Sentiment Analysis Using Support Vector Machines, Neural Networks, and Random Forests

Colin Kai Wang
Seven Lakes High School, Katy, Texas, 77494, USA
W1002491@students.katyisd.org

Abstract. Sentiment analysis, often referred to as opinion mining, plays a crucial role in the field of natural language processing by computationally analyzing text to extract sentiments or opinions. It has become essential in understanding public sentiment, evaluating customer feedback, and identifying market trends, particularly in the age of online platforms and social media. The background of sentiment analysis stems from the need to process vast amounts of textual data and extract meaningful sentiment information. Manual analysis is impractical due to the exponential growth of online content, prompting the development of automated sentiment analysis techniques. Sentiment analysis holds significant value in uncovering insights related to customer satisfaction, brand perception, and sentiment trends. It empowers businesses to improve strategies, enhance customer experiences, and researchers to understand public opinion on various issues. Government agencies can also benefit from assessing public sentiment for evidence-based decision-making. This paper provides an overview of three prominent sub-topics in sentiment analysis: Support Vector Machines (SVMs), Neural Networks, and Random Forests. These approaches have gained significant attention in both research and industry applications, contributing to sentiment classification advancements.

Keywords: Sentiment Analysis, Support Vector Machines, Neural Networks.

1 Introduction

Sentiment analysis is a field of study in natural language processing that aims to computationally analyze text and extract sentiments or opinions expressed by individuals or groups. In an era dominated by online platforms and social media, understanding and interpreting sentiments have become crucial for various applications, including gauging public opinion, evaluating customer feedback, and identifying market trends. This introduction provides a background on sentiment analysis, highlighting its significance and origins, followed by a brief overview of the main sub-topics covered in this paper.

The exponential growth of online content has made it increasingly challenging to manually process and interpret sentiments expressed in textual data. As a result, sentiment analysis has emerged as a solution to automate this process, enabling efficient analysis of sentiments at scale. The need to extract meaningful sentiment information from vast volumes of text has driven the development of sentiment

P. Kar et al. (eds.), *Proceedings of the 2023 International Conference on Image, Algorithms and Artificial Intelligence (ICIAAI 2023)*, Advances in Computer Science Research 108,
https://doi.org/10.2991/978-94-6463-300-9_4

analysis techniques and methodologies. By leveraging computational algorithms, sentiment analysis has become a valuable tool for analyzing sentiments within texts and extracting actionable insights.

The origins of sentiment analysis can be traced back to the early 2000s when researchers recognized the need to automate the analysis of sentiments expressed in online forums, reviews, and social media platforms. The availability of large-scale datasets and advancements in machine learning algorithms paved the way for the development of sentiment classification models. Since then, sentiment analysis has garnered significant attention from researchers, practitioners, and industry professionals due to its potential for understanding human sentiment at scale.

This paper aims to provide an overview of Support Vector Machines, Neural Networks, and Random Forests. Each of these sub-topics represents distinct approaches to sentiment classification and has gained significant traction in both research and industry applications.

Support Vector Machines are a widely used machine learning algorithm known for its ability to classify data points into different classes by finding an optimal hyperplane. In sentiment analysis, Support Vector Machines have been successfully employed to classify text documents according to their sentiments. SVMs have demonstrated remarkable performance in sentiment classification tasks. The objective of SVMs is to identify a hyperplane in the attribute space that successfully distinguishes data points belonging to different classes. This is achieved by employing the principle of structural risk minimization, which seeks to strike a balance between minimizing classification errors and maximizing the margin between classes. By optimizing this balance, SVMs can effectively classify text documents based on their sentiments.

Sentiment analysis has undergone a revolutionary transformation in the presence of Neural Networks, specifically deep learning architectures like Recurrent Neural Networks (RNNs) and Convolutional Neural Networks (CNNs). These advanced models have reshaped the landscape of sentiment analysis and brought about significant advancements in the field. These models excel at capturing intricate patterns in sequential and textual data, enabling more nuanced sentiment analysis and achieving state-of-the-art performance in sentiment classification tasks. Neural networks are composed of interconnected nodes called neurons, organized into multiple layers, including an input layer, hidden layers (which can vary in number), and an output layer. These neurons in the layers receive inputs, perform calculations, and produce outputs, transferring those outputs to the next layer, and the connections between the neurons are associated with weights. These weights are adjusted accordingly in the training process to minimize a loss function, eventually reaching the optimal weights to accomplish the task given.

Random Forest is a robust algorithm that harnesses the combined insights of multiple decision trees to deliver accurate predictions. The main concept behind them are the combinations of decision trees, where each tree is trained on a random subset of training data with a technique called bootstrap aggregating. Moreover, in each split of a decision tree, a random subset of features is considered, introducing additional randomness and aiding in the decorrelation of the trees. This characteristic has contributed to the popularity of Random Forest in sentiment analysis, as it effectively handles high-dimensional feature spaces and alleviates the issue of overfitting. Random Forest has been employed for sentiment classification, where it leverages the collective wisdom of multiple decision trees to predict sentiments accurately.

This paper aims to offer an extensive overview of the background, significance, and key aspects within the field of sentiment analysis. Furthermore, by examining the distinct approaches of the three topics, this paper contributes to the understanding and advancement of sentiment analysis techniques, facilitating their effective application in real-world scenarios.

2 Support Vector Machines

Support Vector Machines are a type of machine learning algorithm designed to address classification and regression tasks, leveraging a robust mathematical framework for supervised learning [1]. They are particularly strong at solving binary classification, but also can handle multi-class classification.

The key advantages of SVMs are:

- Flexibility: SVMs demonstrate flexibility in their ability to handle various types of data, including both linearly separable and non-linearly separable datasets. This flexibility is achieved by employing kernels that enable the mapping of the original feature vector space to a higher-dimensional space. In this transformed space, the data points can potentially become linearly separable, thus allowing SVMs to effectively classify diverse and complex datasets
- Robustness to Outliers: SVMs are not affected as much by outliers due to their focus on finding a decision boundary that has the largest margin between different classes
- Regularization: SVMs have a regularization parameter C that manages the tradeoff between achieving large margins and minimizing classification error, preventing overfitting and increasing generalization

This paper discusses the following 2 methods to optimize SVMs for Sentiment analysis: Grid-Search Technique and Kernel Function Selection.

2.1 SVM Grid-Search Technique

Ahmad et al. reported the effect of using SVM Grid-search technique in combination with 10-fold cross validation [2]. The proposed method introduced in the paper is referred to as the Optimized Sentiment Analysis Framework (OSAF), which extends the standard Sentiment Analysis Framework.

The study makes use of three datasets, comprising two datasets sourced from Twitter and one dataset sourced from IMDB reviews. The first dataset consists of Twitter tweets about Apple, Google, Twitter, and Microsoft, categorized as positive, negative, or neutral. The second dataset contains positive, negative, or neutral tweets about US airlines, while the third dataset comprises positive or negative IMDB reviews. These datasets provide diverse and real-world texts for evaluating the OSAF framework.

The OSAF framework comprises four phases: Dataset preparation, preprocessing, classification, and result analysis. Initially, the researchers input the data into the Weka software to prepare it for analysis. The text strings are converted into feature vectors using various techniques, including Term Frequency Inverse Document Frequency, word stemming, elimination of stop words, tokenization, and word count restriction. The extracted features are fed into the SVM model. To optimize the model, grid search is employed to find the best hyperparameters, specifically the values of C and ϒ, while k-fold cross-validation helps prevent overfitting. Finally, they calculate recall, precision, and f-measure metrics to analyze the performance of the model.

Through experimentation with three diverse datasets, the authors demonstrate the framework's effectiveness in improving classification performance. The findings highlight the potential of this approach to enhance sentiment analysis in real-world applications.

2.2 SVM Sentiment Analysis Kernel Function Selection

In their study, Prastyo et al. aim to employ sentiment analysis with Support Vector Machines to record the public positive or negative sentiment towards the COVID-19 pandemic, enabling the government to enact better responses. They accomplish this by selecting a suitable kernel function for the SVM, which provides the necessary learning abilities for the learning algorithm to correctly classify sentiments [3].

Their dataset comprises of scraped data from the Twitterscraper library, scraping 20,444 tweets for general tweets about COVID in Indonesia and 14,451 tweets about economic impacts, later being reduced to 2,203 for general tweets and 1,941 for economic tweets by removing irrelevant and repeated data. They were then separated into two datasets: one had positive, neutral, and negative classes, while the other had positive and negative. This was done in order to determine the effectiveness of the model at handling 2 classes compared to 3.

Preprocessing data was then done, consisting of case folding, filtering, tokenization, stopword removal, and stopword removal, before feature extraction was applied. The researchers then utilized Term Frequency-Inverse Document Frequency (TF-IDF) to extract features, converting the text into vectors that can be fed to the machine learning model. This measures the frequency a term appears while also computing the relative importance of that term to the overall sentiment. The data was then divided into distinct classes by utilizing the Support Vector Machine algorithm, which aimed to identify the ideal decision boundary; however, a suitable kernel function was needed for it to be effective, and the researchers found that the Normalized Poly kernel was the most suitable for the application as it performed the best. Finally, to mitigate overfitting and evaluate the model's performance, k-fold cross-validation was employed, splitting training and testing data.

The SVM performed relatively well on both datasets; however, it performed better on the dataset with only 2 classes (positive and negative). Overall, SVMs are effective for working with high dimensional data and large feature vector spaces, providing a comprehensive solution for sentiment analysis of large collections of text.

3 Neural Networks

Neural Networks are very powerful tools, capable of capturing complex patterns and relationships in data. Modeled after the brain, they contain neurons that process data and generate outputs.

The key advantages to neural networks are:

- Adaptability: Neural networks are very flexible as they learn from experience, making them suitable for a variety of tasks from image recognition to time series analysis
- Parallel Processing: Neural networks can perform parallel computations across multiple layers, allowing for more efficient training and utilization
- Non-linearity: Neural networks excel at capturing intricate patterns and dependencies that are often challenging for traditional linear models to comprehend. By leveraging their capacity to model complex, non-linear relationships between inputs and outputs, neural networks can uncover and represent the underlying structures present in the data. This ability allows them to capture fine-grained details and subtle connections that linear models may struggle to capture effectively

This paper will discuss three types of neural networks for sentiment analysis: convolutional neural networks with feature extraction, deep neural networks, and deep convolutional neural networks that utilize feature extraction.

3.1 Convolutional Neural Networks

Yang et al. discuss convolutional neural networks as an aspect of sentiment analysis, highlighting the effectiveness of CNNs in learning semantic features and their superiority over other models [3].

The paper depicts a CNN as a multilayer feedforward neural network, having layers composed of two-dimensional planes that each contain independent neurons. It can be used for image processing, but similar techniques can be applied to sentiment analysis as well. By inputting a matrix that represents a sentence or document through feature extraction, it creates an “image” as a result.

Word2vec is a common tool for natural language processing. It can perform a variety of tasks such as finding synonyms, analyzing part of speech, etc, without outside intervention. Although CNNs don’t require external feature extraction, Word2vec is an effective tool to help better train the CNN model.

Results: The theoretical accuracy described in the study for Neural Networks is high, meaning it has a high probability of correctly classifying the sentiment of texts, in contrast with other theoretical accuracies described in the paper such as SVMs. The study also describes the theoretical training speed of neural networks as low, reiterating the benefits of using them. However, it also presents some challenges for neural networks. Although they are effective for tasks like sentiment analysis, there lacks development in other languages, and additionally, due to the variety of neural network approaches, aspects such as feature sets differ between them.

3.2 Deep Neural Networks

Deep learning techniques have been applied to sentiment analysis of Twitter data, with Jianqiang et al. reporting their findings on the use of machine learning methods for handling the vast amount of data from Twitter and extracting valuable insights into aspects like public perception [6].

To train and evaluate the model, the authors compiled a dataset comprising both Korean and English tweets. The dataset was divided into two main sets: a training set consisting of a combination of 2,000 positive and negative tweets , and a separate testing set containing another 2,000 tweets of varied sentiments.

The employed approach for sentiment analysis involved a series of steps. Firstly, text mining techniques were applied to retrieve relevant tweets from the Twitter database using the Twitter API. Then, text cleaning and preprocessing were performed, which included steps such as lowercasing, stemming or lemmatization, removing numbers, stop words, punctuation, and eliminating extra white spaces.

For sentiment analysis, a deep neural network with three hidden layers was implemented. The first layer filters words, the second filters sentences, and the third filters based on word popularity in online dictionaries. The neural network architecture utilized the rectified linear unit and sigmoid activation functions. To train the network, the authors employed the mean square unit and stochastic gradient descent optimization techniques. The training process consisted of 100 epochs, where different learning rates (0.1 and 0.001) were utilized to improve the performance of the model.

The accuracy of the sentiment analysis model was evaluated on both the training and testing datasets. The deep learning approach demonstrated promising results, achieving an accuracy of 77.45% during training and maintaining high accuracy with 75.03% during testing. In contrast, the performance of the multi-layer perceptron model was notably lower, with an accuracy of 67.45% during training and 52.60% during testing. These findings highlight the superior performance and effectiveness of the proposed neural network model.

3.3 Deep Convolutional Neural Networks

Ramadhani and Goo present another novel approach for sentiment analysis of tweets, utilizing deep learning techniques [7]. The authors address the challenge of determining the sentiment polarity of tweets. While existing methods focus on analyzing lexical and syntactic features, this paper introduces a word embedding generated from unsupervised learning based on large Twitter data. They leverage word embeddings, n-grams features, and word sentiment polarity scores to construct a robust set of sentiment features for tweets, enhancing the efficacy of machine learning models. A deep convolution neural network (DCNN) is then used for training and predicting sentiment classification labels. Results are then compared to baseline approaches, demonstrating the effectiveness of the DCNN.

The study assesses the effectiveness of the GloVe-DCNN model by evaluating its performance on five different Twitter tweet datasets.

The methodology of the proposed GloVe-DCNN model is composed of a number of essential elements. To begin, the authors utilize word N-grams features, including both unigrams and bigrams, as the foundational feature model. These features capture the fundamental characteristics of the tweet text. Furthermore, Twitter-specific attributes such as the count of hashtags, capitalization, and emoticons are incorporated into the feature extraction process.

Furthermore, the researchers incorporate word sentiment polarity score features, which leverage lexicon-based sentiment analysis. The authors utilize the AFINN lexicon as a foundation and augment it with SentiWordNet to derive sentiment polarity scores for each word contained within the tweet. This comprehensive approach enables the assignment of sentiment values to effectively capture the overall sentiment expressed in the tweet. The calculation of a tweet's sentiment score involves adding up the sentiment polarity scores of all the words contained within the tweet.

The primary significance of the paper lies in the incorporation of word representation features into the model, enhancing its overall performance and capabilities. The authors use unsupervised learning to extract word vector representations from a sizable Twitter dataset using the GloVe model. These word embeddings capture semantic and grammatical characteristics of words that assist in improving the model's performance. The GloVe-DCNN model effectively integrates the word embeddings with the baseline features, sentiment polarity scores, and Twitter-specific features, creating a rich and comprehensive set of sentiment features that enhances its performance.

Subsequently, the sentiment feature set is utilized as input for training and classification within a deep convolutional neural network. The DCNN architecture consists of multiple convolutional layers, and these extract relevant features from the sentiment feature set. The model is trained using the labeled data from the Twitter datasets and classifies new tweets based on the trained model.

The paper introduces the GloVe-DCNN model, which combines word embeddings, sentiment polarity scores, and Twitter-specific features for Twitter sentiment analysis. The model outperforms baseline approaches, demonstrating its effectiveness in accurately classifying tweet sentiment.

4 Random Forest

Random Forest is a highly regarded and versatile machine learning technique that has gained popularity due to its strong performance. It utilizes an ensemble of trees to address a wide range of classification and regression tasks, making it a powerful and reliable method in the field of machine learning [8].

The key advantages of Random forest are:

- Out-of-Bag Error Estimation: Random Forests use the out-of-bag samples, which are not used during training, to estimate the model's performance. This estimation acts as a validation set without the need for a separate validation set, making the training process more efficient
- Handling High-Dimensional Data: Random Forests are effective in managing datasets with a high dimensionality, where the number of features surpasses the number of samples. This ability allows them to effectively capture and model complex relationships between features and target variables in such scenarios
- High Accuracy: By aggregating the predictions of multiple trees, Random Forests can achieve high accuracy for predictions. The voting mechanism also helps to reduce overfitting and improves generalization

This paper will discuss three aspects of random forest: general application, comparison with other methods, and multilingual performance.

4.1 General Application

Fauzi focuses on sentiment analysis as a text classification problem, specifically in the Indonesian language [9]. The objective is to explore the use of the Random Forest algorithm for sentiment classification and investigate the impact of different term weighting methods on the performance of the sentiment analysis system.

The experiment utilizes a dataset of 386 reviews obtained from FemaleDaily, a platform where users share their opinions and experiences in Indonesian. The reviews serve as the basis for sentiment classification, with the sentiments categorized as positive or negative.

In the preprocessing stage, the reviews undergo several steps to prepare the text data for analysis. These steps include tokenization, case folding, and cleaning processes. However, stemming and filtering techniques are excluded in this study, as previous studies have shown the futility of the techniques for the Indonesian language.

In this study, a bag-of-words approach is employed for feature extraction. The unique terms obtained during preprocessing act as the features. Each review is represented as a vector, where the value of each feature indicates the presence or absence of the corresponding term in the review.

During the sentiment classification phase, the ensemble learning approach known as Random Forest, which builds upon decision tree algorithms, is employed. Multiple classification trees are grown to create a forest in the Random Forest algorithm. When classifying a new data point, each tree generates its prediction, and the final category (positive or negative) assigned to the review is determined through a majority vote.

The findings suggest that the sentiment analysis system utilizing Random Forest exhibits favorable performance, with an average Out of Bag score of 0.829. All four term weighting methods, namely Raw TF, Binary TF, TF.IDF, and Logarithmic TF, show promising outcomes. Nevertheless, the variances in scores are not substantial, indicating that the selection of term weighting method holds limited significance in Random Forest-based sentiment analysis. Additionally, it's to be noted that in another study, Al Amrani et al. found that when compared to other methods such as SVM, the random forest performed better [10]. However, this isn't a conclusive result and could be due to several factors such as the dataset.

4.2 Random Forest Comparison with Other Methods

Gupte et al. reports a comparative study of classification algorithms used in sentiment analysis [11]. In this study, they focus on four widely used algorithms: Naive Bayesian Classifier, Maximum Entropy Model, Gradient Boosting, and Ensemble of Decision Trees.

The study utilizes a dataset consisting of textual reviews, which are a common application of sentiment analysis. The dataset includes reviews where individuals express their thoughts about a product, allowing for more precise feedback on different aspects of the product. Then, Sentiment analysis engines analyze these reviews iteratively and generate output by assigning polarities, such as positive, negative, or neutral, to them.

The study evaluates the performance of four selected classification methods, namely the Naive Bayesian Classifier, Maximum Entropy Model, Gradient Boosting, and Ensemble of Decision Trees. Every algorithm is explained in terms of its fundamental theory, typical application scenarios, and benefits and drawbacks. Naïve Bayes is a baseline classification algorithm based on Bayes theorem, while Max Entropy utilizes logistic regression and features-based models. Boosted Trees combines boosting and decision trees, while Random Forest employs an ensemble of decision trees for classification tasks.

The comparison of the classification algorithms provides information on how well they perform in sentiment analysis. Naïve Bayes, despite its simple nature, exhibits reasonable efficiency and is recommended when training time and limited resources are limiting factors. Max Entropy shows improved efficiency compared to Naïve Bayes, particularly when prior results are utilized. Boosted Trees, combining regression trees and boosting, offer fast training without sacrificing accuracy, also having the ability to handle different types of predictor variables. Random Forest, an ensemble learning method, delivers high performance across various datasets and is robust to irrelevant text, making it a good generalizable method.

4.3 Multilingual Performance

Munshi et al. propose a novel approach to sentiment analysis using the Random Forest model [12]. This approach offers valuable insights for various applications, including the investigation of social problems and government decision-making, by effectively analyzing text data.

The study utilizes a Hindi tweets dataset. The dataset consists of tweets written in Hindi, which are later translated into English. This multilingual dataset enables sentiment analysis in the Hindi language, expanding the application of this research to a broader audience.

The proposed method follows several key steps. First, the dataset is visualized using the matplotlib library, fostering a better understanding of the data. Next, preprocessing techniques, such as removing usernames, stop words, and punctuation, are applied to clean the data. Afterwards, additional techniques like stemming, lemmatization, and tokenization from the Natural Language Toolkit are employed.

Feature extraction methods, such as Bag of Words and TF-IDF, are used to create features from the text, facilitating the division of the dataset into training and testing sets. The sentiment analysis model is then developed using the Random Forest Classifier technique, which combines different decision trees to enhance predictive performance.

The model presented in this study demonstrated a high accuracy score of 90.24% and an F1 score of 0.66354, indicating its effectiveness in accurately predicting and classifying the given data. Comparing these results with previous studies using different machine learning algorithms, the Random Forest approach demonstrated similar performance levels, even though most authors mainly utilized Support Vector Machine or neural network models. This suggests that Random Forest can be a competitive alternative for sentiment analysis.

5 Conclusion

This paper presents a comprehensive overview of sentiment analysis, emphasizing its background, application, and future directions. Sentiment analysis, also known as opinion mining, is a vital area of study within natural language processing, enabling the computational analysis of text to discern and understand expressed sentiment or opinion.

Three prominent sub-topics within sentiment analysis were explored in this paper: Support Vector Machines, Neural Networks, and Random Forests.

Support Vector Machines have emerged as a robust technique for sentiment classification. By finding an optimal hyperplane, SVM separates sentiment classes effectively, achieving high accuracy in sentiment prediction. SVM's ability to handle high-dimensional feature spaces makes it a popular choice in sentiment analysis tasks.

Neural Networks, including Recurrent Neural Networks and Convolutional Neural Networks, have revolutionized sentiment analysis by capturing contextual information and local patterns in text data. RNNs excel in modeling sequential dependencies and capturing long-term contextual information, while CNNs effectively extract local features through convolution operations. These deep learning models have demonstrated impressive performance in sentiment classification tasks, outperforming traditional methods in capturing complex sentiment patterns.

Random Forest, an ensemble learning method, has gained significant traction in sentiment analysis. By combining multiple decision trees, Random Forest leverages the power of collective predictions and handles noisy and high-dimensional feature spaces. Its ability to handle large datasets and capture complex interactions between features makes it a valuable tool for sentiment analysis tasks.

Looking ahead, sentiment analysis holds immense potential for further advancements. Integrating sentiment analysis with multimodal analysis, such as incorporating visual and audio cues, will enable a more comprehensive understanding of sentiment expressed across different modalities. Furthermore, advancements in sentiment analysis techniques should focus on handling nuanced sentiments, sarcasm, irony, and cultural variations to improve the accuracy and applicability of sentiment analysis in diverse domains and applications.

These approaches offer distinct methodologies for sentiment classification, each with its strengths and applications. As sentiment analysis continues to evolve, trends anticipate that exciting developments will enhance the understanding of sentiments expressed in textual data, enabling better decision-making and engagement across various domains.

References

1. Noble, W. S.: What is a support vector machine? Nature Biotechnology 24(12), 1565–1567 (2006).
2. Ahmad, M., Aftab, S., Bashir, M. S., Hameed, N., Ali, I., Nawaz, Z.: SVM Optimization for Sentiment Analysis. International Journal of Advanced Computer Science and Applications 9(4), 393–398 (2018).
3. Prastyo, P. H., Sumi, A. S., Dian, A. W., Permanasari, A. E.: Tweets Responding to the Indonesian Government's Handling of COVID-19: Sentiment Analysis Using SVM with Normalized Poly Kernel. Journal of Information Systems Engineering and Business Intelligence 6(2), 112–122 (2020).
4. Taylor, J. G.: Neural Networks and Their Applications. 1st edn. John Wiley & Sons, Inc., USA (1996).
5. Yang, P., Chen, Y.: A Survey on Sentiment Analysis by using Machine Learning Methods. 2017 IEEE 2nd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC), pp. 117–121. IEEE (2017).
6. Jianqiang, Z., Xiaolin, G., Xuejun, Z.: Deep Convolution Neural Networks for Twitter Sentiment Analysis. IEEE Access, 6, 2325 –23260 (2018).
7. Ramadhani, A. M., Goo, H. S.: Twitter sentiment analysis using deep learning methods. 2017 7th International Annual Engineering Seminar (InAES), pp. 1–4. IEEE (2017).
8. Biau, G., Scornet, E.: A Random Forest Guided Tour. TEST 25(2), 197–227 (2016).
9. Fauzi, M. A.: Random Forest Approach for Sentiment Analysis in Indonesian Language. Indonesian Journal of Electrical Engineering and Computer Science 12(1), 46–50 (2018).
10. Al Amrani, Y., Lazaar, M., El Kadiri, K. E.: Random Forest and Support Vector Machine based Hybrid Approach to Sentiment Analysis. Procedia Computer Science, 127, 511–520 (2018).
11. Gupte, A., Joshi, S., Gadgul, P., Kadam, A.: Comparative Study of Classification Algorithms used in Sentiment Analysis. International Journal of Computer Science and Information Technologies 5(5), 6261–6264 (2014).
12. Munshi, A., Sapra, S., Muthusamy, A.: A Novel Random Forest Implementation of Sentiment Analysis. International Research Journal of Engineering and Technology (IRJET) 7(6), 2821–2824 (2020).