



IDENTIFYING IN-ARTICLE ATTRIBUTION AS A SUPERVISED LEARNING ESTIMATOR BY CLASSIFYING FAKE NEWS ARTICLES USING NATURAL LANGUAGE PROCESSING

A.Amrutavalli¹, K. Pavitha², Ch. Hemalatha³, G. Aparna⁴, A. Mallikarjuna⁵

¹ Asst. Professor, Krishna Chaitanya Institute of Technology & Sciences , Markapur, A.P, India

^{2,3,4,5} Scholar, Krishna Chaitanya Institute of Technology & Sciences , Markapur, India

ABSTRACT:

The deliberate misrepresentation of news under the appearance of reputable journalism is a global issue with information accuracy and integrity that has an impact on people's decision-making, voting, and opinion-forming processes. The majority of allegedly "fake news" is first disseminated via social media platforms like Facebook and Twitter before making its way to more established media outlets like traditional television and radio news. Key linguistic traits of the fake news reports that are originally disseminated through social media platforms include an overuse of unsupported exaggeration and unattributed quoted information. The performance of a fake news classifier is examined through the results of a study on the identification of false news. A revolutionary fake news detector that employs cited attribution in a Bayesian machine learning system as a major feature to assess the risk that a news story is fraudulent was created using the Textblob, Natural Language, and SciPy Toolkits. The accuracy of the resulting algorithm is 63.333 percent in determining if a quote-heavy article is likely to be fabricated. Influence mining is the process that enables the identification of propaganda and fake news, and it is touted as an innovative tool for doing so. The classifier performance and findings, technical analysis, technical linguistics work, and the research procedure are all described in this study. The description of how the existing system will transform into an influence mining system serves as the paper's conclusion.

Keywords : Fake News, Machine Learning, Natural Language Processing, Attribution Classification, Influence Mining

[1] INTRODUCTION

It's usually referred to as "fake news," and it influences how people develop opinions, make decisions, and cast ballots. Intentionally false material is disseminated under the appearance of professional journalism. The majority of false information is first spread through social media channels like Facebook and Twitter before making its way to traditional news outlets like television and radio. Key language traits common to fake news articles that are first disseminated on social media platforms include an overuse of unsupported exaggeration and unattributed cited information. This project presents and discusses the findings of a fake news identification investigation that documents the effectiveness of a fake news classifier.

[2] LITERATURE SURVEY

M. Balmas et. al., studied and examined potential relationships between exposure to political humour, or false news, and feelings of ineptitude, alienation, and cynicism toward political candidates. The study shows evidence for an indirect beneficial effect of fake news watching in developing the sentiments of inefficacy, alienation, and cynicism through the mediator variable of perceived realism of false news, using survey data gathered during the 2006 Israeli election campaign. The relationship between watching false news and its perceived reality in this process is moderated by hard news consumption. It was also shown that persons with high exposure to fake news and limited exposure to hard news are more likely to regard fake news as real than individuals with high exposure to both types of news. Overall, this study advances scientific understanding of how different media forms combine to have an impact on political outcomes.

Miley Cyrus gave a twerk-heavy performance on the VMAs programme, and CNN put the item at the top of their website. Then a parody essay defending this choice was published by The Onion, a fake news outlet, under the guise of CNN's Web editor. This research reveals through textual analysis how a Fifth Estate made up of bloggers, columnists, and fake news organisations fought to re-locate mainstream journalism back to inside its professional bounds.

P. R. Brewer et. al., studied and examined how news coverage of political satire affects audience members by drawing on studies on political humour, press meta coverage, and inter textuality. The research tests whether press coverage of Stephen Colbert's Super PAC affected knowledge and opinions on Citizens United, political trust, and internal political efficacy using experimental data. It examines if prior exposure to regular news and Colbert's satirical television show, The Colbert Report, has any impact on these impacts. According to the findings, exposure to news coverage of satire can affect people's knowledge, beliefs, and level of political trust. Additionally, when reading news articles on topics previously featured in satirical programmes, habitual satire viewers may experience larger effects on opinion as well as enhanced internal efficacy.

M. Haigh, T. Haigh, and N. I. Kozak et. al discussed in terms of how people consume news, social media is serving as a two-edged sword for the entire world. On the one hand, individuals use social media to obtain news because of its accessibility, popularity, and inexpensive dissemination method. On the other side, it serves as a vehicle for the dissemination of "false news." The widespread dissemination of false information on social media and websites has a harmful effect on society. This makes it crucial to educate people and fight the spread of fake news. In this essay, we provide a review that identifies the sources, types, production, motivation, and instances of false news. Some methods are also provided for identifying and halting the spread of false information.

R. Marchi et. al., Teenagers, an understudied cohort in studies on young people and the news media, are the focus of this article's examination of news habits and opinions. It analyses how teenagers learn about current events and why they choose some news forms over others. It is based on interviews with 61 ethnically diverse high school students. The findings show changed ideas about what it means to be informed as well as a preference among young people for opinionated news over objective news. This shows that young people want more genuine representations of the fundamental principles of professional journalism rather than a contempt for them.

H. Allcott and M. Gentzkow et. al., discussed After the 2016 US presidential election, many people have voiced worry over the consequences of untrue articles ("fake news"), which are mostly spread through social media. We talk about the economics of false news and give a fresh look at how it was consumed in the run-up to the election. Using information from recent online surveys, fact-checking website archives, and statistics from web browsing, we discover: 1) 14 percent of Americans cite social media as their "most important" source of election news, indicating that it was a significant but not dominant source of news; 2) of the known false news stories that surfaced in the three months leading up to the election, those favouring Trump were shared on Facebook 30 million times, while those favouring Clinton were shared 8 million times; 3) In the months leading up to the election, the average American adult saw one or possibly several fake news stories, with just over half of those who remembered seeing them believing them. 4) People are much more likely to believe stories that support their preferred candidate, especially if they have ideologically segregated social media networks.

C. Shao, G. L. Ciampaglia, O. Varol, A. Flammini, and F. Menczer et. al., has been claimed that the widespread dissemination of false news might affect elections and endanger democracy, making it a significant worldwide concern. While search and social media companies are starting to implement countermeasures, communication, cognitive, social, and computer scientists are working to understand the complicated causes of the viral dissemination of digital disinformation and to propose remedies. However, anecdotal evidence rather than systematically collected data has largely guided these initiatives up to this point. In this article, we examine 14 million tweets from the 2016 U.S. presidential campaign and election that propagated 400 thousand allegations. We discover proof that social bots are crucial to the propagation of false information. Significantly more accounts that deliberately disseminate false information are likely to be automated. Automated accounts often target important people and are particularly active during the first stages of the dissemination of viral claims. Humans are easily tricked by this manipulation by retweeting bots that spread misleading information. Social bots are one of the most effective sources of untrue and biased statements. These findings imply that limiting social bots may be a successful tactic for reducing the spread of false information online.

A. Gupta et. al., discussed in Online social media, especially during times of crisis, is incredibly important in today's globe. Social media coverage of crises has both good and bad consequences; it may be utilised by the government to manage disasters effectively or by evil actors to propagate rumours and false information. The purpose of this essay is to draw attention to how Twitter helped spread false information regarding Hurricane Sandy (2012). During Hurricane Sandy, we found 10,350 distinct tweets that contained fraudulent photos and were shared on Twitter. To comprehend the temporal, social reputation, and influence patterns for the dissemination of fraudulent photographs, we conducted a characterisation study. Few original tweets were used to propagate the fraudulent photographs, since 86% of them were retweets. Our findings indicated that the top 30 users (0.3%) out of 10,215 people were responsible for 90% of the retweets of phoney pictures; network ties, such as Twitter follower relationships, contributed relatively little (only 11%) to the dissemination of these bogus photo URLs. Next, we employed classification models, to identify phoney photographs from actual images of Hurricane Sandy. The Decision Tree classifier produced the best results, with a 97 percent accuracy rate in distinguishing between actual and fraudulent photos. Additionally, user-based features performed very poorly compared to tweet-based features when it came to telling phoney photos

tweets from from real ones. Our findings demonstrated that authentic photographs from false images posted on Twitter may be distinguished using automated algorithms.

E. Mustafaraj and P. T. Metaxas are suggested , 2010's Web Science 2010 Conference's Best Work Award was to a paper with the title "From Obscurity to Prominence in Minutes: Political Speech and Real-time search." Its discoveries included the identification and recording of a "Twitter-bomb," a planned campaign to disseminate false information on the democratic candidate Martha Coakley via anonymous Twitter accounts. In this essay, we summarise the specifics of that incident and then describe the process through which false information is disseminated on social media. One of the most crucial phases in such a recipe is the "infiltration" of a user community that is already conversing about a subject in order to employ them as natural disinformation propagators in their extended subnetworks. After that, we demonstrate how this recipe for disinformation was utilised effectively to propagate false information during the 2016 U.S. Presidential Election. The usage of Facebook rather than Twitter and the various reasons are the two key variances between the situations (in 2010: political influence; in 2016: financial benefit through online advertising). We take use of this opportunity to address study results' shortcomings and to start a discussion about how communities of researchers might have a greater influence on real-world social challenges after placing these occurrences in the larger framework of utilising the Web.

M. Farajtabar et al. discussed, By fusing reinforcement learning with a point process network activity model, we put forth the first multistage intervention method to address bogus news in social networks. A multivariate Hawkes process with extra exogenous control factors is used to represent the dissemination of fake news and mitigating events within the network. We translate the challenge of false news mitigation into the reinforcement learning framework by selecting a feature representation of states, defining mitigation actions, and creating reward functions to gauge the success of mitigation operations. With the aim of optimising the actions for the maximum total reward under budget restrictions, we create a policy iteration approach specific to the multivariate networked point process. Our approach beats alternatives on artificial datasets and demonstrates promising performance in real-time intervention studies on a Twitter network to counter a surrogate false news campaign.

[3] SYSTEM ARCHITECTURE

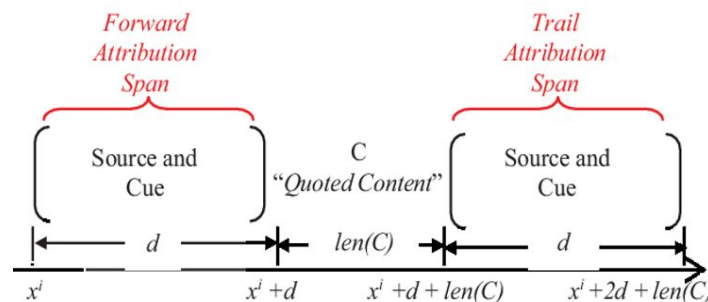


Fig. 1: System Architecture

[4] IMPLEMENTATION

4.1 Modules Description

i) **Social Media Mining System Construction** : We create the system for the assessment of our suggested model in the first module, which includes the system development module with social media mining system. Our topic package space is an expansion of subjects' textual descriptions, like ODP. The topical package space is used to assess how comparable the route topical model package and the user topical model package

are (route package). In this work, we combine two forms of social media—travel diaries and photographs from user-contributed communities—to create the thematic package space. Travelogues are used to mine representative tags, distribution of cost, and visiting time of each topic, and community-contributed photographs are used to mine distribution of visiting time of each topic, in order to create topical package space.

The benefits of combining social media are that: (1) travelogues are more thorough in describing a location than tags with photos, which are rife with noise; (2) it is challenging to determine a user's consumption capacity and the cost of POIs directly from photos or tags with photos; and (3) despite the fact that both media could provide accurate visiting season information of POIs, the number of photos of a POI is much higher than the number of travelogues. (4) The "data taken" of user-contributed images of the places they visit cause the taken time to be erroneous due to the time difference between where they reside and where they took the photos.

ii) User Topical Package Model Mining

- ❖ The mapping of user's photo tags to topical package space yields the user topical package model (user package). The distribution of user subject interests (U), user consumption capacity (U), user preferred travel times (U), and user preferred travel seasons (U) are all included.
- ❖ This module explains how to extract the user package, which includes information about the user's subject interests, consumption capabilities, desired trip times, and preferred travel seasons.
- ❖ We begin by outlining user's topical interest mining, which is derived via mapping user tags to the topical package space. The topical space mapping approach is then explained.
- ❖ We display the user's trip preferences of various themes by mapping the textual descriptions (tags) of user community images to the topical package space, which is defined as user topical interest distribution. When a user's tags show frequently in one topic and infrequently in others, we infer that the person is more interested in that topic.

We demonstrate a user's consumption capabilities using the cost distributions across all subjects and the distribution of topical interest. A user's capacity for consumption is likely to be higher if he often partakes in opulent pursuits like golf and spa visits. If a person often engages in inexpensive activities, his capacity for consuming is likely to be poor, and we avoid advising him on opulent subjects.

iii) Route Package Mining

- ❖ The mapping of travelogues pertaining to the POIs along the route to topical package space allows for the learning of the route topical package model (route package). It includes information on the thematic interest, cost, time, and season distributions of the route.
- ❖ We mine travel routes and the attributes of the routes offline to reduce the amount of time needed for online computation. After mining POIs to create travel routes, we examine the spatio-temporal organisation of the POIs within the records of passengers.
- ❖ According to the "data collected," we create the POIs' spatiotemporal structure. The POI whose timestamp is earlier is referred to as being "in." On the other hand, a POI with a later timestamp is considered "out." After filtering, we tally the number of times users "in" and "out" of POI from other users using their records. The temporal sequence of these POIs is then determined by using a greedy method. Thus, we complete the mining of renowned routes and acquire famous routes for each city.

iv) Travel sequence recommendation

- ❖ In this module, we create our suggestion module for travel routes after mining user and route packages. It consists of two key steps: (1) routes ranking based on user package similarity to routes packages, and (2) route optimization based on records of comparable social users.
- ❖ We receive a set of rated routes following the POI and route rating modules. Here, we go into more detail on how the top routes were optimised based on the trip histories of people with social connections. First, we describe the process for mining socially comparable people's trip histories. Then we go through ways to improve roads based on social users' journey histories.

4.2 Training loss and Accuracy:

Using Natural Language Processing to Classify Fake News Articles and Identify In-Article Attribution as a Supervised Learning Estimator

The notion of identifying false news from social media or document corpora using Natural Language Processing and an attribution supervised learning estimator is described in this work. News articles or documents will be submitted to the programme, which will then use Natural Language Processing to extract quotations, verbs, and name entity recognition (extracting organisations or person names) from the materials to compute scores for each. We will compute the score between the sum of verbs, the sum of name entities, and the amount of quotations divided by the length of the entire sentence using the supervised learning estimator. If the score is more than 0, the news is considered to be TRUE, and if it is less than 0, it is considered to be FAKE.

Three techniques will be used to calculate score

- 1) Source: Any news reporter will identify himself or the person for whom he is writing a story.
- 2) CUE: with the help of this, we may extract VERBS or VERBS phrases; if the news is accurate, it will contain verb-type terms.
- 3) Quotations: Each article will focus on a certain subject, and the subject's name will be described in quotes. So, to identify if news is false or not, we will check for quotations in articles.

Document examples, "We didn't simply slow down our government's rate of expansion while "Mitt Romney" was governor of Massachusetts; we really slashed it," The text above contains quotes about "Mitt Romney" and uses verbs like "was," "didn't," "slow," and "cut." We can determine if news is FAKE or REAL by examining the aforementioned three elements from articles. Since none of the FAKE individuals will make such claims in their papers, we can identify them by using this method. We are utilising the "News" dataset to carry out this research, and by using the aforementioned approach, we can determine if the news is legitimate or phoney. I stored this dataset in a dataset folder. Upload this dataset while your programme is running. After running code in web browser will get below page.

4.3 Screenshots



Fig. 2 Home Page

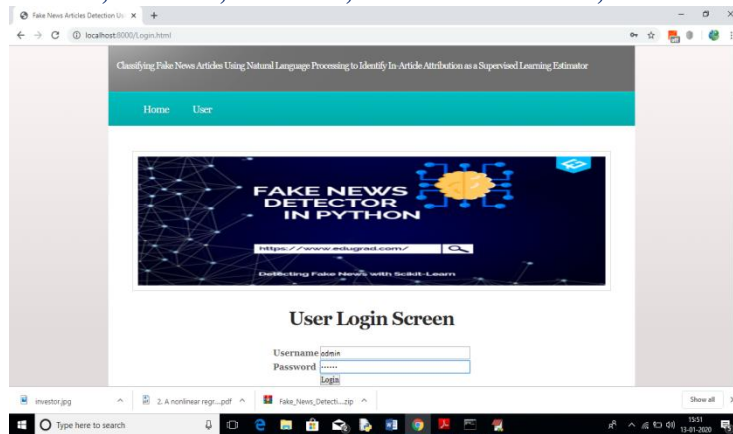


Fig. 3 User Login Page

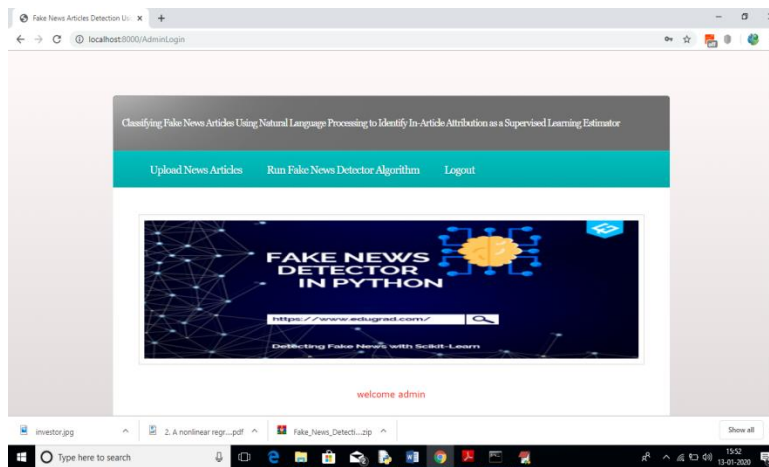


Fig. 4 Upload News Articles ink

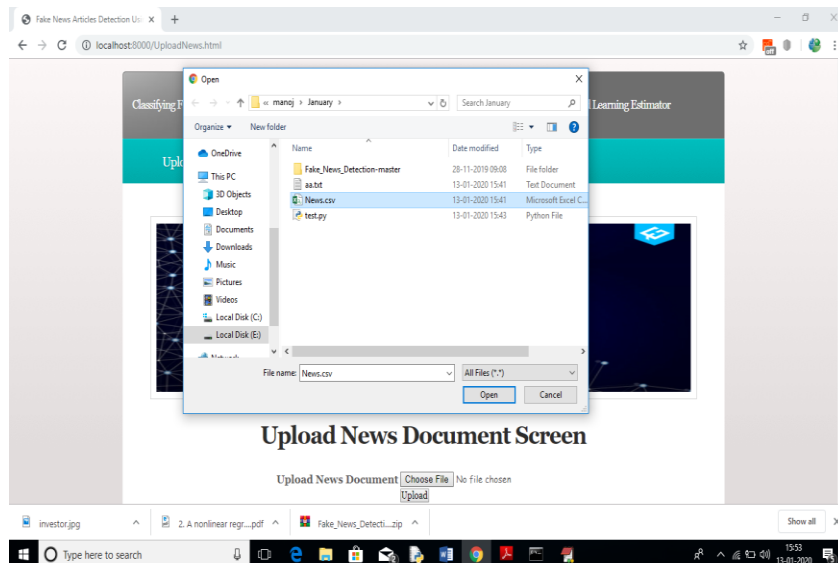


Fig. 5 Uploading 'News.csv' file

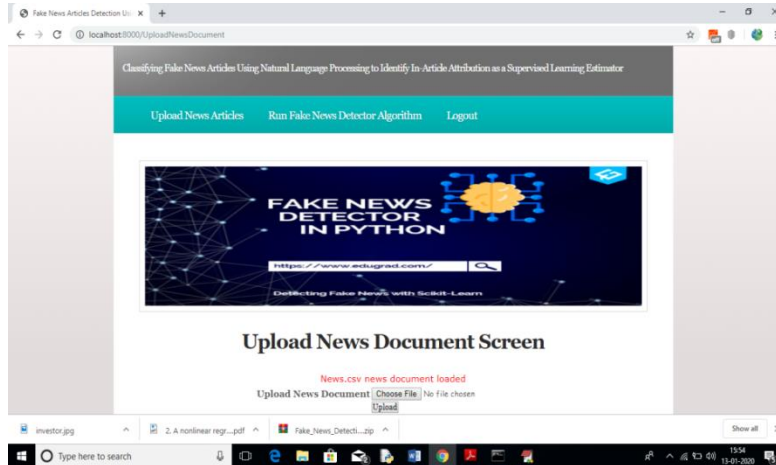


Fig. 6 Uploading 'News.csv file and Run Fake News Detector Algorithm

The link to calculate Fake News Detection algorithm score and based on score and naïve bayes algorithm we will get result.

News Text	Detection Result	Fake Rank Score
Says the Annies List political group supports third-trimester abortions on demand.	Fake News	0.8333333333333333
When did the decline of coal start? It started when natural gas took off that started to begin in (President George W.) Bushs administration.	Real News	2.142857142857143
"Hillary Clinton agrees with John McCain" by voting to give George Bush the benefit of the doubt on Iran."	Real News	3.076923076923077
Health care reform legislation is likely to mandate free sex change surgeries.	Fake News	0.7692307692307693
The economic turnaround started at the end of my term.	Real News	0.9090909090909092
The Chicago Bears have had more starting quarterbacks in the last 10 years than the total number of tenured (UW) faculty fired during the last two decades.	Real News	1.3333333333333333
Jim Dunnam has not lived in the district he represents for years now.	Real News	2.142857142857143
"I'm the only person on this stage who has worked actively just last year passing, along with Russ Feingold, some of the toughest ethics reform since Watergate."	Real News	1.5151515151515151
"However, it took \$19.5 million in Oregon Lottery funds for the Port of Newport to eventually land the new NOAA Marine Operations Center-Pacific."	Real News	2.142857142857143
Says GOP primary opponents Glenn Grothman and Joe Leibham cast a compromise vote that cost \$788 million in higher electricity costs.	Real News	2.1739130434782608
"For the first time in history, the share of the national popular vote margin is smaller than the Latino vote margin."	Fake News	0.8
"Since 2000, nearly 12 million Americans have slipped out of the middle class and into poverty."	Real News	1.5
"When Mitt Romney was governor of Massachusetts, we didnt just slow the rate of growth of our government, we actually cut it."	Real News	2.2222222222222223
The economy bled \$24 billion due to the government shutdown.	Fake News	0.8333333333333333
Most of the (Affordable Care Act) has already in some sense been waived or otherwise suspended.	Real News	2.1052631578947367
"In this last election in November, ... 63 percent of the American people chose not to vote, ... 80 percent of young people, (and) 75 percent of low-income workers chose not to vote."	Real News	0.975609756097561

Fig. 7 Result value - fake or real

In above screen first column contains news text and second column is result value as 'fake or real' the and third column contains score. If score greater > 0.90 then I am considering news as REAL otherwise fake.

News Text	Classification	Score
Some neighborhood schools are closing.	Real News	0.3333333333333333
He told gay organizers in Massachusetts he would be a stronger advocate for special rights than even Ted Kennedy.	Real News	0.5
The year that I was speaker, the Florida House consistently offered leaner budgets than the governor offered.	Real News	0.380952380952381
We are already almost halfway to our 2010 goal of creating 700,000 new jobs in seven years.	Real News	0.5
Says the U.S. Supreme Court found that Social Security is not guaranteed.	Real News	0.8461538461538461
Sen. Michael Bennet wants to close Guantanamo Bay prison and bring terrorists right home to Colorado.	Real News	0.6666666666666666
Oregonians have an amazing no-cost way to fight abortion with free political donations.	Fake News	0.7692307692307693
The president said he was going to bring in 200,000 (Syrian and Iraqi) refugees into this country.	Real News	0.380952380952381
Research shows that a vast majority of arriving immigrants today come here because they believe that government is the source of prosperity, and that's what they support.	Real News	0.6190476190476191
Newt Gingrich's immigration plan offers a new doorway to amnesty.	Real News	0.8181818181818181
Mr. Caprio is a career politician who has never worked in the private sector.	Real News	0.0
In Rhode Island, 9 percent of workers use the states temporary disability insurance program each year while in New Jersey, the rate is only 3 percent.	Real News	0.2903225806451613
In just 17 years, spending for Social Security, federal health care and interest on the debt will exceed ALL tax revenues!	Fake News	0.7692307692307693
President Obama took more money from Wall Street in the 2008 campaign than anybody ever had.	Real News	0.3329411764705882
Donald Trump has said nuclear proliferation is OK.	Real News	0.3333333333333333
Hillary Clinton has taken over \$800,000 from lobbyists.	Real News	0.5
Barack Obama has never even worked in business.	Real News	0.3333333333333333
Says the Arizona immigration law expressly bans racial profiling.	Real News	0.0
Says Gov. Rick Perry has been begging for the federal government to send the Coast Guard to patrol two lakes on the U.S.-Mexico border.	Real News	0.9230769230769231
On the VA: Over 300,000 veterans have died waiting for care.	Real News	0.6666666666666666

Fig. 8 News text articles result - fake or real

[5] CONCLUSION

The findings of a study that resulted in a crude false news detecting system were provided in this publication. The work given here is unique in this field of study since it shows the outcomes of a full-spectrum investigation that began with qualitative observations and led to a functional quantitative model. The research reported in this study is especially encouraging since it shows that massive fake news publications may be classified using machine learning at a pretty high degree of effectiveness using just one extraction characteristic. Furthermore, effort is being done to discover and develop new grammars for categorising fake news, which should result in a more sophisticated categorization system for both false information and direct quotations. To produce tools that not only identify potentially false content but also influence-based content intended to persuade a reader or target audience to make inaccurate or altered decisions, future research plans call for combining attribution feature extraction with additional factors that emerge from the research.

REFERENCES

- [1] H. Liu, T. Mei, J. Luo, H. Li, and S. Li, "Finding perfect rendezvous on the go: accurate mobile visual localization and its applications to routing," in Proceedings of the 20th ACM international conference on Multimedia. ACM, 2012, pp. 9–18.
- [2] J. Li, X. Qian, Y. Y. Tang, L. Yang, and T. Mei, "Gps estimation for places of interest from social users' uploaded photos," IEEE Transactions on Multimedia, vol. 15, no. 8, pp. 2058–2071, 2013.
- [3] S. Jiang, X. Qian, J. Shen, Y. Fu, and T. Mei, "Author topic model based collaborative filtering for personalized poi recommendation," IEEE Transactions on Multimedia, vol. 17, no. 6, pp. 907–918, 2015.
- [4] J. Sang, T. Mei, and C. Sun, J.T. and Xu, "Probabilistic sequential pois recommendation via check-in data," in Proceedings of ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems. ACM, 2012.
- [5] Y. Zheng, L. Zhang, Z. Ma, X. Xie, and W. Ma, "Recommending friends and locations based on individual location history," ACM Transactions on the Web, vol. 5, no. 1, p. 5, 2011.
- [6] H. Gao, J. Tang, X. Hu, and H. Liu, "Content-aware point of interest recommendation on location-based social networks," in Proceedings of 29th International Conference on AAI. AAI, 2015.
- [7] Q. Yuan, G. Cong, and A. Sun, "Graph-based point-of-interest recommendation with geographical and temporal influences," in Proceedings of the 23rd ACM International Conference on Information and Knowledge Management. ACM, 2014, pp. 659–668.
- [8] H. Yin, C. Wang, N. Yu, and L. Zhang, "Trip mining and recommendation from geo-tagged photos," in IEEE International Conference on Multimedia and Expo Workshops. IEEE, 2012, pp. 540–545.
- [9] Y. Gao, J. Tang, R. Hong, Q. Dai, T. Chua, and R. Jain, "W2go: a travel guidance system by automatic landmark ranking," in Proceedings of the international conference on Multimedia. ACM, 2010, pp. 123–132.
- [10] X. Qian, Y. Zhao, and J. Han, "Image location estimation by salient region matching," IEEE Transactions on Image Processing, vol. 24, no. 11, pp. 4348–4358, 2015.
- [11] Sreedhar Jinka, Sk. Althaf Hussain Basha, Suresh Dara, Baijnath Kaushik, "Sequence Labelling for Three Word Disambiguation in Telugu Language Sentences", International Journal of Scientific Research in Computer Science,

- Engineering and Information Technology (IJSRCSEIT) Volume 2 , Issue 7 , pp.311-315, 2017, ISSN :2456-3307.
- [12] Baijnath Kaushik, Sk. Althaf Hussain Basha, Sreedhar Jinka, D Praveen Kumar, “ Sequence Labelling for Two Word Disambiguation in Telugu Language Sentences”, International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT) Volume 2 , Issue 7 , pp.321-327, 2017, ISSN :2456-3307.
- [13] D. Praveen Kumar, Sk. Althaf Hussain Basha, Sreedhar Jinka, Baijnath Kaushik, A. Jagan, “ Empirical Analysis of Context Sensitive Grammars and Parse Trees for Disambiguating Telugu Language Sentences”, International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT) Volume 2 , Issue 7 , pp.328-331, 2017, ISSN :2456-3307.
- [14] Sk.AlthafHussainBasha,SreedharJinka,BaijnathKaushik,D.PraveenKumar, AJagan, “NLP: Context Free Grammars and Parse Trees for Disambiguating Telugu Language Sentences”, International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT) Volume 2 , Issue 7 , pp.332-337, 2017, ISSN : 2456-3307.
- [15] X. Lu, C. Wang, J. Yang, Y. Pang, and L. Zhang, “Photo2trip: generating travel routes from geo-tagged photos for trip planning,” in Proceedings of the international conference on Multimedia. ACM, 2010, pp. 143–152.
- [16] Y. Zheng, L. Zhang, X. Xie, and W. Ma, “Mining interesting locations and travel sequences from gps trajectories,” in Proceedings of the 18th international conference on World wide web. ACM, 2009, pp. 791–800.
- [17] V. W. Zheng, Y. Zheng, X. Xie, and Q. Yang, “Collaborative location and activity recommendations with gps history data,” in Proceedings of the 19th international conference on World wide web. ACM, 2010, pp. 1029–1038.
- [18] N. J. Yuan, Y. Zheng, X. Xie, Y. Wang, K. Zheng, and H. Xiong, “Discovering urban functional zones using latent activity trajectories,” IEEE Trans. Knowl. Data Eng., vol. 27, no. 3, pp. 712–725, 2015.[Online]. Available: <http://dx.doi.org/10.1109/TKDE.2014.2345405>
- [19] J. Liu, Z. Huang, L. Chen, H. T. Shen, and Z. Yan, “Discovering areas of interest with geo-tagged images and check-ins,” in Proceedings of the 20th ACM international conference on Multimedia. ACM, 2012, pp. 589–598.
- [20] Y. Pang, Q. Hao, Y. Yuan, T. Hu, R. Cai, and L. Zhang, “Summarizing tourist destinations by mining user-generated travelogues and photos,” Computer Vision and Image Understanding, vol. 115, no. 3, pp. 352–363, 2011.
- [21] L. Cao, J. Luo, A. Gallagher, X. Jin, J. Han, and T. Huang, “A worldwide tourism recommendation system based on geo-tagged web photos,” in IEEE International Conference on Acoustics Speech and Signal Processing. IEEE, 2010, pp. 2274–2277.
- [22] H. Huang and G. Gartner, “Using trajectories for collaborative filtering–based poi recommendation,” International Journal of Data Mining, Modelling and Management, vol. 6, no. 4, pp. 333–346, 2014.
- [23] C. Zhang and K. Wang, “Poi recommendation through crossregion collaborative filtering,” Knowledge and Information Systems, pp. 1–19, 2015.
- [24] A. Majid, L. Chen, G. Chen, H. Mirza, and I. Hussain, “Gothere: travel suggestions using geotagged photos,” in Proceedings of the 21st international conference companion on World Wide Web. ACM, 2012, pp. 577–578.
- [25] P M Yohan, Sk Althaf Hussain Basha, B Sasidhar, A.Govardhan , “ Automatic Named Entity Identification and Classification using Heuristic Based Approach for Telugu”, IJCSI(International Journal of Computer Science Issues), Volume 10, Issue 6,November 2013,ISSN: 1694-0814.
- [26] Naga Raju Devarakonda, SK Althaf Hussain Basha, ShaikSubhani,“Outliers Detection in Regression Analysis using Partial Least Square Approach”, ICT and Critical Infrastructure: proceedings of the 48th Annual Convention of Computer Society of India- Springer, Vol II Advances in Intelligent Systems and Computing, Volume 249, pp. 125-135, Visakhapatnam, December 2013,ISBN: 978-3-319-03095-1.
- [27] Jinka Sreedhar SK Althaf Hussain Basha, Pammi Pavan Kumar, “Innovative Techniques and Technologies in Translation in a Multilingual Context -2012”, Third International Conference on Translation, Technology and Globalization in Multilingual Context, ITA, NewDelhi
- [28] Y.SriLalitha, Sk Althaf Hussain Basha, N.Sandhya, A.Govardhan, "Web Usage Mining- Survey ", Published in proceeding of National Conference on "Soft Computing and Artificial Intelligence" (NCSCAI 2009) at Lingaya's University, Faridabad, 222-227, 2009
- [29] P. Lou, G. Zhao, X. Qian, H. Wang, and X. Hou, “Schedule a rich sentimental travel via sentimental poi mining and recommendation,” in Proceedings of the 20th ACM international conference on Multimedia Big Data, 2016.
- [30] M. Xie, L. Lakshmanan, and P. Wood, “Breaking out of the box of recommendations: from items to packages,” in Proceedings of the fourth ACM conference on Recommender systems. ACM, 2010, pp. 151–158.
- [31] SkAlthaf Hussain Basha, Jinka Sreedhar, “Introduction: Behaviour of Cuckoo Algorithm”, Advanced Engineering Research and Applications Vol. VIII (Book Code: AERA), Research India Publications, 978-93-87374-32-4, 2018.
- [32] Sk Althaf Hussain Basha, Jinka Sreedhar, Introduction: Behaviour of BAT Algorithm”, Advanced Engineering

INTERNATIONAL JOURNAL OF COMPUTER ENGINEERING AND APPLICATIONS
VOLUME X, ISSUE II, FEB 2016, WWW.IJCEA.COM, ISSN 2321-3469

Research and Applications Vol. VIII (Book Code: AERA), Research India Publications, 978-93-87374-32-4, 2018.

[33] Jinka Sreedhar, Sk Althaf Hussain Basha, "A Case Study : t - Closeness through Micro aggregation: Strict Privacy with Enhanced Utility Preservation in Sensitive Microdata", Hand book on Computer and Information Technology Vol.II (Book Code: HCIT), Research India Publications, ISBN: 978-93-87374-25-6, 2018.

[34] SkAlthaf Hussain Basha, Jinka Sreedhar, Implementation: Ranking and Suggesting Widespread Items In E-Commerce Applications ", Handbook on Computer and Information Technology Vol. II (Book Code: HCIT), Research India Publications, ISBN: 978-93-87374- 25-6, 2018.

[35] X. Wang, M. Yu, L. Zhang, R. Cai, and W. Ma, "Argo: intelligent advertising by mining a user's interest from his photo collections," in Proceedings of the Third International Workshop on Data Mining and Audience Intelligence for Advertising. ACM, 2009, pp. 18–26.

[36] Q. Hao, R. Cai, C. Wang, R. Xiao, J. Yang, Y. Pang, and L. Zhang, "Equip tourists with knowledge mined from travelogues," in Proceedings of the 19th international conference on World wide web. ACM, 2010, pp. 401–410.

[37] Q. Hao, R. Cai, X. Wang, J. Yang, Y. Pang, and L. Zhang, "Generating location overviews with images and tags by mining user generated travelogues," in Proceedings of the 17th ACM international conference on Multimedia. ACM, 2009, pp. 801–804.

[38] K. Sparck Jones, "A statistical interpretation of term specificity and its application in retrieval," Journal of documentation, vol. 28, no. 1, pp. 11–21, 1972.

[39] J. Hays, A. Efros et al., "Im2gps: estimating geographic information from a single image," in IEEE Conference on Computer Vision and Pattern Recognition, 2008. IEEE, 2008, pp. 1–8.

[40] Y. Xue and X. Qian, "Visual summarization of landmarks via viewpoint modeling," in 19th IEEE International Conference on Image Processing. IEEE, 2012, pp. 2873–2876.