Rumors Detection on Social Platforms Using NLP Methods

Nsrin Ashraf

Computer Science Department

Faculty of Computers and Artificial Intelligence

Benha University

Abstract

Social media platforms have grown rapidly in recent years, with billions of people worldwide using them for communication, entertainment, and information. Social media development has dramatically impacted society, affecting how people interact, communicate, and consume information. While social media has numerous advantages, it has also prompted worries about privacy, misinformation, and the influence on mental health, especially among young people. The dissemination of rumors has been significantly impacted by social media platforms. The major platform that has been used for spreading news regarding the Covid-19 pandemic is Twitter. The Covid-19 pandemic has spread a considerable deal of false material on social media. Artificial intelligence proposed several methods to relieve the spread of fake news.

In this study, we proposed a model that can discriminate between "fake" and "true" news tweets capable of working with any up-to-date problem. To address this issue, this research explored various learning approaches to detect fake news. We compare different deep learning and machine learning methods for fake news detection, such as CNN, LSTM, Naïve Bayes, and Support Vector Machine. The efficiency of these models was evaluated on benchmark datasets and self-collected dataset. This research aims to improve the model used in classifying rumors by utilizing various techniques for text representation such as Word Embedding and TF-IDF. It involves extracting the underlying meanings in texts by searching for semantic relationships between words, phrases, and texts. These processes help in analyzing and understanding texts. The efficiency of these models was tested by training data on a set of tweets. New tweets were collected using `Snscrape` to track different writing methods and build a model capable of detecting errors with all the changes that occur in a word and returning to the origin of the word.

The results of the first model using TF-IDF algorithms and machine learning algorithms showed the superiority of Multi-Layer Perceptron algorithm, achieving an accuracy of 93.8% and an F-score of 93.6% when applied to the English language. The results of the Arabic language models showed the superiority of the Support Vector Machine algorithm, achieving an accuracy of 82.90%, while the K-Nearest Neighbor achieved better results with an F-score of 57.5%. The results showed the superiority of Uni-gram text vectorization over Bi-gram. GloVe word embedding was used with deep learning algorithms to improve text understanding and discover relationships between words. Recurrent neural networks achieved the best results for the English language with an accuracy of 99%, but the ensemble learning model achieved better results in terms of F-score achieved 97%. The Convolutional Neural Network achieved the best results with the Arabic language achieved an accuracy of 83% using the Accuracy measure, while the Ensemble learning model achieved better results using the F-score at a rate of 81.7%.

The second step was to test the model on a new test set that had not been tested before. A significant decline of about 25% was found in the English language model, achieving an accuracy of 74%. The experiments showed that adding some modifications to the evidence processing stage to develop the model made it capable of dealing with all the changes that occur in a word and showed an improvement of about 8% achieving an accuracy of 83%. As for the proposed model for the Arabic language, there was a decline of about 5%, achieving an accuracy of 70%. The results vary between deep learning models, but the BI-LSTM showed the difference between the differences in the data. With some modifications to the word processing stage to develop the model and make it capable of dealing with all the changes that occur in a word, there was an improvement of about 8% achieving an accuracy of 78%.

References

[1] Hamada Nayel and H L Shashirekha. Improving ner for clinical texts by ensemble approach using segment representations. In *Proceedings of the 14th International Conference on Natural Language Processing (ICON-2017)*, pages 197–204, Kolkata, India, December 2017. NLP Association of India.

[2] Hamada A. Nayel and Shashirekha H. L. DEEP at HASOC2019: A machine learning framework for hate speech and offensive language detection. In Parth Mehta, Paolo Rosso, Prasenjit Majumder, and Mandar Mitra, editors, *Working Notes of FIRE 2019 - Forum for Information Retrieval Evaluation, Kolkata, India, December 12-15, 2019*, volume 2517 of *CEUR Workshop Proceedings*, pages 336–343. CEUR-WS.org, 2019.

[3] Hamada A. Nayel, Walaa Medhat, and Metwally Rashad. BENHA@IDAT: Improving Irony Detection in Arabic Tweets using Ensemble Approach. In Parth Mehta, Paolo Rosso, Prasenjit Majumder, and Mandar Mitra, editors, *Working Notes of FIRE 2019 - Forum for Information Retrieval Evaluation, Kolkata, India, December 12-15, 2019*, volume 2517 of *CEUR Workshop Proceedings*, pages 401–408. CEUR-WS.org, December 2019.

[4] Qingyu Chen, Robert Leaman, Alexis Allot, Ling Luo, Chih-Hsuan Wei, Shankai Yan, and Zhiyong Lu. Artificial intelligence in action: Addressing the covid-19 pandemic with natural language processing. *Annual Review of Biomedical Data Science*, 4(1):313–339, 2021. PMID: 34465169.

[5] Irene Li, Yixin Li, Tianxiao Li, Sergio Alvarez-Napagao, Dario Garcia-Gasulla, and Toyotaro Suzumura. What are we depressed about when we talk about covid-19: Mental health analysis on tweets using natural language processing. In *Artificial Intelligence XXXVII: 40th SGAI International Conference on Artificial Intelligence, AI 2020, Cambridge, UK, December 15–17, 2020, Proceedings 40*, pages 358–370. Springer, 2020.

[6] Hamada Nayel, Eslam Amer, Aya Allam, and Hanya Abdallah. Machine learning-based model for sentiment and sarcasm detection. In *Proceedings of the Sixth Arabic Natural Language Processing Workshop*, pages 386–389, Kyiv, Ukraine (Virtual), April 2021. Association for Computational Linguistics.

[7] Mohammed Ali Al-Garadi, Yuan-Chi Yang, and Abeed Sarker. The role of natural language processing during the covid-19 pandemic: Health applications, opportunities, and challenges. In *Healthcare*, volume 10, page 2270. Multidisciplinary Digital Publishing Institute, 2022.

[8] Jackie Ayoub, X Jessie Yang, and Feng Zhou. Combat covid-19 infodemic using explainable natural language processing models. *Information processing & management*, 58(4):102569, 2021.

[9] Hamada A. Nayel, H L Shashirekha, Hiroyuki Shindo, and Yuji Matsumoto. Improving Multi-Word Entity Recognition for Biomedical Texts. *International Journal of Pure and Applied Mathematics*, 118(16):301–3019, 2017.

[10] Kabir Alabi Sulaiman, Ismail Olatunji Adeyemi, and Ibrahim Ayegun. Information sharing and evaluation as determinants of spread of fake news on social media among nigerian youths: Experience from covid-19 pandemic. *International Journal of Knowledge Content Development & Technology*, 10(4), 2020.

[11] Amor Ben El Khettab Lalmi, Abderrahim Djaballah, and Mohamed Gharzouli. Covid-19 fake news detection using machine learning techniques: A comparative study. *EAI Endorsed Transactions on Cloud Systems*, 7(22):e3–e3, 2022.

[12] Fatima Haouari, Maram Hasanain, Reem Suwaileh, and Tamer Elsayed. ArCOV-19: The first Arabic COVID-19 Twitter dataset with propagation networks. In *Proceedings of the Sixth Arabic Natural Language Processing Workshop*, pages 82–91, Kyiv, Ukraine (Virtual), April 2021. Association for Computational Linguistics.

[13] Masood Hamed Saghayan, Seyedeh Fatemeh Ebrahimi, and Mohammad Bahrani. Exploring the impact of machine translation on fake news detection: A case study on persian tweets about covid-19. In *2021 29th Iranian Conference on Electrical Engineering (ICEE)*, pages 540–544, 2021.

[14] Oussama Maakoul, Sabah Boucht, Karima EL HACHIMI, and Salma Azzouzi. Towards evaluating the covid'19 related fake news problem: Case of morocco. In *2020 IEEE 2nd International Conference on Electronics, Control, Optimization and Computer Science (ICECOCS)*, pages 1–6, 2020.

[15] Tom Mould. Introduction to the special issue on fake news: Definitions and approaches. *Journal of American Folklore*, 131(522):371–378, 2018.

[16] Jan Deriu, Aurelien Lucchi, Valeria De Luca, Aliaksei Severyn, Simon Müller, Mark Cieliebak, Thomas Hofmann, and Martin Jaggi. Leveraging large amounts of weakly supervised data for multi-language sentiment classification. In *Proceedings of the 26th International Conference on World Wide Web*, WWW '17, page 1045–1052, Republic and

Canton of Geneva, CHE, 2017. International World Wide Web Conferences Steering Committee.

[17] Mohammed Al-Sarem, Abdullah Alsaeedi, Faisal Saeed, Wadii Boulila, and Omair AmeerBakhsh. A novel hybrid deep learning model for detecting covid-19-related rumors on social media based on lstm and concatenated parallel cnns. *Applied Sciences*, 11(17), 2021.

[18] Sheng How Kong, Li Mei Tan, Keng Hoon Gan, and Nur Hana Samsudin. Fake news detection using deep learning. In *2020 IEEE 10th Symposium on Computer Applications Industrial Electronics (ISCAIE)*, pages 102–107, 2020.

[19] N. Smitha and R. Bharath. Performance comparison of machine learning classifiers for fake news detection. In *2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA)*, pages 696–700, 2020.

[20] M. Irfan Uddin, Iftikhar Ahmad, Muhammad Yousaf, Suhail Yousaf, and Muhammad Ovais Ahmad. Fake news detection using machine learning ensemble methods. *Complexity*, 2020:8885861, 2020.

[21] Shafayat Bin Shabbir Mugdha, Sayeda Muntaha Ferdous, and Ahmed Fahmin. Evaluating machine learning algorithms for bengali fake news detection. In *2020 23rd International Conference on Computer and Information Technology (ICCIT)*, pages 1–6, 2020.

[22] Harm de Wet and Vukosi Marivate. Is it fake? news disinformation detection on south african news websites. In *2021 IEEE AFRICON*, pages 1–6, 2021.

[23] Andrei Korobeinikov, Waqas Haider Bangyal, Rukhma Qasim, Najeeb ur Rehman, Zeeshan Ahmad, Hafsa Dar, Laiqa Rukhsar, Zahra Aman, and Jamil Ahmad. Detection of fake news text classification on covid-19 using deep learning approaches. *Computational and Mathematical Methods in Medicine*, 2021:5514220, 2021.

[24] Mohammed N. Alenezi and Zainab M. Alqenaei. Machine learning in detecting covid-19 misinformation on twitter. *Future Internet*, 13(10), 2021.

[25] M. Irfan Uddin, Maha Al-Yahya, Hend Al-Khalifa, Heyam Al-Baity, Duaa AlSaeed, and Amr Essam. Arabic fake news detection: Comparative study of neural networks and transformer-based approaches. *Complexity*, 2021:5516945, 2021.

[26] Faisal Saeed Sultan Noman Qasem, Mohammed Al-Sarem. An ensemble learning based approach for detecting and tracking covid19 rumors. *Computers, Materials & Continua*, 70(1):1721–1747, 2022.

[27] Abhishek Koirala. Covid-19 fake news classification with deep learning. *Preprint*, 4, 2020.

[28] Yahya Tashtoush, Balqis Alrababah, Omar Darwish, Majdi Maabreh, and Nasser Al-saedi. A deep learning framework for detection of covid-19 fake news on social media platforms. *Data*, 7(5):65, 2022.

[29] Omayma Mahmoudi, Mouncef Filali Bouami, and Mustapha Badri. Arabic language modeling based on supervised machine learning, Jun 2022.

[30] Ali Bou Nassif, Ashraf Elnagar, Omar Elgendy, and Yaman Afadar. Arabic fake news detection based on deep contextualized embedding models. *Neural Computing and Applications*, 34(18):16019–16032, 2022.

[31] Ghada Amoudi, Rasha Albalawi, Fatimah Baothman, Amani Jamal, Hanan Al-ghamdi, and Areej Alhothali. Arabic rumor detection: A comparative study. *Alexandria Engineering Journal*, 61(12):12511–12523, 2022.

[32] Karen Sparck Jones. A statistical interpretation of term specificity and its application in retrieval. *Journal of documentation*, 28(1):11–21, 1972.

[33] Ho Chung Wu, Robert Wing Pong Luk, Kam Fai Wong, and Kui Lam Kwok. Interpreting tf-idf term weights as making relevance decisions. *ACM Transactions on Information Systems (TOIS)*, 26(3):1–37, 2008.

[34] Thushan Ganegedara. *Natural Language Processing with TensorFlow: Teach language to machines using Python's deep learning library*. Packt Publishing Ltd, 2018.

[35] Heyoung Yang and Eunsoo Sohn. Expanding our understanding of covid-19 from biomedical literature using word embedding. *International Journal of Environmental Research and Public Health*, 18(6), 2021.

[36] Tshilidzi Marwala. *Handbook of Machine Learning: Volume 1: Foundation of Artificial Intelligence*. World Scientific, 2018.

[37] Hamada A. Nayel. NAYEL@APDA: Machine Learning Approach for Author Profiling and Deception Detection in Arabic Texts. In Parth Mehta, Paolo Rosso, Prasenjit

Majumder, and Mandar Mitra, editors, *Working Notes of FIRE 2019 - Information Retrieval Evaluation, Kolkata, India, December 12-15, 2019*, volume 2517 of *CEUR Workshop Proceedings*, pages 92–99. CEUR-WS.org, December 2019.

[38] Hamada A. Nayel and H. L. Shashirekha. Mangalore-University@INLI-FIRE-2017: Indian Native Language Identification using Support Vector Machines and Ensemble Approach. In Prasenjit Majumder, Mandar Mitra, Parth Mehta, and Jainisha Sankhavara, editors, *Working notes of FIRE 2017 - Forum for Information Retrieval Evaluation, Bangalore, India, December 8-10, 2017.*, volume 2036 of *CEUR Workshop Proceedings*, pages 106–109. CEUR-WS.org, 2017.

[39] Hamada A. Nayel and H. L. Shashirekha. Mangalore university inli@fire2018: Artificial neural network and ensemble based models for INLI. In Parth Mehta, Paolo Rosso, Prasenjit Majumder, and Mandar Mitra, editors, *Working Notes of FIRE 2018 - Forum for Information Retrieval Evaluation, Gandhinagar, India, December 6-9, 2018.*, volume 2266 of *CEUR Workshop Proceedings*, pages 110–118. CEUR-WS.org, 2018.

[40] Hamada Nayel, Ahmed Hassan, Mahmoud Sobhi, and Ahmed El-Sawy. Machine learning-based approach for Arabic dialect identification. In *Proceedings of the Sixth Arabic Natural Language Processing Workshop*, pages 287–290, Kyiv, Ukraine (Virtual), April 2021. Association for Computational Linguistics.

[41] Nsrin Ashraf, Mohamed Taha, Ahmed Abd Elfattah, and Hamada Nayel. NAYEL @LT-EDI-ACL2022: Homophobia/transphobia detection for equality, diversity, and inclusion using SVM. In *Proceedings of the Second Workshop on Language Technology for Equality, Diversity and Inclusion*, pages 287–290, Dublin, Ireland, May 2022. Association for Computational Linguistics.

[42] Fathy Elkazzaz, Fatma Sakr, Rasha Orban, and Hamada Nayel. Bfcai at comma@ icon 2021: Support vector machines for multilingual gender biased and communal language identification. In *Proceedings of the 18th International Conference on Natural Language Processing: Shared Task on Multilingual Gender Biased and Communal Language Identification*, pages 70–74, 2021.

[43] Jakjoud Fatimazahra, Hatim Anas, and Abella Bouaaddi. Fpga-based hardware acceleration for svm machine learning algorithm. In *E3S Web of Conferences*, volume 229, page 01024. EDP Sciences, 2021.

[44] Shimaa Ismail, Mai K. Gallab, and Hamada Nayel. BoNC: Bag of n-characters model for word level language identification. In *Proceedings of the 19th International Confer-*

ence on Natural Language Processing (ICON): Shared Task on Word Level Language Identification in Code-mixed Kannada-English Texts, pages 34–37, IIIT Delhi, New Delhi, India, December 2022. Association for Computational Linguistics.

[45] Hamada Nayel and Ghada Amer. A simple n-gram model for urdu fake news detection. In Parth Mehta, Thomas Mandl, Prasenjit Majumder, and Mandar Mitra, editors, *Working Notes of FIRE 2021 - Forum for Information Retrieval Evaluation, Gandhinagar, India, December 13-17, 2021*, volume 3159 of *CEUR Workshop Proceedings*, pages 1150–1155. CEUR-WS.org, 2021.

[46] Guillaume Garrigos and Robert M Gower. Handbook of convergence theorems for (stochastic) gradient methods. *arXiv preprint arXiv:2301.11235*, 2023.

[47] Divakar Kapil. Stochastic vs batch gradient descent, Jun 2019.

[48] Nsrin Ashraf, Fathy Elkazzaz, Mohamed Taha, Hamada Nayel, and Tarek Elshishtawy. BFCAI at SemEval-2022 task 6: Multi-layer perceptron for sarcasm detection in Arabic texts. In *Proceedings of the 16th International Workshop on Semantic Evaluation (SemEval-2022)*, pages 881–884, Seattle, United States, July 2022. Association for Computational Linguistics.

[49] Mahmoud Sobhy, Ahmed H. Abu El-Atta, Ahmed A. El-Sawy, and Hamada Nayel. Word representation models for Arabic dialect identification. In *Proceedings of the The Seventh Arabic Natural Language Processing Workshop (WANLP)*, pages 474–478, Abu Dhabi, United Arab Emirates (Hybrid), December 2022. Association for Computational Linguistics.

[50] Carolina Bento. Recurrent neural networks explained with a real life example and python code, May 2022.

[51] John J Hopfield. Artificial neural networks. *IEEE Circuits and Devices Magazine*, 4(5):3–10, 1988.

[52] Trevor Hastie, Robert Tibshirani, Jerome H Friedman, and Jerome H Friedman. *The elements of statistical learning: data mining, inference, and prediction*, volume 2. Springer, 2009.

[53] Shamsa Khalid, Muhammad Anees Khan, MS Mazliham, Muhammad Mansoor Alam, Nida Aman, Muhammad Tanvir Taj, Rija Zaka, Muhammad Jehangir, et al. Predicting risk through artificial intelligence based on machine learning algorithms: A case of pakistani nonfinancial firms. *Complexity*, 2022, 2022.

[54] Koby Crammer, Ofer Dekel, Joseph Keshet, Shai Shalev-Shwartz, and Yoram Singer. Online passive aggressive algorithms. *Journal of Machine Learning Research*, page 551–585, 2006.

[55] Xiaojun Yang. Artificial neural networks. In *Handbook of research on geoinformatics*, pages 122–128. IGI Global, 2009.

[56] Shun ichi Amari. *The handbook of brain theory and neural networks*. MIT press, 2003.

[57] Pijush Samui, Sanjiban Sekhar Roy, and Valentina Emilia Balas. *Handbook of neural computation*. Academic Press, 2017.

[58] Neha Agrawal and Shashikala Tapaswi. The performance analysis of honeypot based intrusion detection system for wireless network. *International Journal of Wireless Information Networks*, 24:14–26, 2017.

[59] Yoon Kim. Convolutional neural networks for sentence classification. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 1746–1751, Doha, Qatar, October 2014. Association for Computational Linguistics.

[60] Jonas Teuwen and Nikita Moriakov. Convolutional neural networks. In *Handbook of medical image computing and computer assisted intervention*, pages 481–501. Elsevier, 2020.

[61] Ying Xie, Linh Le, Yiyun Zhou, and Vijay V. Raghavan. Chapter 10 - deep learning for natural language processing. In Venkat N. Gudivada and C.R. Rao, editors, *Computational Analysis and Understanding of Natural Languages: Principles, Methods and Applications*, volume 38 of *Handbook of Statistics*, pages 317–328. Elsevier, 2018.

[62] Michael Phi. Illustrated guide to lstm's and gru's: A step by step explanation, Jun 2020.

[63] Changchun Cai, Yuan Tao, Tianqi Zhu, and Zhixiang Deng. Short-term load forecasting based on deep learning bidirectional lstm neural network. *Applied Sciences*, 11(17), 2021.

[64] Thomas G Dietterich et al. Ensemble learning. *The handbook of brain theory and neural networks*, 2(1):110–125, 2002.

[65] Tae-Ki An and Moon-Hyun Kim. A new diverse adaboost classifier. In *2010 International conference on artificial intelligence and computational intelligence*, volume 1, pages 359–363. IEEE, 2010.

[66] Lior Rokach. Ensemble methods for classifiers. *Data mining and knowledge discovery handbook*, pages 957–980, 2005.

[67] Bayu Adhi Tama and Kyung Hyune Rhee. An extensive empirical evaluation of classifier ensembles for intrusion detection task. *Computer Systems Science and Engineering*, 32:149–158, 03 2017.

[68] Bayu Adhi Tama and Kyung-Hyune Rhee. An extensive empirical evaluation of classifier ensembles for intrusion detection task. *Comput. Syst. Sci. Eng*, 32(2):149–158, 2017.

[69] Fatima Haouari, Maram Hasanain, Reem Suwaileh, and Tamer Elsayed. ArCOV19-rumors: Arabic COVID-19 Twitter dataset for misinformation detection. In *Proceedings of the Sixth Arabic Natural Language Processing Workshop*, pages 72–81, Kyiv, Ukraine (Virtual), April 2021. Association for Computational Linguistics.

[70] Hamdy Mubarak, Sabit Hassan, Shammur Absar Chowdhury, and Firoj Alam. Arcovidvac: Analyzing arabic tweets about covid-19 vaccination. *arXiv preprint arXiv:2201.06496*, 2022.

[71] Nsrin Ashraf, Hamada Nayel, and Mohamed Taha. A comparative study of machine learning approaches for rumors detection in covid-19 tweets. In *2022 2nd International Mobile, Intelligent, and Ubiquitous Computing Conference (MIUCC)*, pages 384–387, 2022.

[72] Nsrin Ashraf, Hamada Nayel, and Mohamed Taha. Misinformation detection in arabic tweets: A case study about covid-19 vaccination. *Benha Journal of Applied Sciences*, 7(5):265–268, 2022.