NAAN MUDHALVAN PROJECT(IBM)

IBM AI 101 ARTIFICIAL INTELLIGENCE-GROUP 1

DONE BY

GOKUL NATH S

(Email: 12345.gokulnath.s@gmail.com)

(NM ID: au110321106013)

ECE 3Rd Year
From the Department of

ELECTRONICS AND COMMUNICATION ENGINEERING

PROJECT:

TEAM-6 FAKE NEWS DETECTION USING NLP

Abstract:

Fake news is a real problem in today's world, and it has become more extensive and harder to identify. A major challenge in fake news detection is to detect it in the early phase. Another challenge in fake news detection is the unavailability or the shortage of labelled data for training the detection models. We propose a novel fake news detection framework that can address these challenges. Our proposed framework exploits the information from the news articles and the social contexts to detect fake news. The proposed model is based on a Transformer architecture, which has two parts: the encoder part to learn useful representations from the fake news data and the decoder part that predicts the future behaviour based on past observations. We also incorporate many features from the news content and social contexts into our model to help us classify the news better. In addition, we propose an effective labelling technique to address the label shortage problem. Experimental results on real-world data show that our model can detect fake news with higher accuracy within a few minutes after it propagates (early detection) than the baselines.

In today's digital era, the spread of information via social media and internet platforms has given people the power to access news from many different sources. The growth of fake news, meanwhile, is a drawback of this independence. Fake news is inaccurate information that has been purposefully spread to confuse the public and undermine confidence in reputable journalism. Maintaining an informed and united global community requires identifying and eliminating fake news. NLP, a subfield of artificial intelligence, gives computers the capacity to comprehend and interpret human language, making it a crucial tool for identifying deceptive information. This article examines how NLP can be used to identify fake news and gives examples of how it can be used to unearth misleading data.

Introduction:

Fake news detection is a subtask of text classification and is often defined as the task of classifying news as real or fake. The term ‘fake news’ refers to the false or misleading information that appears as real news. It aims to deceive or mislead people. Fake news comes in many forms, such as clickbait (misleading headlines), disinformation (with malicious intention to mislead the public), misinformation (false information regardless of the motive behind), hoax, parody, satire, rumour, deceptive news and other forms as discussed in the literature.

Fake news is not a new topic; however, it has become a hot topic since the 2016 US election. Traditionally, people get news from trusted sources, media outlets and editors, usually following a strict code of practice. In the late twentieth century, the internet has provided a new way to consume, publish and share information with little or no editorial standards. Lately, social media has become a significant source of news for many people. According to a report by Statistica, there are around 3.6 billion social media users (about half the population) in the world. There are obvious benefits of social media sites and networks in news dissemination, such as instantaneous access to information, free distribution, no time limit, and variety. However, these platforms are largely unregulated. Therefore, it is often difficult to tell whether some news is real or fake.

Recent studies show that the speed at which fake news travels is unprecedented, and the outcome is its wide-scale proliferation. A clear example of this is the spread of anti-vaccination misinformationFootnote2 and the rumour that incorrectly compared the number of registered voters in 2018 to the number of votes cast in US Elections 2020. The implications of such news are seen during the anti-vaccine movements that prevented the global fight against COVID-19 or in post-election unrest. Therefore, it is critically important to stop the spread of fake news at an early stage.

A significant research gap in the current state-of-the-art is that it focuses primarily on fake news detection rather than early fake news detection. The seminal works on early detection of fake news usually detect the fake news after at least 12 h of news propagation, which may be too late . An effective model should be able to detect fake news early, which is the motivation of this research.

Another issue that we want to highlight here is the scarcity of labelled fake news data (news labelled as real or fake) in real-world scenarios. Existing state-of-the-art works [4, 7, 8] generally use fully labelled data to classify fake news. However, the real-world data is likely to be largely unlabelled. Considering the practical constraints, such as unavailability of the domain experts for labelling, cost of manual labelling, and difficulty of choosing a proper label for each news item, we need to find an effective way to train a large-scale model. One alternative approach is to leverage noisy, limited, or imprecise sources to supervise labelling of large amounts of training data. The idea is that the training labels may be imprecise and partial but can be used to create a strong predictive model. This scheme of training labels is the weak supervision technique.

Usually, the fake news detection methods are trained on the current data (available during that time), which may not generalize to future events. Many of the labelled samples from the verified fake news get outdated soon with the newly developed events. For example, a model trained on fake news data before the COVID-19 may not classify fake news properly during COVID-19. The problem of dealing with a target concept (e.g. news as ‘real’ or ‘fake’) when the underlying relationship between the input data and target variable changes over time is called concept drift. In this paper, we investigate whether concept drift affects the performance of our detection model, and if so, how we can mitigate them.

This paper addresses the challenges mentioned above (early fake news detection and scarcity of labelled data) to identify fake news. We propose a novel framework based on a deep neural network architecture for fake news detection. The existing works, in this regard, rely on the content of news, social contexts, or both. We include a broader set of news-related features and social context features compared to the previous works. We try to detect fake news early (i.e. after a few minutes of news propagation). We address the label shortage problem that happens in real-world scenarios. Furthermore, our model can combat concept drift.

Inspired by the bidirectional and autoregressive Transformer (BART) model from Facebook that is successfully used in language modelling tasks, we propose to apply a deep bidirectional encoder and a left-to-right decoder under the hood of one unified model for the task of fake news detection. We choose to work with the BART model over the state-of-the-art BERT model , which has demonstrated its abilities in NLP (natural language processing) tasks (e.g. question answering and language inference), as well as the GPT-2 model , which has impressive autoregressive (time-series) properties. The main reason is that the BART model combines the unique features (bidirectional and autoregressive) of both text generation and temporal modelling, which we require to meet our goals.

Though we take inspiration from BART, our model is different from the original BART in the following aspects: in comparison with the original BART, which takes a single sentence/document as input, we incorporate a rich set of features (from news content and social contexts) into the encoder part; we use a decoder to get predictions not only from previous text sequences (in this case, news articles) as in the original BART but also from previous user behaviour (how users respond to those articles) sequences, and we detect fake news early by temporally modelling user behaviour; on top of the original BART model, we add a single linear layer to classify news as fake or real.

Problem Statement:

News consumption is a double-edged sword. On the one hand, its low cost, easy access, and rapid dissemination of information lead people to seek out and consume news. It enables the wide spread of “fake news”, i.e., low quality news with intentionally false information. The extensive spread of fake news has the potential for extremely negative impacts on individuals and society. Therefore, fake news detection has recently become an emerging research that is attracting tremendous attention. First, fake news is intentionally written to mislead readers to believe false information, which makes it difficult and nontrivial to detect based on news content. To develop a FAKE NEWS DETECTION system using natural language processing and its accuracy will be tested using machine learning algorithms. The algorithm must be able to detect fake news in a given scenario.

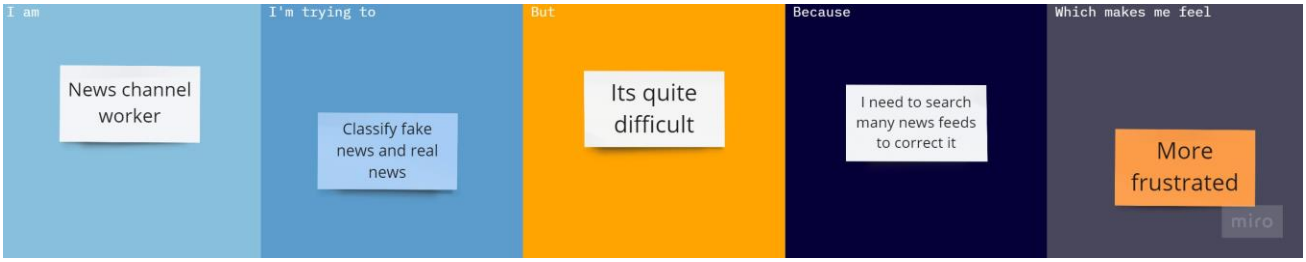
Some challenges in fake news detection include:

1. Detecting it in the early phase
2. The unavailability or shortage of labeled data for training the detection models
3. The lack of a massive dataset and a labeled benchmark dataset with ground-truth labels

Some approaches to fake news detection include:

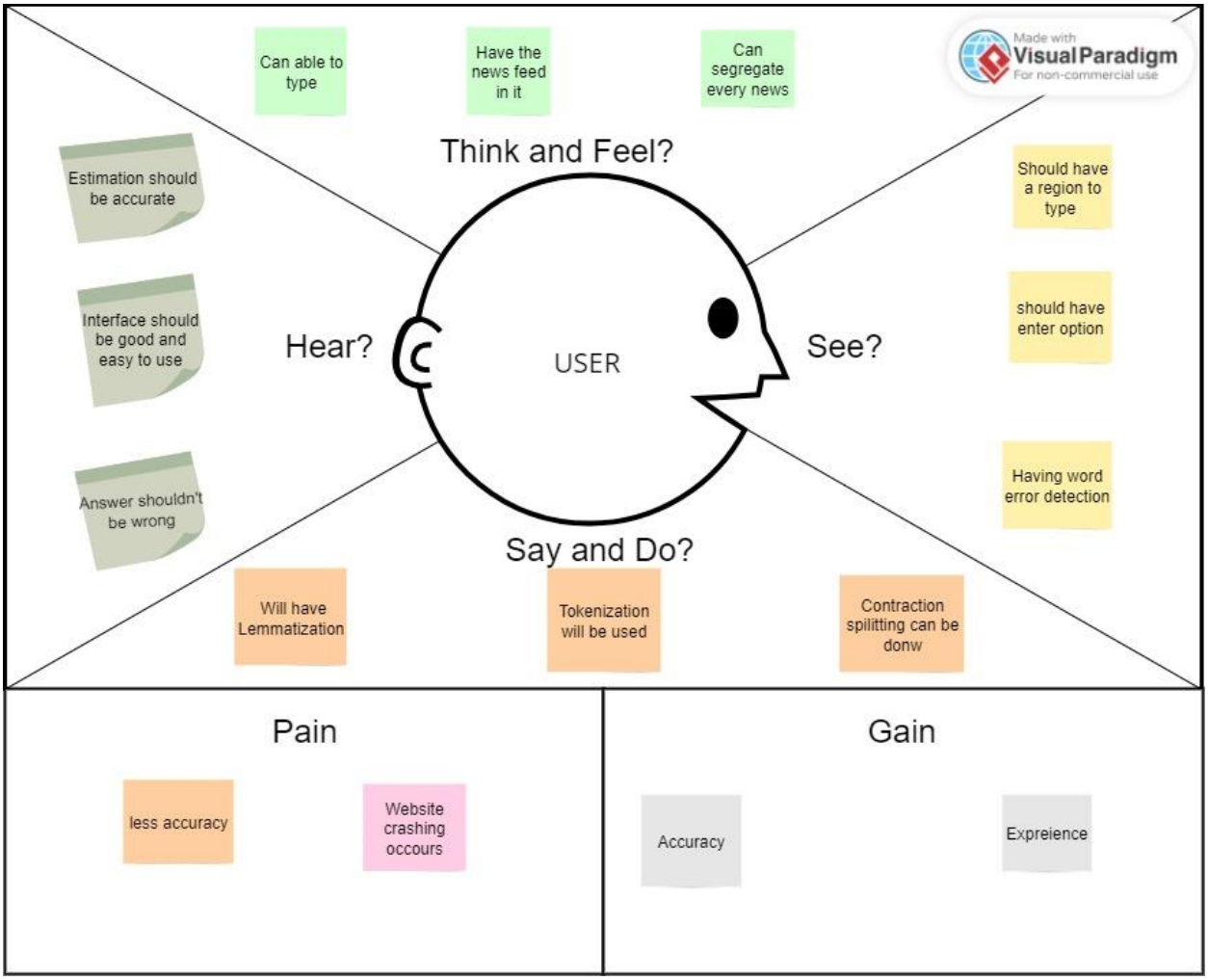
1. We are using NLP based algorithm to build a model to predict whether a given news article is real or fake based on its text
2. We will use Bag of Words to convert text to vector format
3. BY using TF-IDF to extract the feature
4. By using CNN for classification
5. Using classification algorithms such as logistic regression, extreme gradient boosting, decision tree, and random forest classifier

Customer Problem Statement



Nowadays, online fake news has become the main aspect of the growing interest in online media, social-networking sites, and online news portals (Bondielli and Marcelloni 2019). However, most people are generally incompetent for spending adequate time cross-checking the references and for ensuring the credibility of news (Zhou and Zafarani 2020; D'Ulizia 2021). Thus, more attention to fake news detection inspires the research community. In recent days, more research works regarding fake news detection have been implemented (Rama Krishna et al. 2021), though several studies only concentrated on news of specific categories like political or e-commerce reviews. Consequently, they have designed and developed certain features with some standard datasets with their topic of interest. These studies face poor performance in detecting news of another topic and also dataset bias (Beer and Matthee 2020). Therefore, it is necessary for studying whether these models are suitable for diverse classes of news propagated in social media through the evaluation of diverse datasets on different models and investigating their efficiency or performances (Ahmad et al. 2020). On the other hand, conventional studies on fake news detection techniques are focused on either a limited number of models or a particular category of the dataset (Dabbous et al. 2020a). Thus, there is a need of reviewing a fake news detection model.

Empathize & Discover:



This is empathize map and discovery which is been implemented by our team and these are some of the ideologies that we want to implement in it.