DELHI TECHNOLOGICAL UNIVERSITY

B TECH PROJECT SYNOPSIS

Submitted By:
Tejas Harish Borkar (2K18/CO/373) TejasBorkar.
Tushar Ahuja (2K18/CO/374) Tushar

Supervisor's Signature:

Rajni Jindal

Dr Rajni Jindal
Professor

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

FAKE NEWS DETECTION USING MACHINE LEARNING

INTRODUCTION:

A large body of recent works has focused on understanding and detecting fake news stories that are disseminated on social media. To accomplish this goal, these works explore several types of features extracted from news stories, including sources and posts from social media. In addition to exploring the main features proposed in the literature for fake news detection, we present a new set of features and measure the prediction performance of current approaches and features for automatic detection of fake news.

Types of data in Social media: There are three major forms in which social media networking Sites read a news item :

- Text (Multilingual) is analyzed by computational linguistics which focuses the genesis of text semantically and systematically. Since much of the posts are produced in the form of texts much work has been carried out on its analysis.
- Multimedia: Multiple forms of media are integrated in a single post. This may include audio, video, images, and graphics. This is very much attractive and it fetches the attention of the viewers without bothering about the text.
- Hyperlinks enable the originator of the post to cross reference to different sources and thus gains viewers the trust by certifying genesis of the post.

Types of Fake News: The various types of fake news are given below:

- Visual-based: These fake news posts use graphics a lot more in as content, which may include morphed images, doctored video, or combination of both
- User-based: This type of fabricated news is generated by fake accounts and is targeted to specific audiences which may represent certain age groups, gender, culture, political affiliations.
- Knowledge-based : these types of posts give scientific explanations to some unresolved issues and make users believe they are authentic.
- Style-based posts are written by pseudo-journalists who pretend and copy style of some accredited journalists
- Stance-based: It actually is representation of truthful statements in such a way which changes its meaning and purpose.

OBJECTIVE:

- Our objective is to build a fake news detection model using machine learning that can differentiate between fake news and legit news with maximum possible accuracy.
- We will implement this machine learning model using various techniques for example Naive Bayes, Random Forest, SVM etc
- We will also try to optimise accuracy or other parameters as much as we can, then we will compare the results obtained and will conclude the best approach among them.

MATERIAL AND METHODS:

Setting :

Research will be conducted by members remotely from home.

Duration of Study :

This is a B-Tech Major project and hence the maximum time duration is 4th year of Btech curriculum that is 7th and 8th sem.

Sample Size :

The research will be conducted in a group consist of 2 members

- Tejas Harish Borkar (2K18/CO/373)
- Tushar Ahuja (2K18/CO/374)

under the supervision of Dr Rajni Jindal Mam.

Data Collection Procedure :

For reference we will use several research papers and for implementation of machine learning model of fake news detection we will use data set from kaggle.

Data Analysis Procedure :

We will majorly use following softwares and technologies:

- Python
- Jupyter notebook
- Various libraries like numpy (for easy data analysis) , matplotlib (to plot curves for analysis)
- Natural Language Processing for preprocessing of data.

REFERENCES

[1]. Supervised learning for fake news detection, Julio C. S. Reis, Andre Correia, Fabrıcio Murai, Adriano Veloso, and Fabrıcio Benevenuto Universidade Federal de Minas Gerais Editor: Erik Cambria, Nanyang Technological University, Singapore. IEEE

[2]. Parikh, S. B., & Atrey, P. K. (2018, April). Media-Rich Fake News Detection: A Survey. In 2018 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR) (pp. 436-441). IEEE.

[3]. Conroy, N. J., Rubin, V. L., & Chen, Y. (2015, November). Automatic deception detection: Methods for finding fake news. In Proceedings of the 78th ASIS&T Annual Meeting: Information Science with Impact: Research in and for the Community (p. 82). American Society for Information Science.

[4]. Proceedings of the Third International Conference on Trends in Electronics and Informatics (ICOEI 2019) IEEE Xplore Part Number: CFP19J32-ART; ISBN: 978-1-5386-9439-8

[5]. Bondielli, A., Marcelloni, F.: A survey on fake news and rumour detection techniques. Inf. Sci. 497, 38–55 (2019). https://doi.org/10.1016/j.ins.2019.05.035. http://www.sciencedirect.com/science/article/pii/S0020025519304372

[6]. D. M. J. Lazer et al., “The science of fake news,” Science, vol. 359, no. 6380, pp. 1094–1096, 2018.

[7]. W. Y. Wang, “Liar, liar pants on fire: A new benchmark dataset for fake news detection,” in Proc. Annu. Meeting Assoc. Comput. Linguistics, 2017, pp. 422–426

[8]. S. Volkova, K. Shaffer, J. Jang Yea, and N. Hodas, “Separating facts from fiction: Linguistic models to classify suspicious and trusted news posts on twitter,” in Proc. 55th Annu. Meeting Assoc. Comput. Linguistics, 2017, pp. 647–653.

[9]. J. C. S. Reis, A. Correia, F. Murai, A. Veloso and F. Benevenuto, "Supervised Learning for Fake News Detection," in *IEEE Intelligent Systems*, vol. 34, no. 2, pp. 76-81, March-April 2019, doi: 10.1109/MIS.2019.2899143.

[10]. K. Shu, A. Sliva, S. Wang, J. Tang, and H. Liu, “Fake news detection on social media: A data mining perspective,” ACM SIGKDD Explorations Newslett., vol. 19, no. 1, pp. 22–36, 2017.

[11]. E. Cambria, S. Poria, A. Gelbukh, and M. Thelwall, “Sentiment analysis is a big suitcase,” IEEE Intell. Syst., vol. 32, no. 6, pp. 74–80, Nov./Dec. 2017.