iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 Issue: VI Month of publication: June 2024

DOI: https://doi.org/10.22214/ijraset.2024.63404

A Review of: Ensemble Feature Selection Scheme-Based Performance Evaluation of Several Classifiers for Sentiment Analysis

Anupriya Singh[1], Shyamol Banerjee[2]
[1]Research Scholar, [2]Faculty, Dept. of Computer Science and Engineering, SRCEM COLLEGE Banmore, Morena, M.P. India .

Abstract: The rise in popularity of sentiment analysis can be attributed to the growing amount of user-generated content available on the internet in recent times. Robust machine learning models and effective feature selection procedures are necessary for effectively extracting sentiment from textual data. This paper provides a thorough examination of the evaluation of several classifiers' performances in sentiment analysis using an ensemble feature selection scheme. The suggested ensemble feature selection methodology aims to improve the overall efficacy of sentiment analysis models by combining the best aspects of several feature selection techniques. To find the most pertinent features for sentiment classification, feature selection techniques like information gain, chi-square, and recursive feature reduction are combined into an ensemble framework. By reducing the drawbacks of individual feature selection methods, the ensemble approach yields a feature subset that is more extensive. Multiple state-of-the-art classifiers, including as support vector machines, decision trees, and neural networks, are used to assess the efficacy of the ensemble feature selection approach. Benchmark sentiment analysis datasets covering a wide range of subjects and linguistic subtleties are used to train and evaluate the classifiers. The outcomes of the experiment show that the ensemble feature selection strategy greatly enhances sentiment analysis model performance for a variety of classifiers. Comparative evaluations highlight the advantages and disadvantages of each classifier in different scenarios, providing insight into how well-suited they are for sentiment analysis jobs in the real world. The research investigates how ensemble size and diversity affect overall performance, providing information about the ideal setup for sentiment analysis applications. The results of this study provide a methodical assessment of classifier performance in combination with an ensemble feature selection approach, which advances sentiment analysis techniques. This research offers a new method for sentiment analysis that combines several classifiers with an ensemble feature selection methodology. The findings emphasize how crucial it is to choose the right features for sentiment classification tasks and show how ensemble approaches can improve overall model performance.
Keywords: Sentiment Analysis, Classifiers, SVM, Ensemble, Machine Learning

I. INTRODUCTION

Sentiment analysis is becoming more and more popular in the information era because to the abundance of user-generated content on social media, blogs, and online forums. Sentiment analysis and interpretation of textual data are essential for many applications, including market research, brand monitoring, and customer feedback analysis. Opinion mining, or sentiment analysis, is the process of using computer methods to ascertain the sentiment—positive, negative, or neutral—expressed in a text. Though sentiment analysis has a lot of promise, selecting the right machine learning models and feature selection techniques are critical to its efficacy. Ensemble approaches have been a potent tool for enhancing the durability and performance of machine learning models in recent years. When compared to individual models, ensembles of models produce predictions that are more reliable and accurate. In the field of sentiment analysis, feature selection through the use of ensemble approaches has drawn interest as a potential means of improving model performance. This paper explores the complexities of sentiment analysis by evaluating the performance of several classifiers in conjunction with an ensemble feature selection scheme[1]–[5].

Due to its ability to impact decision-making processes and offer insightful information about public opinion, sentiment analysis has become essential in a number of industries. People can openly express their opinions on a wide range of topics, from goods and services to political developments and social challenges, thanks to the growth of internet platforms. It takes advanced computer technologies that can reliably identify feelings across several disciplines and languages to analyze this massive and dynamic textual data.Conventional methods for sentiment analysis frequently make use of rule-based systems, lexicons, or basic machine learning models.

However, because emotions can be conveyed in nuanced and context-dependent ways, these approaches struggle to capture the richness and diversity of language. In order to overcome these obstacles, attention has switched to the use of machine learning techniques, which allow models to adapt to the dynamic nature of language and learn from patterns in data. The accuracy and robustness of a variety of machine learning tasks have been remarkably improved by ensemble learning, which combines numerous models to create predictions. The idea behind ensemble approaches is to build a more robust and dependable forecasting system by utilizing the diversity of different models. The inclusion of ensemble approaches becomes especially interesting in the context of sentiment analysis since it can improve overall model performance and capture a wider range of linguistic nuances[6]–[10].

A key factor in the effectiveness of sentiment analysis algorithms is feature selection. To find the most pertinent and instructive aspects among the vast amount of features in textual data, thorough curation is necessary. This procedure is essential for lowering computing complexity, increasing interpretability, and increasing model correctness. The goal of feature selection techniques is to isolate a group of features that make up the majority of the model's predictive power. Inspired by the desire to develop more efficient techniques for sentiment analysis, this work investigates the relationship between ensemble learning and feature selection. We seek to elucidate the complexities of ensemble feature selection and its influence on model efficiency by evaluating the performance of several classifiers within the framework of sentiment analysis. The goal of this research is to develop more reliable and accurate sentiment analysis algorithms that can be adjusted to the way language changes over time and yield insightful data in a variety of fields[11]–[17].

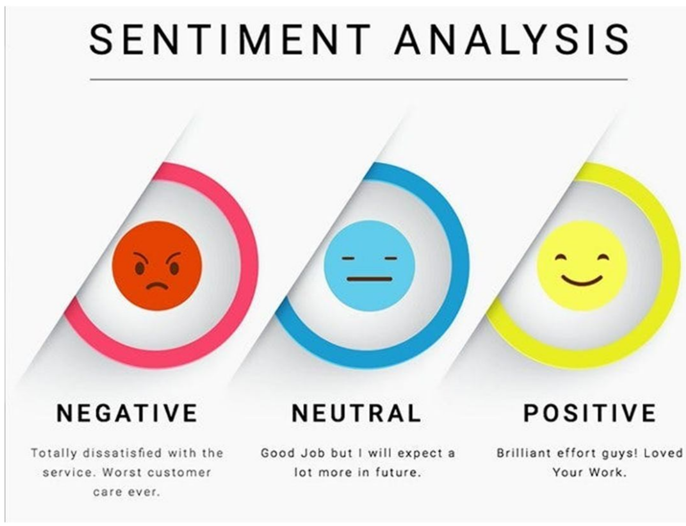

Fig. 1 Sentiment analysis

This study is unique in that it focuses only on ensemble feature selection strategies for sentiment analysis. The integration of many methods within an ensemble framework is a novel approach, while individual feature selection methods have been studied previously. By utilizing all of the methodologies' strengths, this comprehensive approach seeks to overcome the shortcomings of each one and determine which features are most informative for sentiment classification. By introducing a new dimension of intricacy to conventional sentiment analysis techniques, ensemble feature selection's inventiveness may open up new possibilities for enhanced model performance. Beyond the confines of academia, this work has practical ramifications that provide practitioners and decision-makers with useful information. The research offers a sophisticated understanding of which models are better suited for particular sentiment analysis tasks by assessing several classifiers under various circumstances. A useful guide for choosing pertinent features, cutting down on computational complexity, and improving model interpretability is provided by investigating ensemble feature selection strategies. These observations are extremely beneficial for companies trying to glean insights for strategic decision-making from massive volumes of textual data[18]–[24].

II. RELATED WORK

Başarslan 2023 et.al Sentiment analysis studies, which is a task of natural language processing are carried out to give people an idea and even guide them with such comments. In this study, sentiment analysis was implemented on public user feedback on websites in two different areas. TripAdvisor dataset includes positive or negative user comments about hotels. And Rotten Tomatoes dataset includes positive (fresh) or negative (rotten) user comments about films.

Sentiments analysis on datasets have been carried out by using Word2Vec word embedding model, which learns the vector representations of each word containing the positive or negative meaning of the sentences, and the Term Frequency Inverse Document Frequency text representation model with four machine learning methods (Naïve Bayes-NB, Support Vector Machines-SVM, Logistic Regression-LR, K-Nearest Neighbour-kNN) and two ensemble learning methods (Stacking, Majority Voting-MV). Accuracy and F-measure is used as a performance metric experiments. According to the results, Ensemble learning methods have shown better results than single machine learning algorithms. Among the overall approaches, MV outperformed Stacking[25].

Sarkar 2023 et. al illustrates the ROC, PRC, and cost curves that are used to represent the classification performance of the models. Additionally, parameters for evaluating the models' performance have been calculated for four testing options: the test data itself, cross-validation fold (CVF) 4, CVF 10, and percentage split (PS) 34% of the test data. These parameters include accuracy, sensitivity, specificity, false positive rate, false negative rate, precision, f-measure, kappa statistics, MCC, ROC area, and PRC area. Using the AlexNet CNN+BayesNet, AlexNet CNN+SMO, AlexNet CNN+NB, and AlexNet CNN+RF models, we have obtained accuracy values of 88.75%, 98.15%, 86.25%, and 100% for the test data. The outcomes suggest that our method is exceptional and highly successful[26].

Rijal 2023 et. al They will get prejudiced after a long period of time. Sentiment categorization classifies user reviews according to their positivity, negativity, or neutrality in an attempt to address this issue. The source of the dataset is Drone Emprit Academic. It is composed of up to 4887 data points that were retrieved from tweets that contained the phrase "online learning method." The feature selection approach makes use of information gain and adaboost on the C4.5 (FS+C4.5) method. Feature options help us reduce bias and increase accuracy. The experiment outcomes will be contrasted with those of other algorithms, such as random forest and C4.5. The findings showed that the two common decision tree models (random forest and C4.5) had increased their accuracy from 48.21% and 50.35% to 94.47%. The accuracy value increased by forty-four percent. In contrast, the FS+C4.5 model has a correlation of 0.944 and an RMSE of 0.204. Therefore, the C4.5 algorithm can become even more accurate by incorporating the feature selection technique into the sentiment analysis of bold learning instruction[27].

Ahmed 2023 et. al employs medical datasets for patients diagnosed with sepsis, and it analyses the efcacy of ensemble machine learning techniques compared to nonensemble machine learning techniques and the signifcance of data balancing and conditional tabular generative adversarial nets for data aug- mentation in producing reliable diagnosis. Te average F Score obtained by the nonensemble models trained in this paper is 0.83 compared to the ensemble techniques average of 0.94. Nonensemble techniques, such as Decision Tree, achieved an F score of 0.90, an AUC of0.90, and an accuracy of90%. Histogram-basedgradient boosting classifcation tree achieved an F score of0.96, an AUC of 0.96, and an accuracy of 95%, surpassing the other models tested. Additionally, when compared to the current state-of- the-art sepsis prediction models, the models developed in this study demonstrated higher average performance in all metrics, indicating reduced bias and improved robustness through data balancing and conditional tabular generative adversarial nets for data augmentation. Te study revealed that data balancing and augmentation on the ensemble machine learning algorithms boost the efcacy ofclinical predictive models and can help clinics decide which data types are most important when examining patients and diagnosing sepsis early through intelligent human-machine interface[28].

Yang 2022 et. al taking the county of Ziyang, China, as the study area, through historical reports, aerial-photo interpretations, and field investigations, 690 inventory maps of landslide locations were constructed and randomly divided into the 70/30 ratio for a training and validation dataset. Secondly, considering geological conditions, and landslide-induced disease and its characteristics, 14 landslide-conditioning factors was selected. Thirdly, the variance-inflation factor (VIF) and tolerance (TOL) were used to analyze the 14 factors, and the prediction ability was calculated with information-gain technology. Ultimately, the receiver-operating-characteristic (ROC) curve was applied to verify and compare model performance. Results showed that the LMT-RF model (0.897) was superior to other models, and the performance of LMT single model (0.791) was the worst. Therefore, it can be inferred that the LMT-RF model is a promising model, and the outcome of this study will be useful to planners and scientists in landslide sensitivity studies in similar situations[29].

TABLE .1 Literature Summary

Author/ year	Title	Method/ model	Parameters	References
Li 2022	An Ensemble Semantic Textual Similarity Measure Based on Multiple Evidences for Biomedical Documents	multi-evidence-based semantic text similarity calculation method	Purity =0.947 ARI=0.818	[30]

Isikdemir 2022	The Scalable Fuzzy Inference-Based Ensemble Method for Sentiment Analysis	fuzzy inference mechanism, Naive Bayes Logistic regression	Correct classification rates =92.87%	[31]
Sabeena 2022	Optimization-Based Ensemble Feature Selection Algorithm and Deep Learning Classifier for Parkinson's Disease	DMTs (data mining techniques) and MLTs (machine learning techniques) can	Accuracy =98.02 Error=1.97	[32]
Deepa 2022	Comprehensive Performance Analysis of Classifiers in Diagnosis of Epilepsy	Gaussian mixture model outperforms	Accuracy=92.19 %, performanceindex 81.43% good detection rate = 83.48%	[33]
Han 2022	A Feature Selection Method of the Island Algorithm Based on Gaussian Mutation	Gaussian mutation IAGM), Feature selection methods	Accuracy =0.92 Optimum value =0.97	[34]

TABLE 2 RESEARCH GAPS

Sr. no.	Authors	Year	Research gap
1.	Yang[35]	2022	the Whale Optimization Algorithm (WOA) for logistics distribution center space location selection, highlighting deficiencies in the initial population, global exploration, and local development stages. The study proposes improvements, including optimizing cross-selection strategies, incorporating direct and hierarchical logistics distribution, and enhancing algorithm performance through second reverse learning, chaotic mapping, logistic chaotic mapping, and adaptive inertia weight
2	Tang[36]	2022	improved feature extraction and selection methods in motor imagery EEG decoding for brain-computer interface (BCI) systems, with a focus on enhancing accuracy and reducing model training time, which is addressed through the introduction of a novel spatial-frequency extraction method and a hybrid Fisher score-SVM feature selection approach in this study
3	Li [37]	2022	a gap in brand image assessment by proposing a method based on consumer sentiment analysis, leveraging fragmented consumer topic data for cognitive label extraction and deep feature fusion. This approach proves effective in providing a nuanced understanding of consumer perceptions and accurately quantifying brand image
4	Wang [38]	2022	personalized music recommendation algorithms and proposes a method based on bipartite graph node structure similarity and restarted random wandering. This approach enhances recommendations by analyzing connections in the music social network, meeting users' personalized music preferences effectively
5	Duan[39]	2022	the challenge of quantifying public environmental emotions in response to atmospheric quality changes, proposing a sentiment prediction model based on public participation perception. The study establishes a regression model showing strong negative correlations between PM2.5, PM10, and public satisfaction, emphasizing the difficulty in achieving quantifiable public environmental emotions.

III.PERFORMANCE ASSESSMENT OF MULTIPLE CLASSIFIER

Performance assessment of multiple classifiers is a critical aspect of machine learning model evaluation, particularly in the context of multiclass classification problems. This process involves analyzing the effectiveness of various classifiers in distinguishing between multiple classes and providing insights into their strengths and weaknesses. Here, we will discuss key aspects of the performance assessment of multiple classifiers, including evaluation metrics, cross-validation, and ensemble methods.One fundamental aspect of assessing classifier performance is the selection of appropriate evaluation metrics. Common metrics for multiclass classification include accuracy, precision, recall, F1 score, and the confusion matrix. Accuracy measures the overall correctness of predictions, while precision and recall provide insights into the classifier's ability to make accurate positive predictions and capture all actual positives, respectively. The F1 score balances precision and recall. The confusion matrix offers a detailed breakdown of true positives, true negatives, false positives, and false negatives, aiding in a more comprehensive assessment.Cross-validation is a crucial technique to ensure the robustness of performance assessment. In multiclass classification, techniques such as k-fold cross-validation help mitigate issues related to data variability and ensure that the model's performance is consistent across different subsets of the dataset. By partitioning the data into multiple folds and training the classifier on different subsets while validating on others, cross-validation provides a more reliable estimate of the model's generalization performance. Ensemble methods play a pivotal role in enhancing classifier performance. Combining predictions from multiple classifiers can lead to improved accuracy and robustness. Common ensemble methods include bagging, boosting, and stacking. Bagging involves training multiple classifiers independently on random subsets of the training data and aggregating their predictions. Boosting focuses on sequentially training classifiers, giving more weight to misclassified instances to improve overall performance. Stacking combines the outputs of multiple classifiers as input to a meta-classifier, leveraging their diverse strengths.Moreover, it is essential to consider the computational efficiency and scalability of classifiers, especially when dealing with large datasets. Some classifiers may be more suitable for specific scenarios based on their training time, memory requirements, and prediction speed. Understanding the trade-offs between accuracy and computational cost is crucial for selecting the most appropriate classifier for a given application. the analysis of receiver operating characteristic (ROC) curves and area under the curve (AUC) can provide insights into the classifier's performance across different probability thresholds. These metrics are particularly relevant when dealing with imbalanced datasets, where some classes may have fewer instances than others. the performance assessment of multiple classifiers in multiclass classification involves a comprehensive analysis of various evaluation metrics, cross-validation techniques, and the incorporation of ensemble methods. A holistic approach that considers both the model's predictive accuracy and its computational efficiency is crucial for selecting the most suitable classifier for a given task. Regularly evaluating and fine-tuning classifiers based on real-world performance metrics ensures the continued effectiveness of machine learning models in handling complex multiclass classification problems.

IV. ENSEMBLE FEATURE SELECTION

Ensemble feature selection plays a pivotal role in enhancing the performance of multiple classifiers, particularly in the domain of sentiment analysis. This technique involves the careful selection and combination of relevant features from diverse sources to improve the overall predictive power of classifiers. In the context of sentiment analysis, where the goal is to discern the sentiment expressed in textual data, ensemble feature selection becomes crucial for capturing the nuances and intricacies of sentiment-laden language.

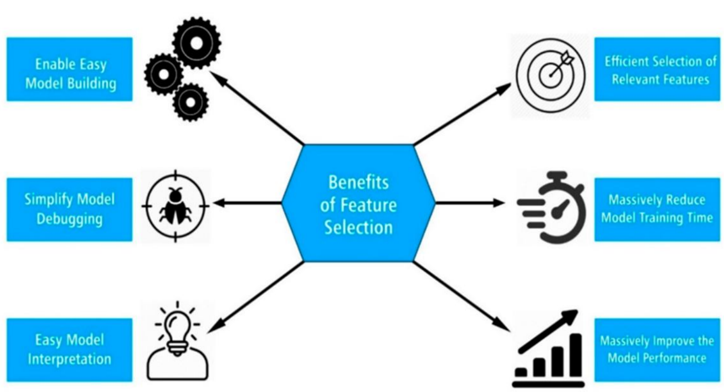

Fig. 2 Benefits of feature selection

One common challenge in sentiment analysis is dealing with high-dimensional feature spaces, especially when considering various linguistic features, syntactic structures, and semantic representations. Ensemble feature selection addresses this challenge by leveraging the strengths of multiple classifiers, each specializing in a subset of features. The process can be divided into several key steps:

1) *Feature Extraction:* Begin by extracting a wide array of features from the textual data. These features can include word frequencies, n-grams, sentiment lexicons, syntactic structures, and semantic embeddings. The goal is to create a rich feature space that captures both surface-level patterns and deeper contextual meanings within the text.
2) *Classifier Diversity:* Select a diverse set of classifiers, each designed to excel in capturing specific aspects of sentiment. For instance, one classifier may focus on lexical features, while another may specialize in syntactic structures. The diversity in classifiers ensures a comprehensive coverage of the feature space.
3) *Individual Classifier Training:* Train each classifier on its respective subset of features. This step involves optimizing the model parameters and fine-tuning to maximize performance based on the chosen evaluation metrics for sentiment analysis, such as accuracy, precision, recall, or F1 score.
4) *Ensemble Construction:* Combine the outputs of individual classifiers using an ensemble method. Common ensemble techniques include bagging, boosting, or stacking. Bagging may involve averaging the predictions, while boosting may assign different weights to classifiers based on their performance. Stacking combines the predictions as inputs to a meta-classifier, allowing for a higher-level integration of diverse feature subsets.
5) *Feature Importance Estimation:* Assess the importance of features within the ensemble. Techniques such as recursive feature elimination, information gain, or permutation importance can be employed to rank and select the most informative features. This step helps identify the features contributing significantly to sentiment prediction.
6) *Performance Evaluation:* Evaluate the performance of the ensemble feature selection scheme using appropriate sentiment analysis metrics. Cross-validation is often applied to ensure robustness and generalizability of the results across different subsets of the dataset.
7) *Iterative Refinement:* Fine-tune the ensemble feature selection scheme iteratively based on performance feedback. This may involve adjusting the ensemble configuration, adding or removing classifiers, or refining feature selection criteria to achieve optimal sentiment analysis results.

Ensemble feature selection for sentiment analysis involves a systematic approach to harness the strengths of multiple classifiers, each specializing in a distinct subset of features. This technique addresses the challenges posed by high-dimensional feature spaces and enhances the overall predictive capability of sentiment analysis models. Through careful construction and evaluation of the ensemble, sentiment analysis systems can achieve improved accuracy and robustness in discerning sentiment from diverse textual data sources.

V. CONCLUSION

In conclusion, the exploration and implementation of a Performance Assessment of Multiple Classifiers Based on Ensemble Feature Selection Scheme for Sentiment Analysis underscore the significance of a nuanced and comprehensive approach in handling the intricacies of sentiment-laden textual data. The integration of ensemble feature selection techniques has demonstrated its efficacy in enhancing the predictive power of sentiment analysis models by effectively addressing the challenges associated with high-dimensional feature spaces. The ensemble feature selection process begins with the extraction of diverse linguistic, syntactic, and semantic features from the textual data, creating a rich and multifaceted feature space. The deliberate selection of a diverse set of classifiers, each specializing in distinct feature subsets, ensures a holistic coverage of sentiment-related patterns and nuances. The subsequent training of individual classifiers and the construction of an ensemble through methodologies like bagging, boosting, or stacking enable a synergistic utilization of the classifiers' strengths. One of the key advantages of ensemble feature selection lies in its ability to weigh the importance of various features within the ensemble. Techniques such as recursive feature elimination and information gain contribute to the identification of the most informative features, shedding light on the aspects of language crucial for sentiment prediction. This feature importance estimation facilitates a more nuanced understanding of the factors influencing sentiment, aiding in the interpretability of the sentiment analysis model. The performance assessment of the ensemble feature selection scheme involves rigorous evaluation using sentiment analysis metrics, with a particular emphasis on accuracy, precision, recall, and F1 score. The application of cross-validation ensures the reliability and generalizability of the model's performance across diverse subsets of the dataset. This iterative process of refinement, guided by performance feedback, allows for continuous improvement and optimization of the ensemble feature selection scheme.

By leveraging the collective intelligence of multiple classifiers and strategically selecting and combining relevant features, the ensemble feature selection scheme not only mitigates the challenges posed by high-dimensional feature spaces but also enhances the robustness and accuracy of sentiment analysis models. This approach holds promise for real-world applications where capturing subtle nuances and context-specific sentiments is crucial, such as in social media monitoring, customer feedback analysis, and market sentiment prediction. In essence, the integration of ensemble feature selection techniques into the performance assessment of multiple classifiers for sentiment analysis represents a sophisticated and effective strategy, contributing to the advancement of sentiment analysis methodologies and their applicability in diverse domains. As sentiment analysis continues to play a vital role in understanding human expressions in text, the refined and nuanced models resulting from this ensemble approach pave the way for more accurate, interpretable, and reliable sentiment predictions in real-world scenarios.

REFERENCES

[1] L. Yu, Y. Wu, J. Yang, and Y. Zhang, "Bullet Subtitle Sentiment Classification Based on Affective Computing and Ensemble Learning," Wirel. Commun. Mob. Comput., vol. 2021, 2021, doi: 10.1155/2021/5563104.

[2] G. Pang et al., "Aspect-Level Sentiment Analysis Approach via BERT and Aspect Feature Location Model," Wirel. Commun. Mob. Comput., vol. 2021, 2021, doi: 10.1155/2021/5534615.

[3] W. Xiao, P. Ji, and J. Hu, "RnkHEU: A Hybrid Feature Selection Method for Predicting Students' Performance," Sci. Program., vol. 2021, 2021, doi: 10.1155/2021/1670593.

[4] C. Huang, "Feature Selection and Feature Stability Measurement Method for High-Dimensional Small Sample Data Based on Big Data Technology," Comput. Intell. Neurosci., vol. 2021, 2021, doi: 10.1155/2021/3597051.

[5] Z. Lu, G. Jiang, Y. Guan, Q. Wang, and J. Wu, "A SAR Target Recognition Method Based on Decision Fusion of Multiple Features and Classifiers," Sci. Program., vol. 2021, 2021, doi: 10.1155/2021/1258219.

[6] N. Alghamdi and S. Khatoon, "Improving Sentiment Prediction using Heterogeneous and Homogeneous Ensemble Methods: A Comparative Study," Procedia Comput. Sci., vol. 194, pp. 60–68, 2021, doi: 10.1016/j.procs.2021.10.059.

[7] J. T. Pintas, L. A. F. Fernandes, and A. C. B. Garcia, Feature selection methods for text classification: a systematic literature review, vol. 54, no. 8. Springer Netherlands, 2021. doi: 10.1007/s10462-021-09970-6.

[8] J. Khan, A. Alam, and Y. Lee, "Intelligent Hybrid Feature Selection for Textual Sentiment Classification," IEEE Access, vol. 9, pp. 140590–140608, 2021, doi: 10.1109/ACCESS.2021.3118982.

[9] L. Zhou, S. M. Jiang, and T. Yan, "A Network Selection Scheme Based on the Analytic Hierarchy Process for Marine Internet," Wirel. Commun. Mob. Comput., vol. 2021, 2021, doi: 10.1155/2021/8861107.

[10] Z. Yuan, T. Zhou, J. Liu, C. Zhang, and Y. Liu, "Fault Diagnosis Approach for Rotating Machinery Based on Feature Importance Ranking and Selection," Shock Vib., vol. 2021, 2021, doi: 10.1155/2021/8899188.

[11] J. Zou, X. Fu, L. Guo, C. Ju, and J. Chen, "Creating Ensemble Classifiers with Information Entropy Diversity Measure," Secur. Commun. Networks, vol. 2021, 2021, doi: 10.1155/2021/9953509.

[12] Y. Li, Z. Sun, X. Liu, W. T. Chen, D. J. Horng, and K. K. Lai, "Feature Selection Based on a Large-Scale Many-Objective Evolutionary Algorithm," Comput. Intell. Neurosci., vol. 2021, 2021, doi: 10.1155/2021/9961727.

[13] X. Liu, T. Li, R. Zhang, D. Wu, Y. Liu, and Z. Yang, "A GAN and Feature Selection-Based Oversampling Technique for Intrusion Detection," Secur. Commun. Networks, vol. 2021, 2021, doi: 10.1155/2021/9947059.

[14] B. Chen, H. Chen, and M. Li, "Feature Selection Based on BP Neural Network and Adaptive Particle Swarm Algorithm," Mob. Inf. Syst., vol. 2021, 2021, doi: 10.1155/2021/6715564.

[15] C. Zhou et al., "A Novel Stacking Heterogeneous Ensemble Model with Hybrid Wrapper-Based Feature Selection for Reservoir Productivity Predictions," Complexity, vol. 2021, 2021, doi: 10.1155/2021/6675638.

[16] M. Peng, W. Liu, T. Wang, and Z. Zeng, "Relay Selection Joint Consecutive Packet Routing Scheme to Improve Performance for Wake-Up Radio-Enabled WSNs," Wirel. Commun. Mob. Comput., vol. 2020, 2020, doi: 10.1155/2020/7230565.

[17] J. Dong et al., "Potential Trend for Online Shopping Data Based on the Linear Regression and Sentiment Analysis," Math. Probl. Eng., vol. 2020, 2020, doi: 10.1155/2020/4591260.

[18] X. Ding, L. Wang, W. Huang, Q. He, and Y. Shao, "Feature Clustering Analysis Using Reference Model towards Rolling Bearing Performance Degradation Assessment," Shock Vib., vol. 2020, 2020, doi: 10.1155/2020/6306087.

[19] K. Sundararajan and A. Palanisamy, "Multi-rule based ensemble feature selection model for sarcasm type detection in Twitter," Comput. Intell. Neurosci., vol. 2020, 2020, doi: 10.1155/2020/2860479.

[20] S. Rani and N. S. Gill, "Hybrid model for twitter data sentiment analysis based on ensemble of dictionary based classifier and stacked machine learning classifiers-SVM, KNN and C5.0," J. Theor. Appl. Inf. Technol., vol. 98, no. 4, pp. 624–635, 2020.

[21] S. Bhola, R. Malhotra, and G. Arora, "' Improved Tweet Sentiment Classification Using semantic Base Features with an ensemble learning Approach ,'" vol. 8, no. 12, pp. 706–713, 2020.

[22] R. Asgarnezhad, S. A. Monadjemi, and M. Soltanaghaei, "A High-Performance Model based on Ensembles for Twitter Sentiment Classification," J. Electr. Comput. Eng. Innov., vol. 8, no. 1, pp. 41–52, 2020, doi: 10.22061/JECEI.2020.7100.357.

[23] M. Soheili, A. M. E. Moghadam, and M. Dehghan, "Statistical Analysis of the Performance of Rank Fusion Methods Applied to a Homogeneous Ensemble Feature Ranking," Sci. Program., vol. 2020, 2020, doi: 10.1155/2020/8860044.

[24] H. Babajanian Bisheh, G. Ghodrati Amiri, and E. Darvishan, "Ensemble Classifiers and Feature-Based Methods for Structural Damage Assessment," Shock Vib., vol. 2020, 2020, doi: 10.1155/2020/8899487.

[25] M. S. Başarslan and F. Kayaalp, "Sentiment analysis with ensemble and machine learning methods in multi-domain datasets," Turkish J. Eng., vol. 7, no. 2, pp. 141–148, 2023, doi: 10.31127/tuje.1079698.

[26] A. Sarkar, M. Maniruzzaman, M. A. Alahe, and M. Ahmad, "An Effective and Novel Approach for Brain Tumor Classification Using AlexNet CNN Feature Extractor and Multiple Eminent Machine Learning Classifiers in MRIs," J. Sensors, vol. 2023, 2023, doi: 10.1155/2023/1224619.

[27] S. Rijal, P. A. Cakranegara, E. M. S. S. Ciptaningsih, P. H. Pebriana, A. Andiyan, and R. Rahim, "Integrating Information Gain methods for Feature Selection in Distance Education Sentiment Analysis during Covid-19," TEM J., vol. 12, no. 1, pp. 285–290, 2023, doi: 10.18421/TEM121-35.

[28] A. Alfakeeh, M. S. Sharif, A. D. Zorto, and T. Pillonetto, "Conditional Tabular Generative Adversarial Net for Enhancing Ensemble Classifiers in Sepsis Diagnosis," vol. 2023, 2023.

[29] P. Yang, N. Wang, Y. Guo, X. Ma, and C. Wang, "Performance Analysis of Logistic Model Tree-Based Ensemble Learning Algorithms for Landslide Susceptibility Mapping," J. Sensors, vol. 2022, 2022, doi: 10.1155/2022/8254356.

[30] M. Li, X. Zhou, K. H. Ryu, and N. Theera-Umpon, "An Ensemble Semantic Textual Similarity Measure Based on Multiple Evidences for Biomedical Documents," Comput. Math. Methods Med., vol. 2022, 2022, doi: 10.1155/2022/8238432

[31] Y. E. Isikdemir and H. S. Yavuz, "The Scalable Fuzzy Inference-Based Ensemble Method for Sentiment Analysis," Comput. Intell. Neurosci., vol. 2022, 2022, doi: 10.1155/2022/5186144.

[32] B. Sabeena, S. Sivakumari, and D. M. Teressa, "Optimization-Based Ensemble Feature Selection Algorithm and Deep Learning Classifier for Parkinson's Disease," J. Healthc. Eng., vol. 2022, 2022, doi: 10.1155/2022/1487212.

[33] R. Deepa, R. Anand, D. Pandey, B. K. Pandey, and B. Karki, "Comprehensive Performance Analysis of Classifiers in Diagnosis of Epilepsy," Math. Probl. Eng., vol. 2022, 2022, doi: 10.1155/2022/1559312.

[34] L. Han, H. Xu, J. Ma, and Z. Jia, "A Feature Selection Method of the Island Algorithm Based on Gaussian Mutation," Wirel. Commun. Mob. Comput., vol. 2022, 2022, doi: 10.1155/2022/1438999.

[35] L. Yang and X. Song, "High-Performance Computing Analysis and Location Selection of Logistics Distribution Center Space Based on Whale Optimization Algorithm," Comput. Intell. Neurosci., vol. 2022, 2022, doi: 10.1155/2022/2055241.

[36] Y. Tang, Z. Zhao, S. Zhang, Z. Li, Y. Mo, and Y. Guo, "Motor Imagery EEG Decoding Based on New Spatial-Frequency Feature and Hybrid Feature Selection Method," Math. Probl. Eng., vol. 2022, 2022, doi: 10.1155/2022/2856818.

[37] Z. Li, "Research on Brand Image Evaluation Method Based on Consumer Sentiment Analysis," Comput. Intell. Neurosci., vol. 2022, 2022, doi: 10.1155/2022/2647515.

[38] D. Wang, "Analysis of Sentiment and Personalised Recommendation in Musical Performance," Comput. Intell. Neurosci., vol. 2022, 2022, doi: 10.1155/2022/2778181

[39] L. Duan, "Correlation Analysis between Atmospheric Environment and Public Sentiment Based on Multiple Regression Model," Wirel. Commun. Mob. Comput., vol. 2022, 2022, doi: 10.1155/2022/3393079.