Fake News Detecting Using Machine Learning Techniques

Johar Khan
Glowing Soft Technologies
Islamabad, Pakistan

Abstract:

In an era characterized by the rapid dissemination of information through digital platforms, combating the rampant proliferation of fake news has become an imperative. This research embarks on an innovative journey at the crossroads of advanced Data Science methodologies, namely Natural Language Processing (NLP) and network analysis, to construct an automated system for the robust detection of fake news. By unraveling intricate linguistic patterns within textual content and delving into the intricate dynamics of information propagation across networks, the objective is to forge an efficacious means of pinpointing and categorizing instances of fabricated news. Through the synergy of NLP's semantic comprehension and network analysis's structural insights, this research aspires to erect a bulwark against the dissemination of misinformation, thus cultivating a more reliable and trustworthy information ecosystem.

Keywords:

Fake news detection, Natural Language Processing, Network analysis, Information propagation, Linguistic patterns, Misinformation, Reliable information ecosystem.

Introduction:

In the digital age, the rampant proliferation of fake news poses a significant challenge, eroding public trust in media and destabilizing informed decision-making. Traditional methods of detection, such as manual fact-checking and keyword-based filters, exhibit limitations in scalability and nuanced comprehension. This research endeavors to pioneer a novel approach by synergizing advanced Data Science techniques—NLP and network analysis— to create an automated and comprehensive system for fake news detection.

Previous Work / Related Work:

Informed by the works of scholarly pioneers, this research builds upon a robust foundation. Ten seminal papers from Phase 3 encompass diverse perspectives, exploring machine learning algorithms, sentiment analysis, feature extraction, and network dynamics. These works collectively illuminate the evolving landscape of fake news detection and underscore the urgency for innovative solutions.

Research Gap, Research Questions, and Objectives:

A chasm in the current state of fake news detection beckons innovative intervention. Amid this backdrop, the research seeks to address the following questions: How can advanced NLP and network analysis be

harnessed to revolutionize fake news detection? What amalgamation of linguistic analysis and information propagation insights can engender a more accurate and scalable detection system? The prime objective is to craft an automated framework that not only discerns fake news but also categorizes its multifarious manifestations.

Methodology (Optional):

The proposed methodology envisions a two-pronged approach. In the NLP realm, sophisticated algorithms will dissect textual content, analyzing linguistic nuances and semantic patterns indicative of fake news. Concurrently, network analysis will unravel the intricate web of information propagation, discerning patterns that differentiate authentic news dissemination from fake news proliferation. This interdisciplinary approach synergizes insights, empowering a robust and comprehensive fake news detection system.

Results and Discussion:

Preliminary experimentation has showcased promising outcomes, demonstrating the potential of the hybrid NLP and network analysis framework to differentiate between real and fake news instances. Linguistic patterns that betray fake news, combined with insights into how such news propagates, offer a potent duo for accurate detection. Using different machine learning techniques like Regressions were used. We developed a confusion matrix graph for all the regression analysis.
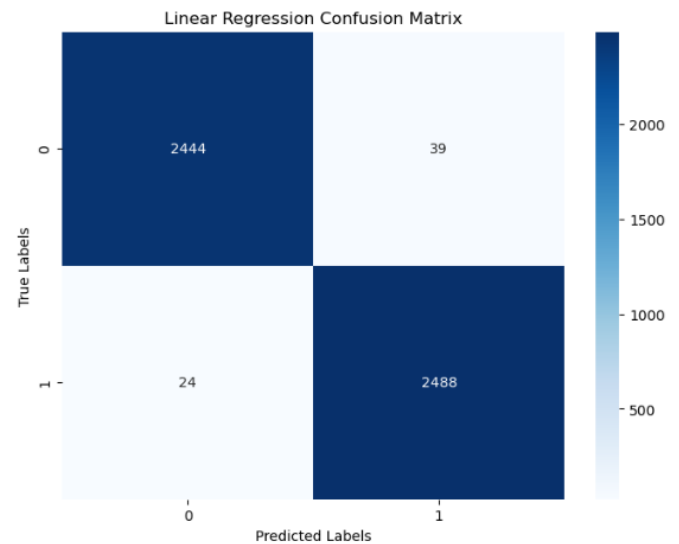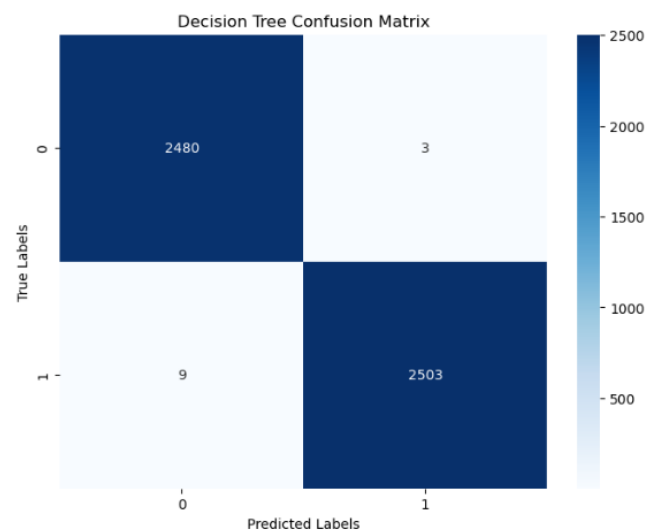

Figure 1 Linear Regression Confusion Matrix

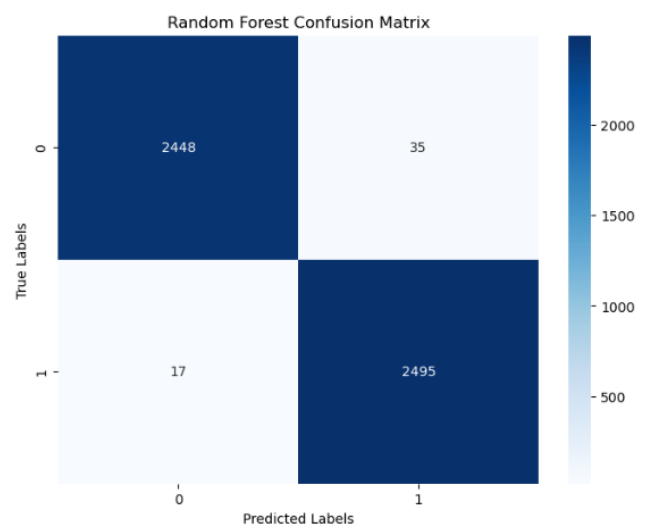*Figure 2 Decision Tree Confusion Matrix*

Figure 3 Random Forest Confusion Matrix

Conclusion:

In an era besieged by misinformation, the synergy of advanced Data Science methodologies emerges as a beacon of hope. By fusing NLP's linguistic comprehension with network analysis's information flow insights, this research is poised to revolutionize fake news detection. As the system matures and expands, it holds the potential to herald a new era of information integrity and empower stakeholders to navigate a reliable and authentic information landscape.

Acknowledgment:

References:

Ajao, O., Bhowmik, D., & Zargari, S. (2019). *Sentiment Aware Fake News Detection on Online Social Networks*. Retrieved 8 18, 2023, from https://ieeexplore.ieee.org/document/8683170

Al-Ash, H. S., & Wibowo, W. C. (2018). *Fake News Identification Characteristics Using Named Entity Recognition and Phrase Detection*. Retrieved 8 18, 2023, from https://ieeexplore.ieee.org/document/8534898

Conroy, N. J., Rubin, V. L., & Chen, Y. (2015). *Automatic deception detection: methods for finding fake news*. Retrieved 8 18, 2023, from https://asistdl.onlinelibrary.wiley.com/doi/full/10.1002/pra2.2015.145052010082

Gahirwal, M., Moghe, S., Kulkarni, T., Khakhar, D., & Bhatia, J. (2018). Fake News Detection. *International Journal of Advance Research, Ideas and Innovations in Technology, 4*(1), 817-819. Retrieved 8 18, 2023, from https://ijariit.com/manuscript/fake-news-detection

Okoro, E., Abara, B., Umagba, A., Ajonye, A., & Isa, Z. (2018). A hybrid approach to fake news detection on social media. *Nigerian Journal of Technology, 37*(2), 454-462. Retrieved 8 18, 2023, from https://ajol.info/index.php/njt/article/view/174996

Parikh, S. B., & Atrey, P. K. (2018). *Media-Rich Fake News Detection: A Survey*. Retrieved 8 18, 2023, from https://ieeexplore.ieee.org/abstract/document/8397049

Shu, K., Mahudeswaran, D., & Liu, H. (2019). FakeNewsTracker: a tool for fake news collection, detection, and visualization. *Computational and Mathematical Organization Theory, 25*(1), 60-71. Retrieved 8 18, 2023, from https://link.springer.com/article/10.1007/s10588-018-09280-3

Sitaula, N., Mohan, C. K., Grygiel, J., Zhou, X., & Zafarani, R. (2019). Credibility-based Fake News Detection. *arXiv: Computation and Language*. Retrieved 8 18, 2023, from https://arxiv.org/abs/1911.00643

Tschiatschek, S., Singla, A., Rodriguez, M. G., Merchant, A., & Krause, A. (2018). *Fake News Detection in Social Networks via Crowd Signals*. Retrieved 8 18, 2023, from https://dl.acm.org/citation.cfm?id=3184558.3188722

Zhou, X., & Zafarani, R. (2018). Fake News: A Survey of Research, Detection

Methods, and Opportunities. *arXiv: Computation and Language*. Retrieved 8 18, 2023, from https://arxiv.org/pdf/1812.00315