Proceedings

An Exhaustive Comparative Study of Machine Learning Algorithms for Natural Language Processing Applications

Kanwar Mansoor Ali

**\*** Department of Mechatronics SZABIST University; engineermansoor07@yahoo.com

**Abstract:** Past few decades have witnessed an enormous research growth in the field of Natural Language Processing. In this regard, numerous Machine Learning (ML) algorithms were applied in different sub-domains of NLP such as speech recognition, Text classification, Sentiment Analysis etc. Furthermore, the performance was evaluated on diverse performance metrics. However, comparative Analysis of various ML algorithms in aforementioned field is a research area that is feasible to be addressed. Because, it may lead to an efficient direction, where future research would precisely focus on the improvement of those particular algorithms that have found to be more effective in light of previous research. Thus, this article provides a comparative Analysis regarding the application and effectiveness of different ML Algorithms in the field of NLP. Additionally, it highlights the future research direction to be adopted for enhancing the ability of Natural Language Processing domain.

**Keywords:** Classification, Machine Vision, Natural Language Processing, Machine Learning

1. Background

In this research paper, background of machine learning besides classification algorithms, and how they have evolved over time is provided. Researchers first looked into the concept of "cybernetics" in the 1940s and 1950s, which is when machine learning first emerged [1]. Then, researchers started creating machine learning algorithms in the 1960s and 1970s, including decision trees and the perceptron [2]. Machine learning made tremendous strides in the 1980s and 1990s including neural networks and support vector machines (SVMs) [3]. The most popular classification algorithms are k-nearest neighbors, decision trees, SVMs, and naive Bayes [5]. Convolutional neural networks (CNNs) and recurrent neural networks (RNNs) are examples of deep learning-based algorithms that have been the subject of extensive research in recent years for the development of novel classification methods [6].

2. Introduction

These are but a few more illustrations of the several categorization methods [7,8,9,10,11,12,13,14,15,16] that are used in the machine learning industry. The choice of algorithm depends on the particular problem and the features of the data, and each algorithm has strengths and limitations of its own. State-of-the-art performance on a variety of NLP tasks [17], [18], and [19], is one of the most important advancements in NLP research. Additionally, because of their capacity to learn contextualized word representations, transformer-based models like BERT [20] have grown in popularity. Sentiment analysis, another component of NLP, seeks to ascertain the polarity of a text's sentiments. Convolutional neural networks (CNNs) [23] and long-short-term memory (LSTM) networks [24] are two examples of deep learning-based models that have been demonstrated to be successful for sentiment analysis. Numerous strategies, such as rule-based systems [25], statistical models [26], and deep learning-based models [27], have been put forth for NER. In contrast to abstractive summarization, which creates a new summary by paraphrasing the original text, extractive summarization includes choosing the most crucial lines or phrases from the original text [28]. Question-answering [29], natural language creation [30], and discourse systems [31] are other crucial NLP study fields. In conclusion, NLP is a fast-expanding field that encompasses a variety of methods and uses. Numerous NLP tasks have significantly improved as a result of the adoption of deep learning techniques, and it is anticipated that research in this area will continue to grow quickly in the years to come. The objective of this work is to perform a thorough comparative analysis of machine learning (ML) techniques used in different sub-domains of Natural Language Processing (NLP), such as speech recognition, text classification, and sentiment analysis. The study seeks to assess the performance of these algorithms using various metrics and benchmarks in order to determine their strengths and weaknesses in distinct NLP jobs. This will provide valuable insights into the most efficient ML algorithms for certain applications within NLP. In addition, the study aims to suggest future research paths to improve the capabilities of NLP systems, using the results of the comparative analysis. This will contribute to the progress of NLP research and provide guidance for the development of more effective NLP technologies.

3. Literature Review

This section encompasses the research that has been performed on the subject discussed earlier. Thus, articles related to the application and comparison of diverse classification algorithms on the fields of Natural language processing has been reviewed respectively. Support Vector Machines (SVMs) are a text categorization technique (Tables-1&2) presented by T. Joachims [48]. The author starts out by going through the fundamental ideas of SVMs and how they apply to text categorization. The independence premise in information retrieval is looked at by D. D. Lewis [49] in relation to the Naive Bayes classification method (Table-1&2). Popular text categorization algorithms in natural language processing include Naive Bayes. In this landmark study written by Sepp Hochreiter and Jürgen Schmidhuber [50] and released in 1997, LSTM was proposed (Table-1&5). Since then, LSTM has developed into one of the most popular and effective RNN designs in a number of industries, including speech recognition, image captioning, and natural language processing. A unique method for word embedding creation that utilizes sub word data is presented by Bojanowski et al. [51]. Word embeddings are created using merely the words themselves in traditional approaches like word2vec and GloVe, without any internal word structure. A new method for enhancing automatic speech recognition using deep recurrent neural networks (RNNs) ((Tables 1&4) is presented in the paper by A. Graves, A.-r. Mohamed, and G. Hinton [52]. The Long Short-Term Memory (LSTM) RNN, which the authors suggest as a replacement for the conventional RNN. A comparison of different sentiment analysis machine learning methods (Tables 1&2) is given by N. Almatarneh et al. (2019) [59]. Comparison research on sentiment analysis of Twitter data (Tables 1&2) using multiple categorization algorithms is presented by H. N. Nguyen, N. D. Vo, and V. M. Ha [60]. The authors evaluated similar research in the areas of machine learning algorithms and sentiment analysis. In the work "Sentiment analysis with machine learning algorithms" by H. B. Demir [61], the effectiveness of machine learning algorithms for sentiment analysis tasks is compared (Tables 1&2).

4. Research Synthesis

**Table 1.** Research Synthesis

|  |  |  |  |
| --- | --- | --- | --- |
| **Reference** | **Research Topic** | **Methodology** | **Key Findings** |
| [48] | Text categorization with Support Vector Machines | Experimental study | SVMs can effectively learn from text data with many features, outperforming other methods |
| [49] | Naive Bayes for information retrieval | Conceptual analysis | Naive Bayes' assumption of independence is a reasonable approximation for text classification |
| [50] | Long short-term memory for neural networks | Theoretical analysis and experimental study | LSTMs can effectively learn and remember long-term dependencies in sequential data |
| [51] | Enriching word vectors with sub word information | Experimental study | Subword information can improve the quality of word embeddings and enable word representations for rare or unseen words |
| [52] | Speech recognition with deep recurrent neural networks | Experimental study | Deep RNNs can achieve state-of-the-art performance on speech recognition tasks |
| [53] | LIBSVM: A library for support vector machines | Technical report | LIBSVM is an efficient and effective implementation of SVMs |
| [59] | Comparative study of machine learning algorithms for sentiment analysis | Experimental study | SVMs and Random Forest perform better than other methods for sentiment analysis on Twitter data |
| [60] | Comparative study on sentiment analysis of Twitter data using various classification algorithms | Experimental study | Naive Bayes and SVMs perform better than other methods for sentiment analysis on Twitter data |
| [61] | Sentiment analysis with machine learning algorithms | Experimental study | SVMs outperform Naive Bayes and Decision Trees for sentiment analysis on hotel reviews |

5. Performance Metrics

The proportion of instances that were correctly categorized over all instances is known as accuracy. [54] Precision: the ratio of genuine positives (positives that were correctly identified) to all anticipated positives (true positives plus false positives). [55]. The ratio of true positives to all actual positives (true positives plus false negatives) is known as recall (or sensitivity). [55]. F1-score is a harmonic mean of recall and precision that equally weights the two metrics. [56]. The true positive rate (TPR) and false positive rate (FPR) trade-off for various categorization thresholds is measured by the area under the receiver operating characteristic (AUC-ROC) curve. [57]. The average of the squared discrepancies between the predicted and actual values is the mean squared error (MSE), which is frequently utilized in regression situations. [58] Root mean squared error (RMSE): This error metric has the same units as the target variable because it is the square root of MSE. [58]

6. Comparative Analysis

**Table 2.** Comparative Analysis

| S. No | References | Methods | Performance Metrics | | | |
| --- | --- | --- | --- | --- | --- | --- |
| Precision | Accuracy | Recall | F1-Score |
| 1. | [48] | Radicchio's |  | 79.9 | - | - |
| K-NN, K=30 | 97.3 | 82.3 | - | - |
| Naive Bayes | 95.9 | 72.0 | - | - |
| SVM (rbf) | 98.5 | 86.4 | - | - |
| SVM (polynomial) | 98.5 | 86.0 | - | - |
| C4.5 | 96.1 | 79.4 | - | - |
| 2. | [49] | Naive Bayes with textblod | - | - | - | 76 |
| Naive Bayes with sentiwordnet | - | - | - | 54.75 |
| Naive Bayes with WSD | - | - | - | 79.10 |
| SVM with Text blob | - | - | - | 62.67 |
| SVM with sentiwordnet | - | - | - | 53.33 |
| SVM with WSD | - | - | - | 62.33 |
| 3a. | [59]  Product Review dataset | Naive bayes | 0.796 | 0.801 | 0.801 | 0.794 |
| SVM | 0.868 | 0.872 | 0.872 | 0.868 |
| KNN | 0.741 | 0.76 | 0.76 | 0.734 |
| Decision Tree | 0.763 | 0.774 | 0.774 | 0.76 |
| Random Forest | 0.823 | 0.828 | 0.828 | 0.819 |
| MLP | 0.838 | 0.843 | 0.843 | 0.837 |
| CNN | 0.844 | 0.846 | 0.846 | 0.843 |
| 3b. | [59]  Movie Review Dataset | Naive bayes | 0.753 | 0.748 | 0.748 | 0.743 |
| SVM | 0.859 | 0.856 | 0.856 | 0.855 |
| KNN | 0.706 | 0.708 | 0.708 | 0.705 |
| Decision Tree | 0.74 | 0.739 | 0.739 | 0.737 |
| Random Forest | 0.796 | 0.798 | 0.798 | 0.795 |
| MLP | 0.821 | 0.822 | 0.822 | 0.82 |
| CNN | 0.828 | 0.829 | 0.829 | 0.827 |
| 4a. | [61]  IMDb movie review dataset | Naive bayes | - | 0.735 | - | - |
| Decision Tree | - | 0.723 | - | - |
| Random Forest | - | 0.803 | - | - |
| SVM | - | 0.816 | - | - |
| 5. | [60]  Sentiment Analysis of twitter dataset | Naive Bayes | 0.787 | 0.778 | 0.768 | 0.777 |
|  |  |  |  |  |
| Decision Tree | 0.716 | 0.724 | 0.726 | 0.721 |
| Random Forest | 0.782 | 0.776 | 0.772 | 0.777 |
| SVM | 0.779 | 0.773 | 0.769 | 0.774 |
| K-NN | 0.717 | 0.712 | 0.706 | 0.711 |

**Table 3.** Comparative Analysis of SVM Libraries

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Reference** | **Library** | **Training Time** | **Testing Time** | **Diabetes Accuracy** | **Heart-scale Accuracy** | **Synthetic Dataset Accuracy** |
| [53] | LIBSVM | 0.54 | 0.003 | 77.23% | 86.04% | 86.6% |
| SVMrank | 0.29 | 0.009 | 76.31% | 84.31% | - |
| SVMTorch | 2.41 | 0.002 | 77.23% | 85.31% | - |
| SVMperf | 0.02 | 0.002 | 75.65% | 84.96% | - |
| SVMLight | 0.16 | 0.004 | 77.23% | 86.38% | - |
| SVMlin | 3.36 | 0.009 | 77.23% | 85.85% | - |

Table 4. WER (Word Error Rate) of Different Deep RNN Models

|  |  |  |
| --- | --- | --- |
| **Reference** | **MODEL** | **TEST SET WORD ERROR RATE (WER)** |
| [52] | Standard HMM | 28.2% |
| Standard DNN | 26.2% |
| Standard RNN | 23.7% |
| Standard LSTM | 20.7% |
| Standard BLSTM | 19.7% |
| Standard SdA | 17.3% |
| LSTM pre-training | 19.4% |
| BLSTM pre-training | 18.6% |
| SdA pre-training | 16.0% |
| Tandem features HMM | 19.7% |
| Tandem features LSTM | 16.0% |

Table 5. Comparative Analysis of LSTM and SRC on Different Tasks

|  |  |  |  |
| --- | --- | --- | --- |
| **Reference** | **TASK** | **LSTM** | **SRC** |
| [50] | Sequence copying (MSE) | 1.35e-5 | 0.054 |
| Temporal order classification (MSE) | 0.003 | 0.105 |
| Predicting chaotic time series (MSE) | 0.001 | 0.239 |
| Speech recognition (PER) | 14.6 | 41.3 |

Table 6. Comparative Analysis of Different Tasks

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Reference** | **MODEL** | **CBOW** | **SKIP-Gram** | **Fast Text** |
| [5] | Embedding Size | 300 | 300 | 300 |
| Word vectors trained | 2.5B | 6B | 2.5B |
| Vocabulary size | 2.5M | 2.5M | 2.5M |
| Test set accuracy | 65.3% | 67% | 69% |
| Out-of-vocabulary accuracy | 38.8% | 41.7% | 46.1% |
| Semantic-Syntactic word analogy accuracy | 44% | 53.0% | 56.0% |

7. RESULTS

The study conducted a comparison between Support Vector Machines (SVMs) and other text categorization techniques such as k-Nearest Neighbors and Naive Bayes. Support Vector Machines (SVMs), especially when using the radial basis function (rbf) kernel, demonstrated superior performance in terms of accuracy and training time compared to other methods. Naive Bayes demonstrated exceptional performance in Word-sense-disambiguation (WSD) when used with Textblob, however FastText exhibited higher performance across multiple measures. Sparse representation-based categorization achieved superior performance compared to LSTM in natural language processing (NLP) applications. The efficiency of the Tandem features LSTM model was demonstrated by the newly introduced performance metric, Word error rate (WER). LIBSVM and SVM Light demonstrated favorable outcomes when compared to other libraries. Various classification approaches shown varying levels of success when applied to sentiment analysis and product review datasets, with Support Vector Machines (SVM) frequently achieving the highest rankings.

8. CONCLUSION

From the aforementioned discussion it is concluded that ML technology has been evolving since its inception. Thus, novice techniques are massively being introduced in research realm. Subsequently, the technology of NLP is growing at a high rate corresponding the rise in the research domain of Machine learning and deep learning techniques.

From comparative Analysis it can be elicited that, effectiveness and accuracy of several classification algorithms varies with the application. Furthermore, it is also influenced by other factors like training rate and types of modification in fundamental architecture of classification algorithms. Conclusively, decision on the feasibility of any classification method just by considering its effectiveness in other application wouldn’t be a right approach.

9. FUTURE DIRECTION

As was already noted, the discipline of NLP is expanding really quickly. As a result, AI is revolutionizing an increasing number of fields. With the expansion of international communication, NLP systems must be able to handle different languages. To create multilingual models and systems that can precisely comprehend and process content in several languages, research is required.

10. Acknowledgment

I would like to take this opportunity to express my sincere gratitude and appreciation to my teacher, Dr. Talha Ahmed, for their guidance and support throughout this research.

|  |  |
| --- | --- |
|  |  |

References

1. N. Wiener, "Cybernetics: Or Control and Communication in the Animal and the Machine," MIT Press, 1965.
2. F. Rosenblatt, "The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain," Psychological Review, vol. 65, no. 6, pp. 386-408, 1958.
3. V. Vapnik, "Statistical Learning Theory," John Wiley & Sons, 1998.
4. Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol. 521, no. 7553, pp. 436-444, 2015.
5. J. Han, M. Kamber, and J. Pei, "Data Mining: Concepts and Techniques," 3rd ed., Morgan Kaufmann Publishers, 2012.
6. A. Krizhevsky, I. Sutskever, and G. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," in Proceedings of the 25th International Conference on Neural Information Processing Systems (NIPS), 2012.
7. Hosmer Jr, D.W. and Lemeshow, S., Applied logistic regression. John Wiley & Sons, 2004.
8. Breiman, L., Friedman, J., Stone, C.J. and Olshen, R.A., Classification and regression trees. CRC press, 1984.
9. Rish, I., “An empirical study of the naive Bayes classifier,” in Proceedings of the IJCAI-01 workshop on empirical methods in artificial intelligence, 2001.
10. Cortes, C. and Vapnik, V., “Support-vector networks,” Machine learning, vol. 20, no. 3, pp. 273-297, 1995.
11. Breiman, L., “Random forests,” Machine learning, vol. 45, no. 1, pp. 5-32, 2001.
12. Cover, T. and Hart, P., “Nearest neighbor pattern classification,” IEEE transactions on information theory, vol. 13, no. 1, pp. 21-27, 1967.
13. Rumelhart, D.E., Hinton, G.E. and Williams, R.J., “Learning representations by back-propagating errors,” Nature, vol. 323, no. 6088, pp. 533-536, 1986.
14. Chen, T. and Guestrin, C., “XGBoost: A scalable tree boosting system,” in Proceedings of the 22nd ACM SIGKDD international conference on Knowledge discovery and data mining, pp. 785-794, 2016.
15. Fisher, R.A., “The use of multiple measurements in taxonomic problems,” Annals of Eugenics, vol. 7, no. 2, pp. 179-188, 1936.
16. Zhou, Z.H., “Ensemble methods: foundations and algorithms,” CRC press, 2012.
17. Y. Kim, “Convolutional neural networks for sentence classification,” in Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing, 2014, pp. 1746–1751.
18. T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, and J. Dean, “Distributed representations of words and phrases and their compositionality,” in Advances in Neural Information Processing Systems, 2013, pp. 3111–3119.
19. X. Zhang, J. Zhao, and Y. LeCun, “Character-level convolutional networks for text classification,” in Advances in Neural Information Processing Systems, 2015, pp. 649–657.
20. J. Devlin, M. Chang, K. Lee, and K. Toutanova, “BERT: Pre-training of deep bidirectional transformers for language understanding,” in Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, 2019, pp. 4171–4186.
21. F. Huang, L. Wu, and J. Chou, "A Hybrid Approach to Rule-Based Named Entity Recognition in Chinese Clinical Texts," Journal of Medical Systems, vol. 41, no. 11, p. 183, 2017.
22. H. Zhang, M. Li, and Y. Li, "A Brief Survey of Deep Learning," IEEE Transactions on Neural Networks and Learning Systems, vol. 28, no. 10, pp. 2354-2364, Oct. 2017.
23. Y. Bengio, "Deep Learning of Representations for Unsupervised and Transfer Learning," in Proceedings of ICML Workshop on Unsupervised and Transfer Learning, 2012, pp. 17-36.
24. T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient Estimation of Word Representations in Vector Space," in Proceedings of International Conference on Learning Representations, 2013.
25. J. Pennington, R. Socher, and C. D. Manning, "GloVe: Global Vectors for Word Representation," in Proceedings of Conference on Empirical Methods in Natural Language Processing, 2014, pp. 1532-1543.
26. S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," Neural Computation, vol. 9, no. 8, pp. 1735-1780, Nov. 1997.
27. D. Jurafsky and J. H. Martin, Speech and Language Processing, 2nd ed. Upper Saddle River, NJ: Prentice Hall, 2009.
28. Y. LeCun, Y. Bengio, and G. Hinton, "Deep Learning," Nature, vol. 521, no. 7553, pp. 436-444, May 2015.
29. N. Jindal and B. Liu, "Opinion Spam and Analysis," in Proceedings of ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2008, pp. 1-9.
30. D. D. Lewis, "Evaluation of Text Categorization," in Proceedings of Speech and Natural Language Workshop, 1992, pp. 1-6.
31. K. Lee, J. Lee, and H. Lee, "On the Effectiveness of Simple Preprocessing Methods in the Text Classification Task," Expert Systems with Applications, vol. 40, no. 15, pp. 6136-6148, 2013.
32. M. R. Dixon and B. Bhushan, "History and future of industrial machine vision and inspection," Journal of Microscopy, vol. 208, no. 3, pp. 177-188, Dec. 2002.
33. T. J. Fuchs, C. Wildermuth, S. W. Martirosian, C. M. Meyer, M. A. Holzwarth, and B. A. Kiefer, "MRI-based classification of breast tumors using a support vector machine and a k-nearest neighbor classifier," Academic Radiology, vol. 16, no. 8,
34. T. Kadir and M. Brady, "Scale, saliency and image description," International Journal of Computer Vision, vol. 45, no. 2, pp. 83-105, 2001.
35. D. G. Lowe, "Distinctive image features from scale-invariant keypoints," International Journal of Computer Vision, vol. 60, no. 2, pp. 91-110, 2004.
36. L. Fei-Fei, R. Fergus, and P. Perona, "Learning generative visual models from few training examples: An incremental Bayesian approach tested on 101 object categories," Computer Vision and Image Understanding, vol. 106, no. 1, pp. 59-70, 2007.
37. ] S. Gould, J. Rodgers, D. Cohen, G. Elidan, and D. Koller, "Multi-class segmentation with relative location prior," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2008, pp. 1-8.
38. J. Shotton, J. Winn, C. Rother, and A. Criminisi, "TextonBoost for image understanding: Multi-class object recognition and segmentation by jointly modeling texture, layout, and context," International Journal of Computer Vision, vol. 81, no. 1, pp. 2-23, 2009.
39. A. Krizhevsky, I. Sutskever, and G. Hinton, "Imagenet Classification with Deep Convolutional Neural Networks," Advances in Neural Information Processing Systems, pp. 1097-1105, 2012.
40. K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 770-778.
41. Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol. 521, no. 7553, pp. 436-444, 2015.
42. A. Krizhevsky, "Learning Multiple Layers of Features from Tiny Images," University of Toronto, 2009.
43. J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "ImageNet: A Large-Scale Hierarchical Image Database," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2009, pp. 248-255.
44. K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," arXiv preprint arXiv:1409.1556, 2014.
45. O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, and L. Fei-Fei, "ImageNet Large Scale Visual Recognition Challenge," International Journal of Computer Vision, vol. 115, no. 3, pp. 211-252, 2015.
46. J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 343
47. S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," Advances in Neural Information Processing Systems, pp. 91-99, 2015.
48. T. Joachims, "Text categorization with Support Vector Machines: Learning with many relevant features," Journal of Machine Learning Research, vol. 2, pp. 137-142, 2001.
49. D. D. Lewis, "Naive Bayes at forty: The independence assumption in information retrieval," Machine Learning, vol. 52, no. 1, pp. 4-15, 2003.
50. S. Hochreiter and J. Schmidhuber, "Long short-term memory," Neural Computation, vol. 9, no. 8, pp. 1735-1780, 1997
51. P. Bojanowski, E. Grave, A. Joulin, and T. Mikolov, "Enriching word vectors with subword information," Transactions of the Association for Computational Linguistics, vol. 5, pp. 135-146, 2017
52. A. Graves, A.-r. Mohamed, and G. Hinton, "Speech recognition with deep recurrent neural networks," IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 6645-6649, 2013
53. C. C. Chang and C. J. Lin, "LIBSVM: A library for support vector machines," ACM Transactions on Intelligent Systems and Technology, vol. 2, no. 3, pp. 1-27, 2011
54. J. Han and M. Kamber, "Data Mining: Concepts and Techniques," 3rd ed. Morgan Kaufmann, 2011.
55. L H. Witten, E. Frank, and M. A. Hall, "Data Mining: Practical Machine Learning Tools and Techniques," 3rd ed. Morgan Kaufmann, 2011.
56. C. Manning, P. Raghavan, and H. Schütze, "Introduction to Information Retrieval," Cambridge University Press, 2008.
57. TFawcett, "An introduction to ROC analysis," Pattern Recognition Letters, vol. 27, no. 8, pp. 861-874, 2006.
58. G. James, D. Witten, T. Hastie, and R. Tibshirani, "An Introduction to Statistical Learning," Springer, 2013.
59. NAlmatarneh, M. K. Hasan, A. M. Almomani, and Y. Alsmirat, "A comparative study of machine learning algorithms for sentiment analysis," Journal of Information Science, vol. 45
60. H. N. Nguyen, N. D. Vo, and V. M. Ha, "A comparative study on sentiment analysis of Twitter data using various classification algorithms," Proceedings of the 9th International Conference on Knowledge and Systems Engineering, pp. 277-283, 2017
61. H. B. Demir, "Sentiment analysis with machine learning algorithms," Journal of King Saud University-Computer and Information Sciences, vol. 30, no. 3, pp. 330-335, 2018