Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN：1671-5497
E-Publication：Online Open Access
Vol: 41 Issue: 05-2022
DOI 10.17605/OSF.IO/B5XHS

FAKE NEWS DETECTION ON ONLINE SOCIAL NETWORKS BASED ON MACHINE AND DEEP LEARNING METHODS

[1]IRFAN ALI KANDHRO, [2]Dr. SHAFIQ-UR-REHMAN MASSAN, [3]ASIF KHAN, [4]ALI ORANGZEB PANHWAR, [5]RAHOOL GIR GOSWAMI , [6]BAKHTAWAR MALIK and [7]HINA KHAN

[1]Department of Computer Science, Sindh Madressatul Islam University (SMIU) Karachi, Pakistan
*Corresponding Author: irfan@smiu.edu.pk
[2]Department of Computer Science, Newports Institute of Communications and Economics, Karachi, Pakistan
[3]Artificial Intelligence & Mathematical Sciences, Sindh Madressatul Islam University (SMIU), Karachi, Pakistan
[4]Faculty of Computing Science and IT, Benazir Bhutto Shaheed University, Lyari Karachi,Pakistan
[5]Department of Computer Science. Sindh Madressatul Islam University (SMIU) Karachi, Pakistan
[6]Department of Computer Science. Sindh Madressatul Islam University (SMIU) Karachi, Pakistan
[7]Department of Computer Science, Sindh Madressatul Islam University (SMIU) Karachi,Pakistan

ABSTRACT
A rapid rise and out-reach of social data and the web have led to the broadcasting of dubious and untrusted content a wider audience, which is negatively impact on people's life. This research study focuses on fake and original news classification based on features and unseen patterns. Over the past decades, many of research studies have been conducted to tackle the detection and identification of fake news. In this paper, we focus on classifying the fake news using different machine learning algorithms such as LSVM, Perceptron, KNN, Random Forest (RF), KNN and so on. The actual challenge is the lack of an efficient way to tell the difference between real view and fake on, even sometimes humans are also confused and can't differentiate. The proposed system works on two steps 1) get the relevant article data and match with the knowledge database and secondly, it identifies the patterns and underlying style of fake content. The classifier has ability to detect the fake news on newly introduced fake news dataset. The experimental result shows that the given classification model obtains up to (96 %) accuracy on Decision tree, AdaBoost and LR approach as compared to other machine learning algorithms.
Keywords: news classification; machine learning; fake news; text classification; text categorization.

1. Introduction

Concerning to news, the online web offers many possibilities with respect of different challenges. the medium of communication and web growing over the time. It has become simpler for the customers to receive the latest on fingertips through social media sites. These mediums are important to share the latest ideas, group discussion and different issues of education, governance, crime, and health. fake news are propagated on these sites that has become a great challenge. Fake news is designed to spread hoaxes, propaganda, and disinformation. In front of the Internet, this is an intentional spreading that is disliked through traditional news media or social media. Incorrect information is defined significantly. This shows the other when the fake site is removed. As a medium for news updates, the use of social networks is the double edge. On the other hand, social networks are easily accessible, and there are few costs, and provide the distribution of impressive levels of information. As others, social

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN：1671-5497
E-Publication：Online Open Access
Vol: 41 Issue: 05-2022
DOI 10.17605/OSF.IO/B5HXS

networks provide ideal places to create and distribute fake news. It is very influential and there is the ability to spread very quickly. The online carrier thinks about how large information it has. Regardless of the advantage of using social media, the same with the help of digital media is not actually accurate information. Regardless of that, there is a reason for providing information about the network and multiplied by faster and faster information on digital media. Rotate the news to imitate the actual headline and rotate the record. Modern life is suitable for gratitude to the tremendous contribution of Internet technology to transfer and share information. Currently, the current era of the Internet is a double sword that can communicate with the moment of the eye, not to know whether the information sent is actually or incorrect. It is very easy to announce what they want, it is very easy to accept, but it can cause panic to manipulate the wrong information of the Internet by posting incorrect information about the internet. Recently, via a variety of platforms intensively or repeatedly increased the diffusion of fake news. Over the past decade, social networks have been changed quickly with the results of social networks for people, one of the major information resources for people. in the world. The definition of illegal news on social networks is very important, but it is also a technically difficult task at the same time. It is a lot of information on the Internet that will become impossible to decrypt the truth. This leads to a problem with a fake news.

2. Literature Review

Our task is a web application which provides you with the direction of the everyday daily practice of phony news, spam messages in day-by-day news channels, Facebook, Twitter, Instagram, and other online media. We have shown a few information investigations from our dataset which have recovered from numerous internet-based web-based media and show the principal source till now counterfeit news and genuine news are locked in. [20,21,25]. Our venture has gone head-to-head with different models prepared by our own and furthermore some pretrained models removed from Felipe Adachi. The exactness of the model is around 95% for all the independent models and 97% for this pretrained model. This model can distinguish all news and message which relate to Coronavirus 19, political news, topography, and so on. At present, many individuals are using the web as a focal stage to track down the data about reality in the world and should proceed. Subsequently [21,23,25]. the false knowledge it carries, its writing style, its propagation patterns, and the credibility of its source [40]. I have noticed above we will make counterfeit news and message identification models which distinguish the truth of the news and message. Likewise, whose utilization our site can see the state-of-the-art about principal source or watchword are getting most phony information and message and planned up with graphs. Later and everybody needs to know how to forestall this consequently we are giving a few significant hints to keep away from this phony insight about spreading gossip on the planet.

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN：1671-5497
E-Publication：Online Open Access
Vol: 41 Issue: 05-2022
DOI 10.17605/OSF.IO/B5HXS

Design and Analysis of Fake News Classifier based on Machine and Deep Learning Methods

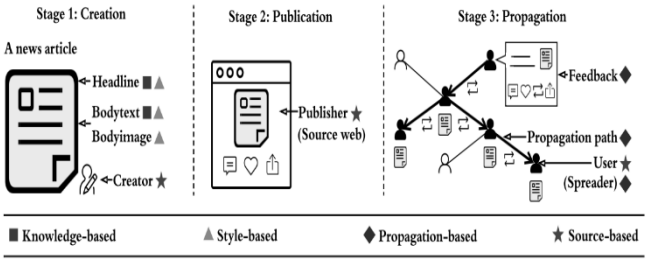

Figure 1. Process of Fake new [40]

3. Methodology

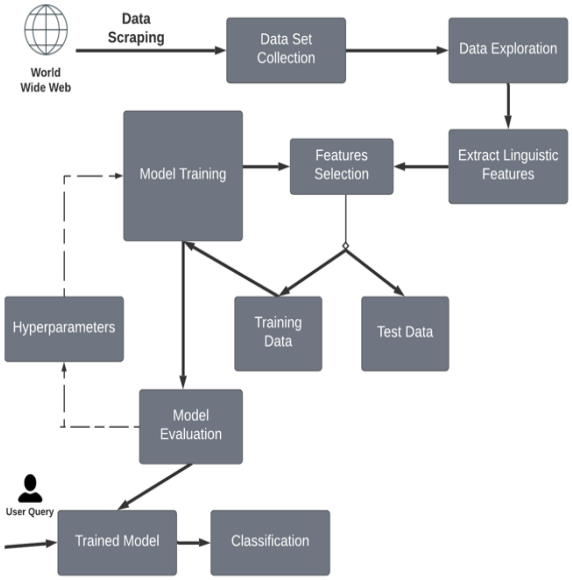

Figure 2 System Architecture of Fake News Classifiers [41].

In our proposed framework, we`re focusing the figure 2 modern-day composition through methods of giving outfit approach wonderful etymological capacities to bunch opinions from numerous areas. There are wonderful, intended places that put up genuine data substance, and or three specific locales, for instance, PolitiFact and Snopes. Additionally, there are open vaults which probably stayed through the manner of method of researchers to hold with the most modern-day evaluation of through manner of method of and through manner of method of open datasets. In any case, we picked three datasets for our tests which contain data from numerous spaces (like administrative issues, redirection, development, and sports) . Data devices are available on the Internet and are isolated from the Internet. The number one dataset is the ISOT Fake News dataset. The 2nd and 1/3 datasets are unconditionally open on Kaggle. A reduced view of the data set is proven. clear out unwanted article

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN：1671-5497
E-Publication：Online Open Access
Vol: 41 Issue: 05-2022
DOI 10.17605/OSF.IO/B5HXS

elements collectively with essay, e-book date, URL, grouping, etc. Similarly, articles without a body text or a great deal less than 20 terms in the body text might be deleted. Multicolumn articles are changed into unequalled segment articles for consistency of format and development. These carrying sports are performed on all the datasets to benefit consistency of affiliation and development. Whenever the right developments are picked after the data cleaning and examination diploma, the following diploma includes extraction of the phonetic features. Semantic components protected unique innovative houses changed over proper right into numerical format loads. These features be a part of level of terms supplying superb or miserable sentiments; level of save you terms; emphasis; artwork terms; easygoing language; and level of unique linguistic form applied in sentences like descriptors, social terms, and interest terms. To accomplish the extraction of components from the corpus, we used the LIWC 2015 machine which orchestrates the text into wonderful discrete and regular variables. The LIWC device isolates 90 3 tremendous features from some peculiar text. As every one of the components removed the usage of the contraption are numerical developments, no encoding is every day for whole elements. Regardless, scaling is used to make certain that wonderful component`s developments lie withinside the amount of (0, 1).)is important as specific developments are withinside the amount of 0 to a hundred, (for instance, charge values), on the identical time as numerous developments have sporadic reach, (for instance, word counts). It then uses the input characteristic to music the numerous artificial intelligence models. Each data set is damaged up separately proper right into a 70/30 cut up plan and test. During event planning and testing, we are improving articles simply so fake and real articles can be allocated fairly. Education scores are prepared with numerous hyperparameters to benefit the maximum accuracy for a given data set with the maximum beneficial fitness amongst variances. Each model is staged for wonderful events with wonderful constraints; the usage of business enterprise seeks to update the model for fantastic effects. Novel to this evaluation, wonderful get together methodologies like terminating, supporting, and projecting a polling form classifier are researched to survey the presentation. We used wonderful vote based completely classifiers made from three gaining knowledge: the chief projecting a polling form classifier is an outfit of essential backslide, sporadic woods, and KNN, however the 2nd surely classifier consists of determined backslide, direct SVM, and portrayal and backslide trees. measures used for installing the majority rule classifiers is to plot person models with the fantastic limits and a quick time later test the model considering the guarantee of the very last effects mark in view of vital votes. We have prepared a pressing outfit concerning a hundred choice trees, however supporting accumulating computations are used, XGBoost and AdaBoost.

3.1. Datasets

The News will be gathered from totally various sources, like squeeze organization landing pages, web search tools, and online media sites. Nonetheless, physically deciding the honesty of reports could be a troublesome errand, [20,26,28] occasionally requiring annotators with area experience UN office performs cautious examination of cases and additional evidence, setting, and reports from legitimate sources. For the most part, news data with comments will be assembled inside the accompanying ways: proficient columnists, Fact-really looking at sites, business indicators, and Crowd-obtained staff.

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN：1671-5497
E-Publication：Online Open Access
Vol: 41 Issue: 05-2022
DOI 10.17605/OSF.IO/B5HXS

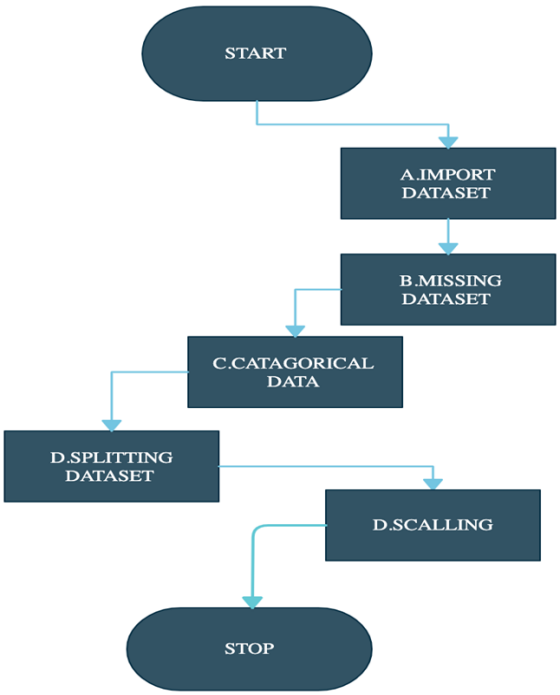

Figure .3 Feature Engineering and Pre-processing

3.1. Feature Engineering and Pre-processing

In this segment, we will investigate three distinct ways one can use to make preparing and testing sets. Prior to bouncing into these methodologies, how about we make a faker dataset that will use for showing purposes. [20,26,28] In the models beneath, we will expect that we have a dataset put away in memory as a panda.

Informational collections are accessible in .csv design. A csv record stores even information in plain text. Each line of the document is an information record. [20,21,26,28] We utilize the read_csv strategy for the panda's library to peruse a nearby CSV document as a data frame. The information we get is seldom homogenous. At times information can be absent and it should be taken care of, so it doesn't lessen the presentation of our AI model. We can't utilize values like "Male" and "Female" in numerical conditions of the model, so we really want to encode these factors into numbers. After that we utilize the fit transform technique on the downright highlights. [20,26,28] : In the wake of Encoding it is important to recognize the factors in a similar section, for this we will utilize OneHotEncoder class from sklearn. Preprocessing library. Presently we partition our information into two sets, one for preparing our model called the preparation set and the other for testing the presentation of our model called the test set. The split is by and large 80/20. To do this we import the "train_test_split"[28,30,32] technique for the "sklearn. model selection" library. The greater part of the AI calculations utilizes the Euclidean distance between two main elements in their calculations. Along these lines, high sizes elements will gauge more somewhere far off estimations than highlights with low sizes. To stay away from this Feature normalization or Z-score standardization is utilized[28,30,32].

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN：1671-5497
E-Publication：Online Open Access
Vol: 41 Issue: 05-2022
DOI 10.17605/OSF.IO/B5HXS

4. Result and Discussion

It is important to have a mechanism for detecting fake news or at least notification for not everything you read on social media may be true. So, we should always think critically. Therefore, we can help people make informed decisions and You won't be fooled into thinking what other people want. Steer them to believe.

We involved gaining knowledge of calculations related to our proposed approach to assess the exhibition of phony records place classifiers. Logistic Regression. As we`re characterizing text based totally mostly on a vast list of skills, with a paired result, a calculated relapse (LR) model is utilized, as it gives the natural situation to order issues into double or numerous trainings. We completed hyperparameters tuning to advantage the notable very last consequences for all individual dataset, whilst numerous limitations are tried preceding to buying the best exactness from the LR model. Numerically, the calculated relapse speculation functionality can be characterized. Support Vector Machine. SVM is the model for double characterization issues and is offered in exceptional bits capacities. The goal of a SVM model is to gauge a hyperplane (or choice limit) based totally mostly on a list of skills to install essential informative elements. The issue of a hyperplane differs as in line with the quantity of highlights. As there may be exceptional opportunities for a hyperplane to exist in a N-layered space. Table 2 sums up the exactness achieved through every calculation on the four considered dataset. It is obvious that the greatest exactness achieved on FAKE NEWS Fake News Dataset is almost all the manner, achieved thru strange woods calculation. Direct SVM, multi-aspect perceptron, stowing classifiers, and helping classifiers achieved anaccuracy of 98%. Benchmark calculations Wang-CNN and Wang-Bi-LSTM carried out more unfortunate than any very last algorithms. On REAL NEWS, stowing classifier (Decision trees) and assisting classifier (XGBoost) are the tremendous performing algorithms, sporting out anaccuracy of 98%. Curiously, without delay SVM, arbitrary woodland, and Perez-LSVM carried out inadequately on REAL NEWS. Individual university students located an accuracy of 47.75%, even though amassing university students` exactness is 81.5%. A comparative pattern is discovered for FAKE NEWS, wherein person university college students' accuracy is 80%, even though organization university college students' precision is 93.5%. Nonetheless, in no manner like REAL NEWS, the tremendous performing calculation on FAKE NEWS is Perez-LSVM which achieved an exactness of 96%. On REAL NEWS (FAKE NEWS, REAL NEWS, and FAKE NEWS consolidated), the tremendous performing set of rules is strange timberland (91% precision). By and large, in-dividual university college students achieved an exactness of 85%, even though troupe university college students achieved 88.16% precision. The most extraordinarily terrible performing calculation is Wang-Bi-LSTM which achieved an exactness of 62%. figure 4 sums up the normal exactness of all algorithms over the dataset. The tremendous performing calculation is packing classifiers (decision trees) (precision 95%), on the identical time because the most terrible performing calculation is Wang-Bi-LSTM (exactness 64.25%). Individual university college students' precision is 77.6% on the identical time because the exactness of organization university college students is 92.25%. Arbitrary backwoods achieved better precision on dataset aside from REAL NEWS. Nonetheless, an exactness score on my own is surely now not a

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN：1671-5497
E-Publication：Online Open Access
Vol: 41 Issue: 05-2022
DOI 10.17605/OSF.IO/B5HXS

first-rate diploma to assess the exhibition of a model; in this manner, we furthermore look at execution of gaining knowledge of models based mostly on evaluation, accuracy, and F1-score.

tables 3 and 4 sum up the evaluation, the precision and recallaccuracy of every calculation on fake and real news dataset. As far as normal accuracy (Table 2), helping the classifier AdaBoost achieved tremendous outcomes. The accuracy of the assisting classifierAdaBoost on dataset with two classes is (Precision 95% and Recall 97%). Irregular timberland (RF) achieved an accuracy of 96%; in any case, at the two datasets (removing the dataset with the most reduced score, i.e., REAL NEWS), the normal accuracy of arbitrary backwoods leaped to 90%. The comparison score for helping classifiers (XGBoost) is 89% also.

In view of the evaluation execution metric, packing classifier (preference trees) stands tremendous thru sporting an evaluation score of 0.942. This is firmly trailed through the helping classifier (XGBoost) which achieved an evaluation of 0.94. Among the benchmark calculations, Perez-LSVM is taken into consideration as a tremendous performing calculation, sporting an evaluation score of 0.92. The calculations showed a comparative presentation conducted on accuracy as that of accuracy. Helping classifiers (XGBoost) achieve accuracy of figure 4is a graphical portrayal of normal performance of gaining knowledge of calculations on dataset using precision, evaluation, and recall. It has a bent to be seen that there isn't always plenty of distinction between several the exhibition of gaining knowledge of algorithms using certainly considered one among a type of execution measurements.The organization scholar XGBoost carried out better in evaluation with one of a kind gaining knowledge of models on all exhibitions.

Table 1 Accuracy of Various Machine Learning algorithm

TECHNIQUES	FAKE NEWS	REAL NEWS
(LR)	**0.96**	0.95
(LSVM)	0.91	0.90
Perceptron	0.93	0.90
(KNN)	0.85	0.85
Rand forest (RF)	0.91	0.90
(RF, LR, KNN)	0.93	0.87
LR, LSVM, CART)	0.94	0.89
(Decision trees)	0.96	**0.97**
(AdaBoost)	0.96	0.94
(XGBoost)	0.90	0.89
Perez-LSVM	0.10	0.78

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN：1671-5497
E-Publication：Online Open Access
Vol: 41 Issue: 05-2022
DOI 10.17605/OSF.IO/B5HXS

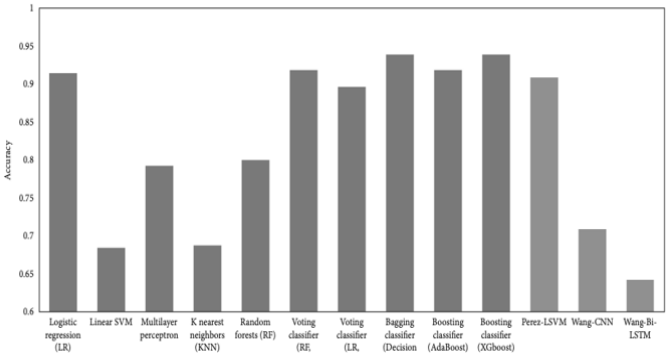

Figure .4Model accuracy on ML and DL

Table 2. Recall on Various ML and DL

TECHNIQUES	RECAL
(LR)	**0.96**
(LSVM)	0.90
Perceptron	0.92
(KNN)	0.84
Rand forest (RF)	0.91
(RF, LR, KNN)	0.94
LR, LSVM, CART)	0.95
(Decision trees)	**0.97**
(AdaBoost)	**0.97**
(XGBoost)	0.90
Perez-LSVM	0.10

Table 3. Precision on Various ML and DL

TECHNIQUES	PRECISION
(LR)	0.96
(LSVM)	0.89
Perceptron	**0.93**
(KNN)	0.82
Rand forest (RF)	0.90
(RF, LR, KNN)	0.93
LR, LSVM, CART)	0.92
(Decision trees)	0.95
(AdaBoost)	**0.95**
(XGBoost)	0.91
Perez-LSVM	0.9

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN：1671-5497
E-Publication：Online Open Access
Vol: 41 Issue: 05-2022
DOI 10.17605/OSF.IO/B5HXS

4. Conclusion

A rapid rise and out-reach of social data and the web have led to the broadcasting of dubious and untrusted content a wider audience, which is negatively impact on people's life. This research study focuses on fake and original news classification based on features and unseen patterns. Over the past decades, many of research studies have been conducted to tackle the detection and identification of fake news. In this paper, we focus on classifying the fake news using different machine learning algorithms such as LSVM, Perceptron, KNN, Random Forest (RF), KNN and so on. The actual challenge is the lack of an efficient way to tell the difference between real view and fake on, even sometimes humans are also confused and can't differentiate. The proposed system works on two steps 1) get the relevant article data and match with the knowledge database and secondly, it identifies the patterns and underlying style of fake content. The classifier has ability to detect the fake news on newly introduced fake news dataset. The experimental result shows that the given classification model obtains up to (96%) accuracy on Decision tree,AdaBoost and LR approach as compared to other machine learning algorithms.

References

1. Douglas), "News consumption and the new electronic media," The International Journal of Press/Politics, vol. 11, no. 1, pp. 29–52, 2006.
2. (J. Wong), "Almost all the traffic to fake news sites is from facebook, new data show," 2016.
3. (D. M. J. Lazer, M. A. Baum, Y. Benkler et al)., "The science of fake news," Science, vol. 359, no. 6380, pp. 1094–1096, 2018.
4. (S. A. Garc´ıa, G. G. Garc´ıa, M. S. Prieto, A. J. M. Guerrero, and C. R. Jime´nez,) "The impact of term fake news on the scientific community scientific performance and mapping in web of science," Social Sciences, vol. 9, no. 5, 2020.
   (A. D. Holan), 2016 Lie of the Year: Fake News, Politifact, Washington, DC, USA, 2016.
5. (S. Kogan, T. J. Moskowitz, and M. Niessner), "Fake News: Evidence from Financial Markets," 2019, https://ssrn.com/ abstract=3237763.
   (A. Robb) "Anatomy of a fake new scandal," Rolling Stone, vol. 1301, pp. 28–33, 2017.
6. (J. Soll), "The long and brutal history of fake news," Politico Magazine, vol. 18, no. 12, 2016.
7. (J. Hua and R. Shaw), "Coronavirus (covid-19) "infodemic" and emerging issues through a data lens: the case of China," International Journal of Environmental Researchand Public Health, vol. 17, no. 7, p. 2309, 2020.
8. (N. K. Conroy, V. L. Rubin, and Y. Chen, "Automatic de)- ception detection: methods for finding fake news," Proceedings of the Association for Information Science and Technology, vol. 52, no. 1, pp. 1–4, 2015.
9. (F. T. Asr and M. Taboada), "Misinfotext: a collection of news articles, with false and true labels," 2019.
10. (K. Shu, A. Sliva, S. Wang, J. Tang, and H. Liu), "Fake news detection on social media," ACM SIGKDD Explorations Newsletter, vol. 19, no. 1, pp. 22–36, 2017.
11. (S. Vosoughi, D. Roy, and S. Aral), "The spread of true and false news online," Science, vol. 359, no. 6380, pp. 1146–1151, 2018.
12. (H. Allcott and M. Gentzkow), "Social media and fake news in the 2016 election," Journal of Economic Perspectives, vol. 31, no. 2, pp. 211–236, 2017.
13. 15. (V. L. Rubin, N. Conroy, Y. Chen, and S. Cornwell), "Fake news or truth? using satirical cues to detect potentially misleading news," in Proceedings of the Second Workshop on Computa- tional Approaches to Deception Detection, pp. 7–17, San Diego, CA, USA, 2016.
14. (H. Jwa, D. Oh, K. Park, J. M. Kang, and H. Lim,) "exBAKE: automatic fake news detection model based on bidirectional encoder representations from transformers (bert)," Applied Sciences, vol. 9, no. 19, 2019.

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN：1671-5497
E-Publication：Online Open Access
Vol: 41 Issue: 05-2022
DOI 10.17605/OSF.IO/B5HXS

15. (H. Ahmed, I. Traore, and S. Saad), “Detection of online fake news using n-gram analysis and machine learning tech- niques,” in Proceedings of the International Conference on Intelligent, Secure, and Dependable Systems in Distributed and Cloud Environments, pp. 127–138, Springer, Vancouver, Canada, 2017.
16. (W. Y. Wang, Liar, Liar Pants on Fire): A New Benchmark Dataset for Fake News Detection, Association for Computa- tional Linguistics, Stroudsburg, PA, USA, 2017.
17. Riedel, I. Augenstein, G. P. Spithourakis), and S. Riedel, “A simple but tough-to-beat baseline for the fake news challenge stance detection task,” 2017, https://arxiv.org/abs/1707. 03264.
18. (N. Ruchansky, S. Seo, and Y. Liu, “Csi): a hybrid deep model for fake news detection,” in Proceedings of the 2017 ACM on Conference on Information and Knowledge Management, pp. 797–806, Singapore, 2017.
19. (V. Pe´rez-Rosas, B. Kleinberg, A. Lefevre, and R. Mihalcea,)“Automatic detection of fake news,” 2017, https://arxiv.org/ abs/1708.07104.
20. (P. Bu¨hlmann), “Bagging, boosting and ensemble methods,” in Handbook of Computational Statistics, pp. 985–1022, Springer, Berlin, Germany, 2012.
21. (H. Ahmed, I. Traore, and S. Saad), “Detecting opinion spams and fake news using text classification,” Security and Privacy, vol. 1, no. 1, 2018.
22. (Kaggle), Fake News, Kaggle, San Francisco, CA, USA, 2018, https://www.kaggle.com/c/fake-news.
23. (J. Bergstra and Y. Bengio), “Random search for hyper-parameter optimization,” Journal of Machine Learning Research, vol. 13, pp. 281–305, 2012.
24. (T. M. Mitchell), The Discipline of Machine Learning, Carnegie Mellon University, Pittsburgh, PA, USA, 2006.
25. (N. Cristianini and J. Shawe-Taylor), An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods, Cambridge University Press, Cambridge, UK, 2000.
26. (T. Hofmann, B. Scho¨lkopf, and A. J. Smola,) “Kernel methods in machine learning,” The Annals of Statistics, vol. 36, no. 3, pp. 1171–1220, 2008. Complexity 11
27. (V. Kecman), Support Vector Machines-An Introduction in “Support Vector Machines: Theory and Applications”, Springer, New York City, NY, USA, 2005.
28. (S. Akhtar, F. Hussain, F. R. Raja et al.), “Improving mispro- nunciation detection of arabic words for non-native learners using deep convolutional neural network features,” Elec- tronics, vol. 9, no. 6, 2020.
29. Ruta and B. Gabrys), “Classifier selection for majority voting,” Information Fusion, vol. 6, no. 1, pp. 63–81, 2005.
30. (B. Gregorutti, B. Michel, andP.Saint-Pierre), “Correlation And variable importance in random forests,” Statistics and Com- puting, vol. 27, no. 3, pp. 659–678, 2017.
31. (L. Breiman, J. Friedman, R. Olshen, and C. Stone,) Classifi- cation and Regression Trees, Springer, Berlin, Germany, 1984.
32. (R. E. Schapire), “A brief introduction to boosting,” IJCAI, vol. 99, pp. 1401–1406, 1999.
33. M. Dos Santos, R. Sabourin, and P. Maupin), “Overfitting cautious selection of classifier ensembles with genetic algo- rithms,” Information Fusion, vol. 10, no. 2, pp. 150–162, 2009.
34. (T. Chen and C. Guestrin,) “Xgboost: a scalable tree boosting system,” in Proceedings of the 22nd ACM SIGKDD Interna- tional Conference on Knowledge Discovery and Data Mining, pp. 785–794, San Francisco, CA, USA, 2016.
35. (T. Hastie, S. Rosset, J. Zhu, and H. Zou), “Multi-class ada- boost,” Statistics and its Interface, vol. 2, no. 3, pp. 349–360, 2009.
36. (L. Lam and S. Y. Suen), “Application of majority voting to pattern recognition: an analysis of its behavior and perfor- mance,” IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans, vol. 27, no. 5, pp. 553–568, 1997.
37. Zhou, Xinyi, and Reza Zafarani. "A survey of fake news: Fundamental theories, detection methods, and opportunities." ACM Computing Surveys (CSUR) 53.5 (2020): 1-40.
38. Ahmad, Iftikhar, et al. "Fake news detection using machine learning ensemble methods." Complexity 2020 (2020).
40. Kandhro, S. Z. Jumani, F. Ali, Z. U. Shaikh, M. A. Arain, and A. A. Shaikh, “Performance Analysis of Hyperparameters on a Sentiment Analysis Model ”, *Eng. Technol. Appl. Sci. Res.*, vol. 10, no. 4, pp. 6016–6020, Aug. 2020.