VISVESVARAYA TECHNOLOGICAL UNIVERSITY

"Jnana Sangama", Belagavi, Karnataka-590018

FAKE TWEET DETECTION USING MULTINOMIAL NAIVE BAYES CLASSIFIER ALGORITHM (CEC/CS/2022/P26)

A LITERATURE REVIEW REPORT

Submitted by

S SREENIVASA SHENOY (4CB19CS087)
P PADMAPRASAD SHENOY (4CB19CS064)
SUBRAMANYA A SHET (4CB19CS105)
SUHAS S KAMATH (4CB19CS106)

In the partial fulfillment of the requirement for the degree of
BACHELOR OF ENGINEERING
IN
COMPUTER SCIENCE & ENGINEERING
Under the guidance of
MR. SHATANANDA BHAT P
Assistant Professor

Department of Computer Science & Engineering
CANARA ENGINEERING COLLEGE
BENJANAPADAVU, BANTWAL - 574219, D.K., KARNATAKA
2022-2023

CANARA
ENGINEERING COLLEGE
BANTWAL, D.K.574219 – KARNATAKA

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

CERTIFICATE

This is to certify that **"Mr. S Sreenivasa Shenoy (4CB19CS087), Mr. P Padmaprasad Shenoy (4CB19CS064), Mr. Subramanya A Shet (4CB19CS105) and Mr. Suhas S Kamath (4CB19CS106)"** has successfully completed the PROJECT WORK PHASE – I (Literature Survey) of the project entitled ***"FAKE TWEET DETECTION USING MULTINOMIAL NAIVE BAYES CLASSIFIER ALGORITHM"*** under the guidance of **Prof. Shatananda Bhat P.** The project report has been approved as it satisfies the academic requirements in respect of Project work.

Signature of the Guide

(**Prof. Shatananda Bhat P)**

Abstract

Nowadays, online social media plays a vital role in real-world applications with potentially positive and negative effects. When some event has occurred, many of us discuss it online through social networking. Once sudden events happen there's conjointly faux news that's broadcasted that makes confusion because of the character of the events. This is important because fake tweets can spread misinformation and cause panic.

The dataset used in this study will be a dataset of tweets that have been labeled as fake or not fake. The machine learning model will be trained on this dataset and then tested on a separate dataset. The model will be evaluated on its ability to correctly classify fake tweets.

TABLE OF CONTENTS

CONTENTS	PAGE NO.
Abstract	iii
1. Introduction	01
2. Literature Review	
2.1 Fake News Detection using Machine Learning	02
2.2 The Detection of Fake Messages using Machine Learning	03
2.3 Can Fake News Detection Models Maintain the Performance through Time? A Longitudinal Evaluation of Twitter Publications	04
2.4 A Novel Stacking Approach for Accurate Detection of Fake News	05
2.5 An Improved Multiple Features and Machine Learning-Based Approach for Detecting Clickbait News on Social Networks	06
2.6 Self Multi-Head Attention-based Convolutional Neural Networks for fake news detection	07
2.7 Fake Sentence Detection Based on Transfer Learning: Applying to Korean COVID-19 Fake News	08
2.8 A comprehensive Benchmark for fake news detection	09
2.9 Collecting a Large Scale Dataset for Classifying Fake News Tweets Using Weak Supervision	10
2.10 Deep Learning for Fake News Detection: A comprehensive Survey	11
2.11 Analyzing User Profiles for Detection of Fake News Spreaders on Twitter	12
2.12 Combining Machine Learning with Knowledge Engineering to detect Fake News in Social Networks-a survey	13

2.13 Detecting Fake News Spreaders in Social Networks via Linguistic and Personality Features 14

2.14 A Benchmark Study on Machine Learning Methods for Fake News Detection 15

2.15 Analyzing User Profiles for Detection of Fake News Spreaders on Twitter 16

2.16 Fake News Detection on Social Media using Geometric Deep Learning 17

2.17 Credibility Analysis on Twitter Considering Topic Detection. 18

2.18 Fake News Analysis Modelling Using Quote Retweet 19

2.19 Fake News Data Exploration and Analytics 20

2.20 Deep Learning for Fake News Detection in a Pairwise Textual Input Schema 21

References 22

Chapter 1

Introduction

Fake News is one of the most controversial stories that has attracted attention over the past year. Social media is a powerful tool for spreading lies. Modern life has become much more relevant and people around the world should appreciate the great contribution of internet technology to the transmission and sharing of information. There is no doubt that the internet has made our lives easier and access to more information has worked. This is an evolutionary process in human history, but at the same time, it does not focus on the line between real media and brutally constructed media. Internet users can often follow their favorite events in online mode and the proliferation of mobile devices makes this process even easier.

Many scientists believe that counterfeit stories can be solved by means of machine learning and artificial intelligence. This is because recently artificial intelligence algorithms have begun to improve the function in many classification problems (image recognition, voice detection, and so on) because the hardware is cheaper and larger databases are available.

Although many false stories are flawless, they are used for entertainment purposes only but readers do not understand the truth of these stories and change them according to the context. It is therefore very difficult for students to understand the motto of stories whether they are issued for entertainment purposes or for any other purpose. That is why it is so important to create such a model that can easily reflect the theme of the story so that readers are not distracted.

Chapter 2

LITERATURE REVIEW

2.1 Fake News Detection using Machine Learning [1]

Fake News is one of the most controversial stories that has attracted attention over the past year. The media reports that social media has played a key role in the outcome of the 2016 US elections. It shows a simple technique for fake news detection using the Naive Bayes classifier. This approach was executed as a software system and tested against a dataset of Facebook news posts. The authors have used a Logistic Regression classifier for the classification of fake news. Machine (SVM) and resolved whether the stories were true or false. Author said that a support vector machine algorithm for binary classification to organize the articles and based on that model work to categorize the articles as either fake or real. The authors have used three main modules to purify their articles in their proposed models: an aggregator, authenticator, and recommendation system. Author also used the Naïve Bayes algorithm to check whether the articles were fake or real and to obtain 93.50% accuracy achieved by combining these three algorithms i.e., Naive Bayes, SVM, and NLP.

There are many existing approaches for detection of fake news. But the author has used three types of existing ML Algorithms which are used for detection of fake news.

1. Naïve bayes
2. Logistic regression
3. Support vector machine

It is significant to identify whether a news article is fake or real.Machine learning-based classification algorithms play a very important role in the detection of fake news from social media, which is a very complicated and difficult process due to the diverse social, political and economic, and many other related factors. The paper gives different approaches for detecting fake news.

2.2 The Detection of Fake Messages using Machine Learning [2]

The proposed research focuses on the social media platform Twitter. Author has told about how the investigation is done to detect fake messages using machine learning. The research will focus on the Dutch election of 2012. To investigate to what extent fake messages have been used during the Dutch election of 2012, they formulated the following research questions:

- Can we train a classifier to detect potential fake media regarding the Dutch election of 2012?
- What kind of fake messages have been used during the Dutch election period of 2012 on Twitter?

The database consists of tweets posted around the Dutch election of September 12, 2012. The classifier was trained on only the text of the tweet.

Data Selection and Gathering: Author researched how online citizens persuaded fellow voters in the Dutch election of 2012. They Used the logic of the snowball sampling method to gather relevant hashtags. A hashtag—written with a # symbol—is used to index messages on Twitter. It allows people to follow topics easily according to their interests.

Training of Classifiers: They used Machine Learning algorithms in this process.

1.Data sampling

2.Pre-processing of data

3.Machine Training

For this research, eight different supervised Machine Learning algorithms have been analyzed.The Python package Scikit-learn was used for this study. This package was selected because it contains multiple implementations of the Naïve Bayes algorithms. Author used the following algorithms: Linear Support Vector Machines (LSVM), Naïve Bayes (NB), Decision Trees (DT), ExtraTrees (ET), Stochastic Gradient Descent (SGD), and Random Forests (RF). Of Naïve Bayes, three different implementations have been used: Gaussian Naïve Bayes (GNB), Bernoulli Naïve Bayes (B-NB), and Multinomial Naïve Bayes (M-NB)**.** As a result, they conclude that the Decision Tree algorithm is the best algorithm for the classification of true and false messages.

2.3 Can Fake News Detection Models Maintain the Performance through Time? A Longitudinal Evaluation of Twitter Publications [3]

Online Social Networks redefined the way we communicate. Since their inception, they evolved from a way to share media and information among small friends networks to an entire medium to consume and share content. Author said they focused on developing fake news detection models in social networks and evaluating their performance over a long-term period. The data used in the training and evaluation of the models spans 18 months of fake and actual news that includes important events favorable to the spreading of false information, such as the 2020 US presidential elections and the rise of the COVID-19 pandemic. Author has said that the definition of fake news is very similar to the one presented. Thus, they defined a post as fake news if it contains at least one reference to a website whose articles are included in one of the following categories:

1. false
2. clickbait
3. bias
4. junksci
5. hate
6. unreliable

Concerning the models and evaluation, they opted by selecting some of the most common models in the literature. Therefore, they used SVMs (radial and linear kernel), Decision Tree, Naive Bayes, and K-Nearest Neighbors (KNN). In addition, they complimented this selection with the following ensemble models: Gradient Boost Classifier (GBC), Random Forest, and AdaBoost.

Author said,they focused on a more pragmatic approach towards the problem of identifying fake news in social networks by assessing the performance of different groups of features and models in a dataset consisting of an 18-month time period of tweets.The results provided in this study contribute towards the development of a more pragmatic fake news detection system that is intended to be used in a real-world scenario.

2.4 A Novel Stacking Approach for Accurate Detection of Fake News [4]

A lot of researchers proposed their models based on machine learning and deep learning techniques. To detect fake news numerous machine learning and deep learning techniques have been recommended by various scholars. Some researchers evaluated a lot of machine learning models on different datasets to choose the best individual model. Ozbay et al.implemented twenty-three supervised artificial intelligence algorithms in three datasets including BayesNet, JRip, OneR, Decision Stump, ZeroR, Stochastic Gradient Descent (SGD), Logistic Model Tree (LMT), etc. The fundamental concepts of the proposed models have been discussed in the below sections.

1.Dataset: This paper involved two datasets, ISOT fake news dataset, and the KDnugget dataset.

2.Pre-processing

3.Machine Learning Classification Models: In this proposed study, for classifying fake news and real news, five machine learning models have been used: LogisticRegression, DecisionTree, K-NearestNeighbor, RandomForest, Support Vector Machine.

4.Deep Learning Models Description: In this proposed study, for classifying fake news and real news, three deep learning models have been used: Convolutional Neural Networks,Long Short Term Memory, Gated Recurrent Unit.

5.Hold Out Cross Validation Method

6.Model Evaluation Criteria

7.McNemar’s Statistical Test

8.Proposed Stacking Mechanism

According to their experiments, some models like K-Nearest Neighbors had better performance on small datasets and other models like Decision Tree, Support Vector Machine, Logistic Regression, CNN, GRU, and LSTM had much worse performance on small datasets. To select the best model, the author used a corrected version of McNemar’s test to determine if the models’ performance is significantly different.

2.5 An Improved Multiple Features and Machine Learning-Based Approach for Detecting Clickbait News on Social Networks [5]

Introduction Currently, social networks have become the main environment for communicating, sharing, and posting news on the Internet. Although social networks provide an easy way to express our opinions, this platform also can be used to share misinformation in the form of news and advertisements. This misinformation can take the form of clickbait, which aims at enticing the users into clicking a link to news items or advertisements, whose titles do not completely reflect the inside contents. Although several machine-learning approaches have been proposed to detect clickbait headlines, most of these recent methods are not very robust. The key contributions of this paper which is mentioned by author are as follows:

- Author constructed the first Arabic clickbait headline news dataset. The raw dataset is available publicly for research purposes.
- Author extracted a set of user-based features and content-based features for the constructed Arabic clickbait dataset.
- Author conducted extensive experiments, and the results show that the proposed model enhances the performance of some classifiers in terms of accuracy, precision, and recall.

The proposed multiple-feature-based approach for detecting clickbait news is presented in this section. Since the difference between clickbait and normal news can be distinguished directly by analysis of the linguistic character of news content, the proposed approach takes into consideration both the headlines and the content of the news features (CFs). They implemented six machine learning-based classifiers, including Random Forest (RF), Stochastic Gradient Descent (SGD), Support Vector Machine (SVM), Logistic Regression (LR), Multinomial Naïve Bayes (NB), and k-Nearest Neighbor (k-NN).

The proposed system gives a comprehensive approach that includes three main phases: data collection, data preparation, and machine learning modeling phases. After collecting the dataset, which is considered the first Arabic clickbait headline news dataset, the pre-processing tasks were performed, which included text cleansing, normalization, stemming, stop words removal, and tokenization. Future work will investigate the application of several deep learning methods on this Arabic dataset in order to enhance detection performance.

2.6 Self Multi-Head Attention-based Convolutional Neural Networks for fake news detection [6]

In recent years, social media has become ubiquitous and most important for social networking and content sharing. In the age of information explosion, without analyzing and judging the massive information effectively, the information will become a burden. Fake news is deceptive, always misleading people. What's worse, it is difficult to distinguish. A study reported that many people typically retweet false information on Twitter.

The attention mechanism can get the relation between the whole and the part in one step and help the classifier grasp the easily overlooked but significant features of the raw data. The attention mechanism was initially applied in the field of image processing; it can effectively capture images' local features. Just like human vision, the partial can generalize the whole, and people may pay more attention to certain parts of the picture.

News has the characteristics of real-time, thus it is necessary to detect fake news as soon as possible. They have summarized two methods for detecting fake news: context-based in fake news detection and content-based in fake news detection.Although the detection results are significant, they are not intuitive enough and have some defects due to the characteristics of the algorithm. For example, the basic CNN architecture can get the local information by the convolution kernels, but because of the characteristics of stride and the size of convolution kernels, it cannot capture the contextual information and the correlations between the widely separated words, it will affect the detection performance of the model.

To purify the environment of the Internet, the detection of fake information is urgent. The author has proposed an effective model to detect fake news only based on news content by using self-multi-head attention-based convolutional neural networks. On detecting a novel topic, the proposed model still had excellent performance.Besides, the experimental result is compared with previous work objectively and the result has shown that the proposed model of using the self multi-head mechanism with the convolutional neural networks indeed can get more outstanding performance.

2.7 Fake Sentence Detection Based on Transfer Learning: Applying to Korean COVID-19 Fake News [7]

Users encounter news from various fields through a variety of platforms due to the development of social media. Unlike how the press and TV networks used to dominate news broadcasting, users of social media can directly participate in the news while the users quickly propagate the news due to the high accessibility and speed. Furthermore, fake news is propagated more quickly. The performance of fake news detection has drastically improved recently through BERT, which undergoes pre-training based on large-scale data.

Due to advancements in BERT, most research on fake news detection was conducted by embedding entire documents. When embedding the entire document of fake news, however, discriminating fake news from real news becomes challenging because fake news maliciously includes truthful sentences.

The paper builds a dataset on COVID-19 fake news collected from multiple fact checking websites and proposes a model for detecting fake sentences related to COVID-19.

BERT, which was released by Google in 2018, demonstrates excellent performance in various tasks as a pre-trained model. The pre-training of BERT involves randomly masking words using a masked language model, which then predicts the masked words, and applying the next sentence prediction to determine whether two sentences are connected. The models pre-trained with large-scale data of the problem being solved are fine-tuned and applied to various tasks. KoBERT demonstrates excellent performance in the problems in Korea, as it was trained with an SKT large-scale Korean Wiki corpus.

The proposed system gives a model for detecting Korean fake news about COVID-19. The datasets were constructed by the researchers because of their non-availability. The proposed model classified sentences into fake or true instead of documents because all sentences in a document are not fake, only one or two sentences. Through experiments, the proposed model was shown to have an accuracy of 78.8%, which was improved by 8% compared with the linear model as a baseline model. However, the experiments showed that a greater amount of data results in a better performance for the BiLSTM model.

2.8 A comprehensive Benchmark for fake news detection [8]

In the last decade, with the disruptive diffusion of social media, people turned towards consuming news from journalistic websites to popular social media platforms. More in detail, social media are nowadays the main medium for large-scale information sharing and communication and they can be considered the main drivers of the Big Data revolution we observed in recent years. Unfortunately, due to the malicious user having fraudulent goals fake news on social media are growing quickly both in volume and potential influence thus leading to very negative social effects. Thus, fake news detection on social media poses particular challenges due to the inherent nature of social networks that requires both the analysis of their content and their social context.

The author has defined a Fake News Detection framework for experimental purposes, based on news flow processing and data management. In particular, a preliminary pre-processing stage executes filtering and aggregation operations over the news content, and in addition, filtered data are processed by two independent modules: the first one performs natural language processing over data, while the second one performs a multimedia analysis. The goal of the Machine Learning Module is to produce a binary classification on a text dataset. Thus, a news article will be labeled as 0 if it is recognized as Real, and as 1 if it is recognized as Fake. They devised a supervised approach since the dataset they worked on is fully labeled. The Machine Learning implementation has been chosen by comparing most of the available classifiers provided by the Scikit- Learn library.

Fake news is a challenging task even though several techniques have been developed over time to mitigate their negative effects. A benchmark analysis of fake news detection using classical Machine Learning and Deep Learning approaches (both for texts and images) has been discussed. Traditional machine learning classifiers still have advantages and some drawbacks. First, these methods are very fast both during training and test steps for supporting real-time analysis. On the other hand, these methods are not still able to detect words' semantic meaning and context of the word picked up from a sentence obtaining low accuracy values.

2.9 Collecting a Large Scale Dataset for Classifying Fake News Tweets Using Weak Supervision [9]

In recent years, fake news shared on social media has become a much-recognized topic. Social media makes it easy to share and spread fake news, i.e., misleading or wrong information, easily reaching a large audience. While other problems of tweet classification, like sentiment detection or topic detection, are rather extensively researched, the problem of fake news detection, although similar from a technical perspective, is just about to gain attention.

The proposed system makes use of two datasets in this work:

1. A large-scale training dataset is collected from Twitter and labeled automatically.
2. A smaller dataset is collected and labeled manually. It is not used for training, only for validation.

Author modeled the problem as a binary classification problem. Author also mentions the usage of Naive Bayes, Decision Trees, Support Vector Machines (SVM), and feed-forward Neural Networks as basic classifiers. Moreover, they used two ensemble methods known to usually work well, i.e. Random Forest and XGBoost. For Neural Networks, the proposed method follows the observation that one hidden layer is sufficient for most classification problems. Optimization in the number of neurons in the hidden layer along with many other parameters, such as activation function, learning rate, regularization penalty, and the maximum number of training epochs using random search is done,as suggested.

The proposed work shows a practical approach for treating the identification of fake news on Twitter as a binary machine-learning problem. While that translation to a machine learning problem is rather straightforward, the main challenge is to gather a training dataset of suitable size. Here, it’s shown that, instead of creating a small, but accurate hand-labeled dataset, using a large-scale dataset with inaccurate labels yields competitive results. The advantages of using our approach are two-fold. First, the efforts for creating the dataset are rather minimal. Second, since the dataset can be created automatically to a large extent, it can update at any point in time, thus accounting for recent topics and trends, as well as changes in the social media service like the change from 140 to 280 characters in Twitter.

2.10 Deep Learning for Fake News Detection: A Comprehensive Survey [10]

The Internet and social media platforms have become indispensable ways for people to obtain news in their daily life. Because they allow the news to be disseminated rapidly and freely, public viewers have unhindered access to consult whenever and wherever they want. As of August 2018, over 68 percent of Americans acquire their news through social media. However, the deficiency of censorship from authority renders the quality of news spread far lower than in the traditional way. The online information ecosystem is extremely noisy and fraught with disinformation and fake news.

Fake news detection aims to identify fake news automatically. Existing traditional ML-based FND methods require feature engineering. Deep learning-based fake news detection. With the success of deep learning in various domains, DL-based FND methods have been proposed and attracted significant attention recently. Most above methods rely heavily on linguistic and semantic features from the news content or social content features. However, they fail to effectively exploit external knowledge, which could help determine whether the news document is trusted. News content is highly condensed and consists of many entity mentions. Due to issues with aliases, abbreviations, and alternative spellings, it is not always possible to understand news text content directly. Thus, several research studies introduce external knowledge to help improve the performance of fake news detection.

Graph-based methods use the similarity between news texts to model the labeled news and the unlabeled news into a single graph. The underlying constraint is: nodes in the graph that are close to each other often have the same label. Based on the above review and analysis, we believe there exists much space for further enhancement in this field. In this section, they discussed the emerging trends for further exploration in FND research. Different types of future directions are:

1. Emergent Events Fake News Detection
2. Multi-domain Fake News Detection
3. Multi-lingual Fake News Detection

2.11 Fake News Detection on Social Media using Machine Learning Techniques [11]

The rise of social media in the past years has changed the way people consume information, especially news. The amount of time spent online, immediacy and lower to non-existent price are decisive factors for this global change in the ways of news consumption. Our society is subjected to a massive exposition of information nowadays, and this has led to a great amount of false or extremely biased information being shared and consumed by Internet users every day. Our proposal focuses on the content created by a user instead of the content created by other users as for example retweets. Then, we use a combination of psychological and linguistic features aimed at modelling the behaviour of a user to detect if they are a spreader or non-spreader.

The dataset provided by the organizers of the task was divided into two collections of tweets: one in Spanish and one in English. The data used is in the English directory in the proposed model. With regards to the pre-processing applied to the data, there are three important steps that we took before the feature extraction:

1. Tokenization
2. Stop words removal
3. Tweet aggregation

The SPOT approach incorporates four main phases, including data transformation, opinion mining, credibility analysis, and fake news detection.

1. Data transformation
2. Opinion mining on user comments
3. Credibility score computation

The proposed model aimed to differentiate fake news spreaders from real news spreaders using a combination of psychological and linguistic traits extracted from the user's data. The paper presented the SPOT approach, a fake news detection method based on semantic knowledge source and deep learning. To detect the false news stories in Twitter effectively, the SPOT method performs the opinion mining on user comments for mining the opinion of the user, which helps to improve the fake news identification.

2.12 Combining Machine Learning with Knowledge Engineering to detect Fake News in Social Networks-a survey [12]

In the news media and social media the information is spread at a high speed. This information is without accuracy and hence detection mechanisms should be able to predict news fast enough to tackle the dissemination of fake news. Fake news and the spread of misinformation have dominated the news cycle after US Presidential elections in 2016. The author has presented what is fake news, importance of fake news, overall impact of fake news on different areas, different ways to detect fake news on social media, existing detections algorithms that can help us to overcome the issue, similar application areas and at the end we proposed combination of data driven and engineered knowledge to combat fake news.

Some of them are to gain political gains or ruin someone else's reputation or for seeking attention. Learning from data and engineered knowledge to overcome fake news issues on social media. To achieve the goal a new combination algorithm approach shall be developed which will classify the text as soon as the news will publish online. Classification of millions of news that is published online manually is time consuming and an expensive task.

Recently after US presidential elections social media often become a trained vehicle of spreading misinformation. Maybe this one has no serious consequences if only to share or spread rumours of less importance but it can be a serious problem when the consumers can purchase products on the basis of these rumours or sometimes serious security issues.

It is better to analyse fake news differently with different measure similarities Eg we can detect whether the same news published by other media agencies or not, We can check the location of the news Maybe a news has a higher probability of being fake, if it is generated somewhere else and not at the location they deal with. Combination of machine learning and knowledge engineering can be useful for fake news detection as it looks like fake news may be the most challenging area of research in the coming years.

2.13 Detecting Fake News Spreaders in Social Networks via Linguistic and Personality Features [13]

The author addresses the problem of automatically detecting fake news spreaders in social networks such as Twitter. Fake news sharing has become a concerning problem in online social networks in recent years. Research has found that fake news is more likely to go viral than real news, spreading both faster and wider and is threatening public health, emergency management and response, election outcomes, and is responsible for a general decline in trust that citizens of democratic societies have for online platforms. It is mainly because people are not able to disguise it from truthful information and often share news online without even reading its content.

The dataset is provided by the PAN'20 shared task for Profiling Fake News Spreaders on Twitter. As people express their emotions, appraisals, and sentiments towards any news or article through the choice of words in their tweets, we leverage sentiment analysis as another feature. For each user, we measured the average sentiment across all their tweets. We used the Valence Aware Dictionary and Sentiment Reasoner (VADER), a library specifically built for capturing sentiments expressed in social media texts, for English tweets and sentiment-analysis-Spanish 5 for Spanish tweets.

Linguistics features refer to features that represent the functionality of text, such as the average number of words per sentence and the rate of misspelling. Personal features target emotional, social processes and cognitive processes. The affective processes (positive and negative emotions), social processes, cognitive processes, perceptual processes, biological processes, time orientations, relativity, personal concerns, and informal language (swear words) can be used to scrutinize the emotional part of the text.

One possible motivation could be that some users keen to spread fake news do not do it intentionally. Hence they are hard to differentiate from users who never shared fake news. It will be devoted to analysing the role of additional sets of features for detecting fake news spreaders in social networks, including behavioural features describing the user activity on Twitter and social network features.

2.14 A Benchmark Study on Machine Learning Methods for Fake News Detection [14]

There was a time when if anyone needed any news, he or she would wait for the next-day newspaper. However, with the growth of online newspapers who update news almost instantly, people have found a better and faster way to be informed of the matter of his/her interest. Nowadays social-networking systems, online news portals, and other online media have become the main sources of news through which interesting and breaking news are shared at a rapid pace. However, many news portals serve special interest by feeding with distorted, partially correct, and sometimes imaginary news that is likely to attract the attention of a target group of people.

The architecture (Character-level C-LSTM) used in the proposed methodology accepts a sequence of encoded characters as input. The encoding is done by prescribing a fixed length of alphabets for the input language and then quantizing each character using one hot encoding. The characters are then transformed into fixed-length vectors.The paper presents an overall performance analysis of different approaches on three different datasets.

1. Liar
2. Combined Corpus
3. Feature Extraction

The proposed system shows that Naive Bayes with n-gram can attain analogous results to neural network-based models on a dataset with less than 100k news articles. Moreover, advanced models like Character level C-LSTM have shown high promise that demands further attention on these models in fake news detection. Finally, we perform a topic based analysis that exposes the difficulty to correctly detect political, health and research related deceptive news. The future plan is to experiment on a larger dataset to find how the traditional model like Naive Bayes competes against highly computational neural network-based models to detect fake news.

2.15 Analyzing User Profiles for Detection of Fake News Spreaders on Twitter [15]

The rise of social media in the past years has changed the way people consume information, especially news. The amount of time spent online, immediacy and lower to non-existent price are decisive factors for this global change in the ways of news consumption. Our society is subjected to a massive exposition of information nowadays, and this has led to a great amount of false or extremely biased information being shared and consumed by Internet users every day.

The proposed system focuses on the content created by a user instead of the content created by other users as for example retweets. Then, we use a combination of psychological and linguistic features aimed at modelling the behaviour of a user to detect if they are a spreader or non-spreader. The dataset provided by the organizers of the task was divided into two collections of tweets: one in Spanish and one in English. With regards to the pre-processing applied to the data, there are three important steps that we took before the feature extraction:

1. Tokenization
2. Stop words removal
3. Tweet aggregation

The main contents that users share in social media can be divided in:

1. content created by the user
2. content created by others.

The proposed model aims to differentiate fake news spreaders from real news spreaders using a combination of psychological and linguistic traits extracted from the user's data, together with characteristics extracted from both the user behaviour in the social network and the news headline analysis. As it can be seen in the results exposed in the previous section, our results in development and test are consistent, which means that, despite the work that needs to be done in order to improve it, it is a robust model with an expectable performance when datasets vary.

The proposed work is at very early stages of development and will continue evolving towards a more efficient and better performing system. Hence preliminary results are shown and there is still room for improvement.

2.16 Fake News Detection on Social Media using Geometric Deep Learning [16]

Social media is one of the main sources of the news around the globe due to easy access and rapid transformation of news. This comes at the cost of dubious trustworthiness and significant risk of exposure to ‘fake news’, intentionally written to mislead thae readers. So detecting fake news poses challenges which is one of the main reasons that often leds to interpretation of the news requires the knowledge of political or social context or ‘common sense’.

Fake news is often intentionally written by bad actors to appear as real news but containing false or manipulative information in ways that are hard even for trained human experts to detect.These rumours and false information has to be analyzed and recognized efficiently Which are missing in the current natural processing algorithms. The proposed system has used learning of fake news specific propagation patterns by exploiting geometric deep learning, a novel class of deep learning methods designed to work on graph-structured data. The system has used a collection of news verified by fact-checking organizations with established reputation in debunking rumors, each source fact-checking organization provides an archive of news with an associated short claim and a label determining its veracity .

Geometric deep learning naturally deals with heterogeneous data such as social network structure, news propagation and content) thus carrying the potential of being a unifying framework for content, social context, and propagation based approaches.The system has presented a geometric deep learning approach for fake news detection on Twitter social network. The proposed method naturally allows integrating heterogeneous data pertaining to the user profile and activity, social network structure, news spreading patterns and content. The key advantage of using a deep learning approach is its ability to automatically learn task-specific features from the data; the choice of geometric deep learning in this case is motivated by the graph-structured nature of the data. Model achieves very high accuracy and robust behavior in several challenging settings involving large-scale real data, pointing to the great potential of geometric deep learning methods for fake news detection.

2.17 Credibility Analysis on Twitter Considering Topic Detection. [17]

Twitter is one of the most popular sources of information available on the internet. Thus,many studies have proposed tools and models to analyze the credibility of the information shared.The credibility analysis on Twitter is generally supported by measures that consider the text, the user,and the social impact of text and user. More recently, identifying the topic of tweets is becoming an interesting aspect for many applications that analyze Twitter as a source of information, for example, to detect trends, to filter or classify tweets, to identify fake news, or even to measure a tweet's credibility.

The motivation of the survey is associated with twitter by the amount of information and is used to spread information and express opinions.This dataset was structured by a tweet id and a topic id.Then topic detection is discovered using the main topics automatically which can help in many applications that analyze. The proposed work extends a credibility model by adding topic analysis for tweets that have hashtags, since currently on Twitter it is very common to use hashtags to somehow label the tweet with words that are trending or relevant.The dataset used for this analysis is a licensed public access database.

The tweet id corresponds to the internal id provided by Twitter and the topic id is the original identifier provided by the dataset between 1 and 5234 associated with an event (or topic for us). The data contain tweets gathered from news headlines from a manually curated list of well-known news media accounts on twitter. The author first compared different topic detection algorithms, evaluated them using precision and stayed with the one that gave the best results, which was NMF.

Even though the NMF was not trained for PHEME dataset, we obtained high score values for events such as Prince Toronto (97.80%), Ebola (greater than 80%),Germanwings (90% threshold), Ferguson (approx 39% F1 score for thresholds greater than 75%) and Sydney Siege (51.90%).the present study shows the feasibility of integrating a topic analysis to our credibility framework and of considering certain associated semantics. However,this concern is still a challenge.

2.18 Fake News Analysis Modelling Using Quote Retweet[18]

Fake news can create confusion in the area of politics, culture, and healthcare. Fake news refers to news containing misleading or fabricated contents. They are intentionally created to provide false information. They can distort reality and cause problems. The proposed study proposes a fake news analysis modelling method by identifying a variety of features and collecting various data from Twitter. The popularity of social media has increased. Social media is the composition of texts as well as multimedia such as photos, videos and so on. If the spread of information in social media is not controlled, there is a risk that current information will spread rapidly.

So the motivation of this survey is to stop or control the spread of false information whose source is unclear. The paper presents a novel approach by using an analysis model to identify the best features of fake news by using information from Quote-Retweet as well as conventional Tweets. So ,not only more text information can be collected compared to conventional. The attributes that showed significant differences between fake news TW and real news TW were the average number of replies, average depth of propagation tree, proportion of including URL, user's influence/activeness/active period, QRT's proportion, proportion of including multimedia, retweets, but also the depth of each can be easily measured.

The numerous amounts of information are generated every day amid advancement of electronic devices and social media, and among such information, false information, called fake news, exists as well. Fake news creates various problems in our society, and endeavours are required to solve them. Accordingly, many researchers have conducted studies to detect rumours and fake news by using Twitter data, one of the popular social media outlets. The results of visualisation and statistical analysis to investigate the best features from the collected data indicated significant differences between fake news and real news in terms of existence/absence of information source, replies for Tweets, influencing power of Tweets, rate of using Quote Retweet, depth of Tweet propagation tree, and rate of quoting picture/video.

2.19 Fake News Data Exploration and Analytics[19]

In the earlier days,people used to acquire their news from the radio,television and newspapers.With the internet,the news was moved to online, and there anyone could post information on websites.Since internet use has taken off dramatically ,fake news is the thing that everyone is very fond of. Social media platforms such as Facebook, instagram etc also. YouTube is one of the biggest culprits in spreading fake news among the population. The paper applies pre-processed and extensive data exploration are applied in this work.They have used a variety of ensemble approaches to training a variety of machine learning algorithms.determined automatic detection approaches based on deep learning, and machine learning techniques.Fake news applies sentiment analysis,the branch of information retrieval and information extraction.

A variety of approaches are implemented in this experiment.The KNN model was used in the beginning but it did not turn out to perform well for this study.Then the author proposed methodology collects fundamental properties from false news datasets,then categorizes them using an ensemble model that combines three main machine learning methods. Then they determined automatic detection approaches based on deep learning, and machine learning was researched to combat the rise and distribution of fake news.Spreading fake news for the sake of entertainment is a terrible act. This type of news should not be shared with the population because when it turns out to be inaccurate, they become disheartened and depressed.

One of the main reasons that fake news is spreading rapidly worldwide is that we rely heavily on the information we acquire from social media or any other news platform.This is the basic reason to implement this experiment to analyse the fake news in order to overcome this situation.Our social media is generating every kind of news; mostly, these are fake. Usually, we see clashing realities for a similar point and wonder whether both are valid. We set ourselves in a fix trying to figure out which source to put our confidence. As we have also discussed in the Discussion section, cleaning the dataset is very important.However, when the dataset has not been cleaned, words such as the, are, and appear the most often. These words on their own have no identity and are considered meaningless until they are used with the other terms.

2.20 Deep Learning for Fake News Detection in a Pairwise Textual Input Schema[20]

Among the social network users, a large volumes of online information is observed.This voluminous information can be easily exploited by malicious users and many entities.Their propaganda is to forge ,distribute and reproduce fake news.Fake news detection is the ability to define the truthfulness of information by analyzing its contents and find the forgeries it created.The news on Twitter is often been exploited by malicious users and various other entities,and is intentionally a forged information ,which is distributed either to deceive and make false information believable.

The present work investigates the modeling process that identifies real vs. fake tweets (text and headers) using the learning schema.Two experiments were conducted for detecting fake news by different pair wise inputs.For all the experiments, we used 10-fold cross-validation, which is effective for small data sets. Cross-validation is a technique to evaluate predictive models by partitioning the original sample into a training set to train the model, and a test set to evaluate it.The fake news detection task is viewed as a binary classification problem. We propose a pairwise ranking approach in detecting tweets with fake content. This uses maximum likelihood estimation, while the classification problem is modeled as a Bernoulli distribution.

In order to model the textual input, an embedding layer (automatically calculated) is used for the two different tweets per input pair.The input dimensions of the embedding layer are in agreement with the vocabulary of each input tweet text, taking into account the most frequent words.This presents the following novelty and contribution. While the problem of fake news detection has been tackled in the past in a number of ways, most reported approaches rely on a limited set of existing, widely accepted, and validated real/fake news data.The process of exploiting the provided fake tweets by Twitter itself, as well as the process of collecting and validating real tweet news pertaining to the particular event, are described in detail and generate a best practice setting for developing fake/real news data sets with significant derived findings.

References

[1] V. Gupta, R. S. Mathur, T. Bansal, and A. Goyal, "Fake News Detection using Machine Learning," *2022 Int. Conf. Mach. Learn. Big Data, Cloud Parallel Comput. COM-IT-CON 2022*, vol. 5, no. 5, pp. 84–89, 2022, doi: 10.1109/COM-IT-CON54601.2022.9850560.

[2] M. Looijenga, "The Detection of Fake Messages using Machine Learning," *Retrieved Sept.*, vol. 1, p. 2021, 2018, [Online]. Available: http://essay.utwente.nl/77385/

[3] N. Guimarães, Á. Figueira, and L. Torgo, "Can fake news detection models maintain the performance through time? A longitudinal evaluation of twitter publications," *Mathematics*, vol. 9, no. 22, 2021, doi: 10.3390/math9222988.

[4] T. Jiang, J. P. Li, A. U. Haq, A. Saboor, and A. Ali, "A Novel Stacking Approach for Accurate Detection of Fake News," *IEEE Access*, vol. 9, pp. 22626–22639, 2021, doi: 10.1109/ACCESS.2021.3056079.

[5] M. Al-Sarem *et al.*, "An improved multiple features and machine learning-based approach for detecting clickbait news on social networks," *Appl. Sci.*, vol. 11, no. 20, 2021, doi: 10.3390/app11209487.

[6] Y. Fang, J. Gao, C. Huang, H. Peng, and R. Wu, "Self Multi-Head Attention-based Convolutional Neural Networks for fake news detection," *PLoS One*, vol. 14, no. 9, pp. 1–13, 2019, doi: 10.1371/journal.pone.0222713.

[7] J. W. Lee and J. H. Kim, "Fake Sentence Detection Based on Transfer Learning: Applying to Korean COVID‐19 Fake News," *Appl. Sci.*, vol. 12, no. 13, 2022, doi: 10.3390/app12136402.

[8] A. Galli, E. Masciari, V. Moscato, and G. Sperlí, "A comprehensive Benchmark for fake news detection," *J. Intell. Inf. Syst.*, vol. 59, no. 1, pp. 237–261, 2022, doi: 10.1007/s10844-021-00646-9.

[9] S. Helmstetter and H. Paulheim, "Collecting a large scale dataset for classifying fake news tweets using weak supervision," *Futur. Internet*, vol. 13, no. 5, 2021, doi: 10.3390/fi13050114.

[10] L. Hu, S. Wei, Z. Zhao, and B. Wu, "Deep learning for fake news detection: A comprehensive survey," *AI Open*, 2022, doi: 10.1016/j.aiopen.2022.09.001.

[11] S. S. Nikam* and P. R. Dalvi, "Fake News Detection on Social Media using Machine Learning Techniques," *Int. J. Innov. Technol. Explor. Eng.*, vol. 9, no. 7, pp. 940–943, 2020, doi: 10.35940/ijitee.g5428.059720.

[12] S. Ahmed, K. Hinkelmann, and F. Corradini, "Combining machine learning with knowledge engineering to detect fake news in social networks - A survey," *CEUR Workshop Proc.*, vol. 2350, 2019.

[13] A. Shrestha, F. Spezzano, and A. Joy, "Detecting Fake News Spreaders in Social Networks via Linguistic and Personality Features Notebook for PAN at CLEF 2020," *CEUR Workshop Proc.*, vol. 2696, no. September, pp. 22–25, 2020.

[14] J. Y. Khan, M. T. I. Khondaker, S. Afroz, G. Uddin, and A. Iqbal, "A benchmark study of machine learning models for online fake news detection," *Mach. Learn. with Appl.*, vol. 4, p. 100032, 2021, doi: 10.1016/j.mlwa.2021.100032.

[15] M. S. Espinosa, R. Centeno, and Á. Rodrigo, "Analyzing User Profiles for Detection of Fake News Spreaders on Twitter," *Work. Notes CLEF 2020 - Conf. Labs Eval. Forum*, no. September, pp. 22–25, 2020.

[16] F. Monti, F. Frasca, D. Eynard, D. Mannion, and M. M. Bronstein, "Fake News Detection on Social Media using Geometric Deep Learning," pp. 1–15, 2019, [Online]. Available: http://arxiv.org/abs/1902.06673

[17] M. Hernandez-Mendoza, A. Aguilera, I. Dongo, J. Cornejo-Lupa, and Y. Cardinale, "Credibility Analysis on Twitter Considering Topic Detection," *Appl. Sci.*, vol. 12, no. 18, pp. 1–25, 2022, doi: 10.3390/app12189081.

[18] Y. Jang, C. H. Park, and Y. S. Seo, "Fake news analysis modeling using quote retweet," *Electron.*, vol. 8, no. 12, 2019, doi: 10.3390/electronics8121377.

[19] M. J. Awan *et al.*, "Fake news data exploration and analytics," *Electron.*, vol. 10, no. 19, 2021, doi: 10.3390/electronics10192326.

[20] D. Mouratidis, M. N. Nikiforos, and K. L. Kermanidis, "Deep learning for fake news detection in a pairwise textual input schema," *Computation*, vol. 9, no. 2, pp. 1–15, 2021, doi: 10.3390/computation9020020.