engineering proceedings

Proceeding Paper

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

Kanwar Mansoor Ali [1], Talha Ahmed Khan [2,3,4,*], Syed Mubashir Ali [4], Asif Aziz [4], Sharfuddin Ahmed Khan [5] and Sadique Ahmad [6]

1 Department Computer Sciences, SZABIST, Karachi 75850, Pakistan; engineermansoor07@yahoo.com
2 Malaysian Institute of Information Technology (MIIT), Universiti Kuala Lumpur, Kuala Lumpur 50250, Malaysia
3 Faculty of Computing, Multimedia University (MMU), Cyber Jaya 63100, Malaysia
4 Department of Computer Science, Bahria University, Lahore 54600, Pakistan; syedmubashir.bulc@bahria.edu.pk (S.M.A.); asifaziz.bukc@bahria.edu.pk (A.A.)
5 Industrial System Engineering, University of Regina, Regina, SK S4S 0A2, Canada; sharfuddin.khan@uregina.ca
6 EIAS: Data Science Laboratory, College of Computer and Information Sciences, Prince Sultan University, Riyadh 12435, Saudi Arabia; saahmad@psu.edu.sa
* Correspondence: talhaahmedkhan.bukc@bahria.edu.pk
† Presented at the 1st International Conference on Industrial, Manufacturing, and Process Engineering (ICIMP-2024), Regina, Canada, 27–29 June 2024.

Abstract: The past few decades have witnessed an enormous research growth in the field of natural language processing. In this regard, numerous machine learning (ML) algorithms have been applied in different sub-domains of NLP such as speech recognition, text classification, sentiment analysis, etc. Furthermore, their performances have been evaluated using diverse performance metrics. However, a comparative analysis of various ML algorithms in the aforementioned field is a feasible research area to explore. This may efficiently guide future research to precisely focus on the improvement of those particular algorithms that have been found to be more effective based on 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 the natural language processing domain.

Keywords: classification; machine vision; natural language processing; machine learning

check for updates

Citation: Ali, K.M.; Ahmed Khan, T.; Ali, S.M.; Aziz, A.; Khan, S.A.; Ahmad, S. An Exhaustive Comparative Study of Machine Learning Algorithms for Natural Language Processing Applications. *Eng. Proc.* **2024**, *76*, 79. https://doi.org/10.3390/engproc2024076079

Academic Editors: Golam Kabir, Mohammad Khondoker and Hussameldin Ibrahim

Published: 13 November 2024

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Background

In this research paper, a background of machine learning, in addition to classification algorithms and how they have evolved over time, is provided. Researchers first explored 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 [4,5]. 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 [7–16] 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 [17–19] is one of the most important advancements brought by 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 [21,22]. 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. MRI based classification has already been completed using support vector machines and kNN [32–36]. 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 [37–40]. This 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. Artificial neural networks with convolution have also been applied for the segmentation of semantic analysis [41–44]. 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. Convolutional Neural Networks (CNNs) have also been applied for the segmentation of semantic analysis. Text categorization has been performed by using the support vector machines [45,46].

3. Literature Review

This section encompasses the research that has been performed on the subject discussed earlier. Thus, studies related to the application and comparison of diverse classification algorithms in the fields of natural language processing are reviewed [47]. The support vector machine (SVM) is a text categorization technique presented by authors. The author starts out by discussing the fundamental ideas of SVMs and how they are applied to text categorization. The independence assumption in information retrieval was explored by authors [48] in relation to the naive Bayes classification methods. Popular text categorization algorithms in natural language processing include naive Bayes. In a landmark study [49] that was released in 1997, LSTM was demonstrated in Tables 1 and 2. Moreover, SVM optimization and convergence issues were resolved. 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 subword data was presented in past research [50]. 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). Comparative analysis was presented in the previous research for XML retrieval and text classification [51]. Big data mining tools have been elaborated to group and cluster the data using machine learning based approaches. Statistical learning and ROC analysis have been studied [52]. The long short-term memory (LSTM) RNN was suggested by the authors as a replacement for the conventional RNN. A comparison of different sentiment analysis machine learning methods has been discussed in the research synthesis [53]. Comparative research on the sentiment analysis of Twitter data using multiple categorization algorithms was presented by the researchers using SVM and KNN [54]. The authors evaluated similar research in the areas of machine learning algorithms and sentiment analysis. In the work "Sentiment analysis with machine learning algorithms", the effectiveness of machine learning algorithms for sentiment analysis tasks was compared.

4. Research Synthesis

Research synthesis was completed and it was found that the Support vector machine and Naïve Bayes were found to be more effective and competent for categorization. Long short-term memory can also learn and remember effectively for neural networks.

Table 1. Research synthesis.

Reference	Research Topic	Methodology	Key Findings
[45]	Text categorization with support vector machines	Experimental study	SVMs can effectively learn from text data with many features, outperforming other methods
[46]	Naive Bayes for information retrieval	Conceptual analysis	Naive Bayes' assumption of independence is a reasonable approximation for text classification
[47]	Enriching word vectors with subword information	Experimental study	Subword information can improve the quality of word embeddings and enable word representations for rare or unseen words
[48]	Speech recognition with deep recurrent neural networks	Experimental study	Deep RNNs can achieve state-of-the-art performance on speech recognition tasks
[49]	LIBSVM	Technical report	LIBSVM is an efficient and effective implementation of SVMs
[50]	Comparative study of machine learning algorithms for sentiment analysis, Data Mining	Experimental study	SVMs and Random Forest perform better than other methods for sentiment analysis on Twitter data
[51]	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
[52]	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 refers to the ratio of genuine positives (positives that were correctly identified) to all anticipated positives (true positives plus false positives). The ratio of true positives to all actual positives (true positives plus false negatives) is known as recall (or sensitivity) [55]. F1-score is the harmonic mean of recall and precision, which equally weighs the two metrics. 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. 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. Root mean squared error (RMSE) is an error metric that has the same units as the target variable because it is the square root of the MSE.

6. Comparative Analysis

The comparative analysis has been demonstrated in Table 3. Table 3 showed the precision, accuracy, recall and F1-score for KNN, Naïve Bayes, SVM (RBF) and other competent algorithms.

Table 2. Comparative analysis of LSTM and SRC on different tasks.

Reference	TASK	LSTM	SRC
[51]	Sequence copying (MSE)	1.35×10^{-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 3. Comparative analysis.

S. No.	Reference	Methods	Performance Metrics			
			Precision	Accuracy	Recall	F1-Score
1.	[46]	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.	[47]	Naive Bayes with TextBlob	-	-	-	76
		Naive Bayes with sentiwordnet	-	-	-	54.75
		Naive Bayes with WSD	-	-	-	79.10
		SVM with TextBlob	-	-	-	62.67
		SVM with sentiwordnet	-	-	-	53.33
		SVM with WSD	-	-	-	62.33
3.	[48] 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
4.	[49] 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
5.	[50]	Naive Bayes	-	0.735	-	-
		Decision Tree	-	0.723	-	-
		Random Forest	-	0.803	-	-
		SVM	-	0.816	-	-

7. Results

Tables 4 and 5 represented 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 in the LSTM model was demonstrated by the newly introduced performance metric, word error rate (WER). LIBSVM and SVMLight demonstrated favorable outcomes when compared to other libraries. Various classification approaches showed varying levels of success when applied to sentiment analysis and product review datasets, with support vector machines (SVM) frequently achieving the highest rankings.

Table 4. Comparative analysis of SVM libraries.

Reference	Library	Training Time	Testing Time	Diabetes Accuracy	Heart-Scale Accuracy	Synthetic Dataset Accuracy
[49]	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 5. WER (Word Error Rate) of different deep RNN models.

Reference	MODEL	TEST SET WORD ERROR RATE (WER)
[50]	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%

8. Conclusions

From the aforementioned discussion, it is concluded that ML technology has been evolving since its inception. Thus, novel techniques are being substantially introduced in the research realm. Subsequently, the technology of NLP is growing rapidly, corresponding to the rise in research on machine learning and deep learning techniques.

From our comparative analysis, it can be elicited that the effectiveness and accuracy of several classification algorithms varies with the application. Furthermore, they are also influenced by other factors like training rate and types of modification in the fundamental architecture of classification algorithms. Ultimately, deciding on the feasibility of any classification method solely based on its effectiveness in other applications would not be the right approach.

9. Future Direction

As was already noted, the discipline of NLP is expanding rapidly. 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, further research is required.

Author Contributions: Conceptualization, K.M.A.; methodology, T.A.K.; software, S.M.A.; validation, A.A.; formal analysis, S.A.; investigation, K.M.A.; resources, S.A.K. and T.A.K.; data curation, A.A.; writing—original draft preparation, K.M.A.; writing—review and editing, S.A.K. and T.A.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: No new data were created.

Acknowledgments: I would like to take this opportunity to express my sincere gratitude and appreciation to my supervisor, Talha Ahmed Khan, for his guidance and support throughout this research.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Wiener, N. *Cybernetics: Or Control and Communication in the Animal and the Machine*; MIT Press: Cambridge, MA, USA, 1965.
2. Rosenblatt, F. The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain. *Psychol. Rev.* **1958**, *65*, 386–408. [CrossRef] [PubMed]
3. Vapnik, V. *Statistical Learning Theory*; John Wiley & Sons: Hoboken, NJ, USA, 1998.
4. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [CrossRef] [PubMed]
5. Han, J.; Kamber, M.; Pei, J. *Data Mining: Concepts and Techniques*, 3rd ed.; Morgan Kaufmann Publishers: Burlington, MA, USA, 2012.
6. Krizhevsky, A.; Sutskever, I.; Hinton, G. ImageNet Classification with Deep Convolutional Neural Networks. In Proceedings of the 25th International Conference on Neural Information Processing Systems (NIPS), Red Hook, NY, USA, 3–6 December 2012.
7. Hosmer, D.W., Jr.; Lemeshow, S. *Applied Logistic Regression*; John Wiley & Sons: Hoboken, NJ, USA, 2004.
8. Breiman, L.; Friedman, J.; Stone, C.J.; Olshen, R.A. *Classification and Regression Trees*; CRC Press: Boca Raton, FL, USA, 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, Seattle, WA, USA, 4–10 August 2001.
10. Cortes, C.; Vapnik, V. Support-vector networks. *Mach. Learn.* **1995**, *20*, 273–297. [CrossRef]
11. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
12. Cover, T.; Hart, P. Nearest neighbor pattern classification. *IEEE Trans. Inf. Theory* **1967**, *13*, 21–27. [CrossRef]
13. Rumelhart, D.E.; Hinton, G.E.; Williams, R.J. Learning representations by back-propagating errors. *Nature* **1986**, *323*, 533–536. [CrossRef]
14. Chen, T.; Guestrin, C. XGBoost: A scalable tree boosting system. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; pp. 785–794.
15. Fisher, R.A. The use of multiple measurements in taxonomic problems. *Ann. Eugen.* **1936**, *7*, 179–188. [CrossRef]
16. Zhou, Z.H. *Ensemble Methods: Foundations and Algorithms*; CRC Press: Boca Raton, FL, USA, 2012.
17. Kim, Y. Convolutional neural networks for sentence classification. In Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing, Doha, Qatar, 25–29 October 2014; pp. 1746–1751.
18. Mikolov, T.; Sutskever, I.; Chen, K.; Corrado, G.S.; Dean, J. Distributed representations of words and phrases and their compositionality. *Adv. Neural Inf. Process. Syst.* **2013**, *26*, 3111–3119.
19. Zhang, X.; Zhao, J.; LeCun, Y. Character-level convolutional networks for text classification. *Adv. Neural Inf. Process. Syst.* **2015**, *28*, 649–657.

20. Devlin, J.; Chang, M.; Lee, K.; Toutanova, K. 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, Minneapolis, MN, USA, 2–7 June 2019; pp. 4171–4186.
21. Huang, F.; Wu, L.; Chou, J. A Hybrid Approach to Rule-Based Named Entity Recognition in Chinese Clinical Texts. *J. Med. Syst.* **2017**, *41*, 183.
22. Zhang, H.; Li, M.; Li, Y. A Brief Survey of Deep Learning. *IEEE Trans. Neural Netw. Learn. Syst.* **2017**, *28*, 2354–2364.
23. Bengio, Y. Deep Learning of Representations for Unsupervised and Transfer Learning. In Proceedings of the ICML Workshop on Unsupervised and Transfer Learning, Bellevue, WA, USA, 2 July 2012; pp. 17–36.
24. Mikolov, T.; Chen, K.; Corrado, G.; Dean, J. Efficient Estimation of Word Representations in Vector Space. In Proceedings of the International Conference on Learning Representations, Scottsdale, AZ, USA, 2–4 May 2013.
25. Pennington, J.; Socher, R.; Manning, C.D. GloVe: Global Vectors for Word Representation. In Proceedings of the Conference on Empirical Methods in Natural Language Processing, Doha, Qatar, 25–29 October 2014; pp. 1532–1543.
26. Hochreiter, S.; Schmidhuber, J. Long Short-Term Memory. *Neural Comput.* **1997**, *9*, 1735–1780. [CrossRef] [PubMed]
27. Jurafsky, D.; Martin, J.H. *Speech and Language Processing*, 2nd ed.; Prentice Hall: Upper Saddle River, NJ, USA, 2009.
28. Jindal, N.; Liu, B. Opinion Spam and Analysis. In Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Las Vegas, NV, USA, 22 February 2008; pp. 1–9.
29. Lewis, D.D. Evaluation of Text Categorization. In Proceedings of the Speech and Natural Language Workshop, Harriman, NY, USA, 23–26 February 1992; pp. 1–6.
30. Lee, K.; Lee, J.; Lee, H. On the Effectiveness of Simple Preprocessing Methods in the Text Classification Task. *Expert Syst. Appl.* **2013**, *40*, 6136–6148.
31. Dixon, M.R.; Bhushan, B. History and future of industrial machine vision and inspection. *J. Microsc.* **2002**, *208*, 177–188.
32. Hussain, L.; Huang, P.; Nguyen, T.; Lone, K.J.; Ali, A.; Khan, M.S.; Li, H.; Suh, D.Y.; Duong, T.Q. Machine learning classification of texture features of MRI breast tumor and peri-tumor of combined pre- and early treatment predicts pathologic complete response. *BioMed. Eng. OnLine* **2021**, *20*, 63. [CrossRef]
33. Kadir, T.; Brady, M. Scale, saliency and image description. *Int. J. Comput. Vis.* **2001**, *45*, 83–105. [CrossRef]
34. Lowe, D.G. Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vis.* **2004**, *60*, 91–110. [CrossRef]
35. Fei-Fei, L.; Fergus, R.; Perona, P. Learning generative visual models from few training examples: An incremental Bayesian approach tested on 101 object categories. *Comput. Vis. Image Underst.* **2007**, *106*, 59–70. [CrossRef]
36. Gould, S.; Rodgers, J.; Cohen, D.; Elidan, G.; Koller, D. Multi-class segmentation with relative location prior. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Anchorage, AK, USA, 23–28 June 2008; pp. 1–8.
37. Shotton, J.; Winn, J.; Rother, C.; Criminisi, A. TextonBoost for image understanding: Multi-class object recognition and segmentation by jointly modeling texture, layout, and context. *Int. J. Comput. Vis.* **2009**, *81*, 2–23. [CrossRef]
38. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 26 June–1 July 2016; pp. 770–778.
39. Krizhevsky, A. *Learning Multiple Layers of Features from Tiny Images*; University of Toronto: Toronto, ON, Canada, 2009.
40. Deng, J.; Dong, W.; Socher, R.; Li, L.-J.; Li, K.; Fei-Fei, L. ImageNet: A Large-Scale Hierarchical Image Database. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 20–25 June 2009; pp. 248–255.
41. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv* **2014**, arXiv:1409.1556.
42. Russakovsky, O.; Deng, J.; Su, H.; Krause, J.; Satheesh, S.; Ma, S.; Huang, Z.; Karpathy, A.; Khosla, A.; Bernstein, M.; et al. ImageNet Large Scale Visual Recognition Challenge. *Int. J. Comput. Vis.* **2015**, *115*, 211–252. [CrossRef]
43. Long, J.; Shelhamer, E.; Darrell, T. Fully convolutional networks for semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; p. 343.
44. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. In Proceedings of the International Conference on Neural Information Processing Systems, Montreal, QC, Canada, 7–12 December 2015; MIT Press: Cambridge, MA, USA, 2015; pp. 91–99.
45. Joachims, T. Text categorization with Support Vector Machines: Learning with many relevant features. *J. Mach. Learn. Res.* **2001**, *2*, 137–142.
46. Lewis, D.D. Naive Bayes at forty: The independence assumption in information retrieval. *Mach. Learn.* **2003**, *52*, 4–15.
47. Bojanowski, P.; Grave, E.; Joulin, A.; Mikolov, T. Enriching word vectors with subword information. *Trans. Assoc. Comput. Linguist.* **2017**, *5*, 135–146. [CrossRef]
48. Graves, A.; Mohamed, A.-R.; Hinton, G. Speech recognition with deep recurrent neural networks. In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, Vancouver, BC, Canada, 26–31 May 2013; pp. 6645–6649.
49. Chang, C.C.; Lin, C.J. LIBSVM: A library for support vector machines. *ACM Trans. Intell. Syst. Technol.* **2011**, *2*, 1–27. [CrossRef]
50. Witten, L.H.; Frank, E.; Hall, M.A. *Data Mining: Practical Machine Learning Tools and Techniques*, 3rd ed.; Morgan Kaufmann: Burlington, MA, USA, 2011.
51. Manning, C.; Raghavan, P.; Schütze, H. *Introduction to Information Retrieval*; Cambridge University Press: Cambridge, UK, 2008.
52. Fawcett, T. An introduction to ROC analysis. *Pattern Recognit. Lett.* **2006**, *27*, 861–874. [CrossRef]
53. James, G.; Witten, D.; Hastie, T.; Tibshirani, R. *An Introduction to Statistical Learning*; Springer: Berlin/Heidelberg, Germany, 2013.

54. Demir, H.B. Sentiment analysis with machine learning algorithms. *J. King Saud Univ.-Comput. Inf. Sci.* **2018**, *30*, 330–335.
55. Khan, T.A.; Alam, M.; Ahmed, S.F.; Shahid, Z.; Mazliham, M.S. A factual flash flood evaluation using SVM and K-NN. In Proceedings of the 2019 IEEE 6th International Conference on Engineering Technologies and Applied Sciences (ICETAS-2019), Kuala Lumpur, Malaysia, 20–21 December 2019. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.