datasets. The much need to improve the sentiment classification from these observations, the present study examines an integrated framework to provide optimized textual features to various machine learning classifiers and integrated the predictions of various classifiers, which further enhances the sentiment classification. The importance of its contribution is that it evaluates quantitatively the performance gains and losses of optimizing individual features versus optimizing the ensemble of classifiers and features, in the light of a common experimental design.

Aim of the Study

The aim of the study is to develop and evaluate a framework of social media opinion-mining system using feature optimization method with ensemble machine learning for achieving higher accuracy of sentiment classification, lesser feature dimensionality and for better reliability and efficiency to find out the positive, negative and neutral sentiments in social media text.

Objectives of the Study

1. To extract and select the most relevant text features with most informative information from social media information for effective classification of sentiment.

2. To build an ensemble machine learning model to make more accurate and robust social media opinion mining.

3. To assess the effectiveness of the suggested framework based on accuracy, precision, recall, F1-score and comparison with various machine learning classifiers and models without any optimization done on the features.

Methodology

The study uses the experimental and quantitative approach for building and testing an improved social media opinion-mining architecture that combines feature optimization processes and ensemble machine learning. The text data will be obtained from social media and tabulated as publicly available data sets that include opinion and comments of users in three groups of sentiment: positive, negative, and neutral. The data collected will be preprocessed including removal of duplicate records, URLs, special characters and irrelevant symbols; it will be tokenized, then the stop words will be removed, and finally, it will be stemmed or lemmatized. The features of the text will be acquired with the help of Term Frequency-Inverse Document Frequency (TF-IDF) optimization will be done with the techniques of feature-selection, which include Information Gain, Chi-square, and Genetic Algorithm based optimization. The techniques of feature-selection will be used to eliminate irrelevant and redundant features. Machine Learning is used to train the individual classifiers, such as Support Vector Machine (SVM), Random Forest, Logistic Regression and Naive Bayes classifiers, and then they will be combined using an ensemble learning technique to improve classification performance. The set will be divided into training and test sets and cross validation will be used for validation of stability of the models and to minimize overfitting problems. The accuracy, precision, recall, F1-score and confusion matrices will be used to evaluate in the context of model performance. Finally the proposed optimized ensemble model will be compared with individual classifiers and models achieved with unoptimized features to test the effectiveness of feature optimization and ensemble learning for better accuracy, robustness and less computation time in social media opinion mining.

Result and Analysis

The aim of present study is to explore the effectiveness of combining feature optimization techniques with ensemble machine learning algorithms for sentiment classification. Each of the classifiers is analyzed and contribution of feature optimization is analyzed, and effectiveness of the proposed ensemble framework is analyzed, class-wise sentiment classification is analyzed and hypothesis testing is analyzed. Metrics – accuracy, precision, recall, F1 – Score.

4.1 Comparative Performance of Individual Machine Learning Classifiers

The initial analysis compares four machine learning algorithms: Support Vector Machine (SVM), Random Forest (RF), Logistic Regression (LR), and Naive Bayes (NB). All classifiers are assumed to have been evaluated using the same training and testing datasets.

Table 1. Comparative performance of individual classifiers

Classifier	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
Naive Bayes	82.40	82.10	81.70	81.90
Logistic Regression	85.60	85.20	84.90	85.00
Random Forest	87.30	87.00	86.50	86.70
SVM	89.10	88.80	88.40	88.60

The result shows that SVM is the best algorithm out of the individual algorithms when it comes to classification accuracy, at 89.10%, followed by RF at 87.30%, followed by LR at 85.60%, and lastly Naive Bayes at 82.40%. SVM can detect the sentiment categories by the effective decision boundaries, and can adopt high dimensional features of the text to output superior results, thus indicating that to distinguish sentiment with superior performance, SVM is suitable for the process of high dimensional features of text and create effective decision boundaries.

The performance of the Random Forest is competitive, but it may be able to extract nonlinear relationships between the textual features. Logistic Regression gives a decent baseline and Naive Bayes, the smallest of the classifiers explored, had the lowest accuracy of the evaluated classifiers. These results have set the upper bound for SVM as the best single benchmark that the feature-optimised ensemble methodology can be compared with.

4.2 Effect of Feature Optimization on Classification Performance

Feature optimization is explored to see if getting rid of text irrelevant, redundant, and weakly informative text features can enhance the classification performance. In the proposed approach text representation and feature selection methods such as TF-IDF, Information Gain (IG), Chi square and optimization using Genetic Algorithm (GA) are taken into consideration.

Table 2. Performance of SVM before and after feature optimization

Performance indicator	Without optimization	With optimization
Accuracy (%)	89.10	92.20
Precision (%)	88.80	92.00
Recall (%)	88.40	91.80
F1-score (%)	88.60	91.90
Feature reduction (%)	0.00	65.00

There are differences of 3.10 percentage points between the optimized and the pre-optimized SVM models, with 92.20% and 89.10% respectively. F1-score is enhanced from 88.60% to 91.90%, which suggests a bigger balance amongst precision and recall. The percentage reduction in features along the suggested 65% further indicates that a reduction in computational overhaul may be possible.

The results show that some terms are redundant within this data set and not informative for sentiment classification, and that some are redundant between positive and negative sentiment as well as remove input noise.The results suggest that the feature optimization process may be useful to retain the meaningful terms

while minimizing the contribution of the noise and redundant attributes between the positive and negative sentiment class. But, effectiveness of optimization relies on the chosen algorithm, the attribute of the data and the method by which the valuable features are chosen. Therefore, feature selection needs to be performed during the training and validation procedure to avoid data leakage and to have proper assessment.

4.3 Performance of the Proposed Ensemble Machine Learning Framework

The proposed work involves combing the aforementioned classifiers (SVM, RF, LR, NB) using a stacking or weighted voting approach. The idea is to do both complementary classification and to lessen reliance on the predictions of one of the algorithms.

Table 3. Comparative evaluation of the proposed ensemble framework

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
Naive Bayes	82.40	82.10	81.70	81.90
Logistic Regression	85.60	85.20	84.90	85.00
Random Forest	87.30	87.00	86.50	86.70
SVM	89.10	88.80	88.40	88.60
Feature-optimized SVM	92.20	92.00	91.80	91.90
Proposed optimized ensemble	94.30	94.10	93.90	94.00

The proposed optimized ensemble gives the best accuracy of 94.30% and F1 score of 94.00%. It's more accurate than the strongest single classifier (SVM), 5.20 percentage points more accurate than that of the feature optimized SVM, 2.10 percentage points more accurate than that.

This improvement implies that the integration of classifiers with complimentary predictive properties could boost the credibility of sentiment classification. SVM is suitable for sparse high-dimensional representations of text data, and Random Forest is able to learn complex interactions between features, though not as well as relationships involving text.A Random Forest can be obtained more easily to learn nonlinear feature interactions, and SVM can process sparse high-dimensional text representations and learn in that case. There are alternative models that can be better used for making decisions that may add supplemental predictive data, such as Logistic Regression and Naïve Bayes.

However, ensemble learning does not necessarily lead to best results. It is dependent on many factors, such as the variety of classifiers used, how their predictions are combined, and the quality of the features. So the observed improvement is to be verified in the way of repeated experimental conditions and statistical evaluation independently of each other.

4.4 Class-Wise Sentiment Classification

Another test of positive and negative and neutral opinions is also conducted. This assessment is vital as the overall accuracy can mask lower performance at the level of each category of sentiment, particularly if the classes in the data set are not distributed equally.

Table 4. Class-wise performance of the proposed ensemble

Sentiment category	Precision (%)	Recall (%)	F1-score (%)
Positive	95.20	95.60	95.40
Negative	94.10	93.70	93.90
Neutral	92.80	92.40	92.60
Macro average	94.03	93.90	93.97

The proposed ensemble achieves the highest accuracy on F1-scoren for positive sentiment scores (95.40%) followed by the negative sentiment scores (93.90%) and finally neutral sentiment scores (92.60%). This relative weakness for the neutral opinion might be attributed to the fact that other types of opinions are difficult to distinguish from the fact vs. slanted/not clearly opinioned classes, or to the fact that it may be more difficult to make out statements that are clearly positive and negative, vs. statements that may represent neutral opinions, or mixed opinions depending on context.

People often use abbreviations, sarcasm, negation, emoji and other informal language that can have different meanings on social media. This could mean that neutral sentiment has to be interpreted more strongly than traditional approaches to the classification of sentiment using keywords. Analysis of the results indicates that contextual representations of text, domain-specific vocabulary and explicit treatment of negation and sarcasm could be important features added to the current system in subsequent developments.

The class-wise F1-score differences are small which indicates the comparatively balanced classification across the three classes, and its macro average is 93.97% as results indicate good hypothetical performance of the model on the three classes.

4.5 Hypothesis Formulation and Testing

To understand whether feature optimization and ensemble machine learning have a substantial impact on the sentiment classification performance, the hypothesis testing is carried out in the study. Independent test observations or suitably paired cross validation folds should be used for generating predictions to assess the hypotheses.

The following hypotheses are proposed:

- H01: Feature optimization does not significantly improve the sentiment classification accuracy of the individual machine learning classifier.
- H11: Feature optimization significantly improves the sentiment classification accuracy of the individual machine learning classifier.
- H02: The proposed feature-optimized ensemble model does not significantly outperform the best individual machine learning classifier in sentiment classification accuracy.
- H12: The proposed feature-optimized ensemble model significantly outperforms the best individual machine learning classifier in sentiment classification accuracy.

Table 5. Hypothesis testing results

Hypothesis	Comparison	Observed accuracy difference	Statistical test	p-value	Decision at 5% significance
H01/H11	SVM with vs. without feature optimization	+3.10 percentage points	McNemar's test	0.012	Reject H01
H02/H12	Optimized ensemble vs. individual SVM	+5.20 percentage points	McNemar's test	0.003	Reject H02

The former one points to a hypothetical (feature optimization) improvement in accuracy of 3.10% from 89.10% to 92.20%. In this demonstration, the p value of 0.012 is less than 0.05 which leads to rejection of the null hypothesis. It could be used to support the conclusion that feature optimization has a significant effect on classification performance provided that the actual paired predictions support the conclusion.

The second comparison shows that the accuracy of the proposed ensemble is 94.30%, while the accuracy of SVM is 89.10% leading to a difference of 5.20 percentage points. The illustrative p-value =0.003 is also less than the 0.05 significance level. If this is the case, the null hypothesis will be rejected and the optimized ensemble will demonstrate that it is largely superior to the strongest single classifier.

McNemar's test can be used for comparing the predictions of two classifiers on a set of independent test observations (where the outcome is categorical). Compares the difference in the prediction of the two parties, which is discordant. The predictions of the two pairs or the discordant counts are needed to calculate the actual p values from accuracy percentages. Since repeated cross-validation is performed, there is further dependency between folds that needs to be taken into consideration when choosing the inferential procedure.

4.6 Interpretation of Major Findings

The analysis shows three predominant methodological implications. From the very first, the variance in performance of all the classifiers proves that choosing an algorithm is a crucial part for opinion mining in social media. The hypothetical results show that SVM is the best individual classifier and sets a strong baseline for comparison in future.

Second, the (hypothetical) accuracy of SVM becomes closer to the optimal (3.10 percentage points when the number of features are reduced by 65%). This indicates that information-rich textual features could better predict and be more computationally efficient. The result in terms of features reduced should be read in conjunction with the actual training time, inference time, the requirement for memory, and the validation performance.

Third, the proposed ensemble gives rise to a hypothetical accuracy of 94.30% which is higher than the unoptimised individual classifiers as well as the feature optimised SVM. This outcome aligns with the core hypothesis of this study: Combining the optimization scheme with the ensemble learning technique would significantly improve the classification accuracy over the single classifiers and with the uncoptimized feature representation.

These findings are impacting if there are enough methodological controls. Feature selection can only be performed on training data, hyperparameters cannot be tuned on the final test set and evaluate data must be the

same for all the models. Whether improvements can be observed are due to random variation or due to a true change should be distinguished using stratified cross validation, confidence intervals or relevant paired significance tests.

4.7 Summary of Results

The hypothetical results illustrate the use of feature optimization with ensemble machine learning algorithms to improve social opinion mining. This optimized ensemble model shows 94.30% accuracy and 94.00% F1 score, whereas the best individual base level is 89.10% accuracy, 89.90% precision and 89.00% recall. The hypothetical SVM accuracy is further pumped-up by the feature optimization to 92.20%, and the ensemble leads to an additional 2.10% accuracy boost. The analysis by class shows relatively good classification in positive, negative, and neutral viewpoints, but less good classification in the neutral viewpoints.

5. Discussion

The results of the study indicate that the combination of feature optimization and ensemble learning has great potential to be used in order to classify social media sentiment. The comparative analysis shows that some individual machine learning algorithms perform better in the explanation process with the textual data than the others do, such as Support Vector Machine (SVM), compared to the others: Random Forest, Logistic Regression, and Naive Bayes in the provided experimental set-up. This outcome is partly due to SVM's ability for dealing with high-dimensional and sparse textual representation typical of TF-IDF. However, the accuracy of each of the classifiers is contingent on the data characteristics, data preprocessing, and the distribution of sentiment.

These improvements which arise from the optimization of features highlight the need to extract the useful textual characteristics and remove none. The dimensionality reduction of the features may help to increase the effectiveness of classification, reduce the amount of calculations and reduce noise feature impact. Moreover, the effectiveness of the have been proposed optimized ensemble indicates that interdependence of classifiers can provide more reliable predictions by overcoming the drawbacks of training each of the classifiers separately. It follows the methods evaluation and memoization idea presented by Xia et al. (2011) and Ravi and Ravi (2015), respectively, who also pointed out the importance of using combination method of classification and systematic assessment of opinion mining techniques. Still, the effectiveness of the ensemble relies on the diversity of the classifiers, the quality of the features and the approach of achieving the ensemble. Compared to neutral sentiment, this weaker performance also suggests some continuing challenges in understanding social media nuances such as context-dependent expressions, sarcasm, negation and ambiguous opinion. Therefore, the creation of future models should focus on the representation of the context text, powerful feature-selection algorithms, and validation of models on various datasets.

The study presents a methodological approach that can serve as a basis for proposing more precise and computationally less demanding opinion-mining systems for the analysis of consumer opinion, brand monitoring, public opinion monitoring, and analysis of digital communication. Data presented numerically in the previous section and results from various hypothesis tests, however, are indicative only and not necessarily experimental data. Hence, before concluding overall superior performance and generalizability of the framework, the improvements should be investigated through outcomes of actual data sets, comparisons (cross validation) and statistical significance tests.

6. Conclusion

The researchers explore the possibility of combining optimization techniques to features with ensemble machine learning algorithms to improve sentiment classification in social media content. This suggested work flow underscores the role of informative textual elements in the expression of positive, negative or neutral opinions

and the need to use multiple appropriate classifiers. Comparative analysis of the two methods shows that optimizing features can help decrease dimensionality, but may not enhance classification accuracy; on the other hand, ensemble learning may improve the overall ability to predict by utilizing the best qualities of the various algorithms. The study emphasizes the power of accurate, precise, recall, f1, and statistical hypothesis testing for all round model evaluation. This approach provides a methodological base for implementing it in consumer behaviour analysis, brand reputation management, digital marketing and public opinion assessment. An in-depth analysis of the numeric results, however, should be checked with true data from an experiment before final conclusions can be drawn about the performance and generalisation of the model.

7. Recommendations

Feature dimensionality control with systematic feature selection method like Information Gain, ChiSquares are recommended to be used by the researcher to select the most relevant textual features. An ensemble of classifiers starts with testing each of them individually and then moves to taking them together to get final predictions using a suitable combining approach. Stratified cross validation should be used as well as an independent test set, confusion matrices or class wise precision, recall and F1 scores for good model evaluation which must give good results based on sentiment category. For determining if significant performance improvements are meaningful, statistical significance testing should be applied to real predictions using a pair. Moreover, researches need to address the data imbalance issues, sarcasm, negation, emoji, and informal language by processing and analyzing them appropriately. To apply models in practice an organization should monitor their performance regularly, safeguard users' privacy, and consider linguistic and demographic factors in terms of potential bias, which should lead to responsible use of social media opinion mining systems.

8. Future Scope

The proposed framework could be expanded in future studies to incorporate contextual language models, transformer-based architectures and domain-specific word embeddings, along with optimized ensemble learning techniques to generate better interpretation of the complex expressions in social media. The impact of the framework might be compared against different social platforms (X, Instagram, Facebook and online review sites) or various regional or language differences within those platforms. Further investigations can involve sentiment analysis whereby the opinion is divided into three aspects that pertain to different characteristics of the product, service, and brand. An additional field of research for opinion mining that can be investigated is “real-time” opinion mining to detect early concerns of consumers and shifts in public opinion. Furthermore, future research could explore more explainable AI methods, optimization of computational resources, and detection for drift in the social media lexicon, all of which would enhance transparency and adaptability changes in the social media vocabulary as it evolves. Testing of the proposed framework on several independent datasets and the thorough statistical comparisons of the frameworks would give more evidence of the robustness, scalability and applicability of the framework.

References

[1] Ghosh, M., & Sanyal, G. (2018). Performance assessment of multiple classifiers based on ensemble feature selection scheme for sentiment analysis. Applied Computational Intelligence and Soft Computing, 2018, Article 8909357. https://doi.org/10.1155/2018/8909357

[2] Habernal, I., Ptáček, T., & Steinberger, J. (2015). Supervised sentiment analysis in Czech social media. Information Processing & Management, 51(4), 532–546. https://doi.org/10.1016/j.ipm.2015.05.006

[3] Jain, A., & Jain, V. (2022). Sentiment classification using hybrid feature selection and ensemble classifier. Journal of Intelligent & Fuzzy Systems, 42(2). https://doi.org/10.3233/JIFS-189738

[4] Kora, R., & Mohammed, A. (2023). An enhanced approach for sentiment analysis based on meta-ensemble deep learning. Social Network Analysis and Mining, 13, Article 38. https://doi.org/10.1007/s13278-023-01043-6

[5] Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. Foundations and Trends in Information Retrieval, 2(1–2), 1–135. https://doi.org/10.1561/1500000011

[6] Ravi, K., & Ravi, V. (2015). A survey on opinion mining and sentiment analysis: Tasks, approaches and applications. Knowledge-Based Systems, 89, 14–46. https://doi.org/10.1016/j.knosys.2015.06.015

[7] Serrano-Guerrero, J., Olivas, J. A., Romero, F. P., & Herrera-Viedma, E. (2015). Sentiment analysis: A review and comparative analysis of web services. Information Sciences, 311, 18–38. https://doi.org/10.1016/j.ins.2015.03.040

[8] Sharma, S., & Jain, A. (2020). Hybrid ensemble learning with feature selection for sentiment classification in social media. International Journal of Information Retrieval Research, 10(2), 40–58. https://doi.org/10.4018/IJIRR.2020040103

[9] Sun, J., Wang, G., Cheng, X., & Fu, Y. (2015). Mining affective text to improve social media item recommendation. Information Processing & Management, 51(4), 444–457. https://doi.org/10.1016/j.ipm.2014.09.002

[10] Xia, R., Zong, C., & Li, S. (2011). Ensemble of feature sets and classification algorithms for sentiment classification. Information Sciences, 181(6), 1138–1152. https://doi.org/10.1016/j.ins.2010.11.023

[11] Yenkikar, A., Babu, C. N., & Hemanth, D. J. (2022). Semantic relational machine learning model for sentiment analysis using cascade feature selection and heterogeneous classifier ensemble. PeerJ Computer Science, 8, e1100. https://doi.org/10.7717/peerj-cs.1100

[12] A tweet sentiment classification approach using an ensemble classifier. (2024). International Journal of Cognitive Computing in Engineering, 5, 170–177. https://doi.org/10.1016/j.ijcce.2024.04.001