CHAPTER 2

## 2.1 Introduction

In recent times, the surge of misinformation and fake news has emerged as a significant issue, especially regarding social media platforms. The literature review acts as an in-depth examination of the current knowledge and studies concerning fake news identification and profiling systems. This chapter explores the importance of the issue, existing solutions, possible alternatives, and the background of the suggested "Fake News Detection and Profiling System" project.

## 2.2 Type of Fake News

Although mainstream media often talks about fake news and rumors, researchers have investigated numerous other aspects of online misinformation (Bondielli A, 2019). This encompasses the analysis of clickbait, social spam, and fake reviews. The academic literature has presented various categorizations for fake news and rumors, typically depending on the source and the type of information used in the analysis.

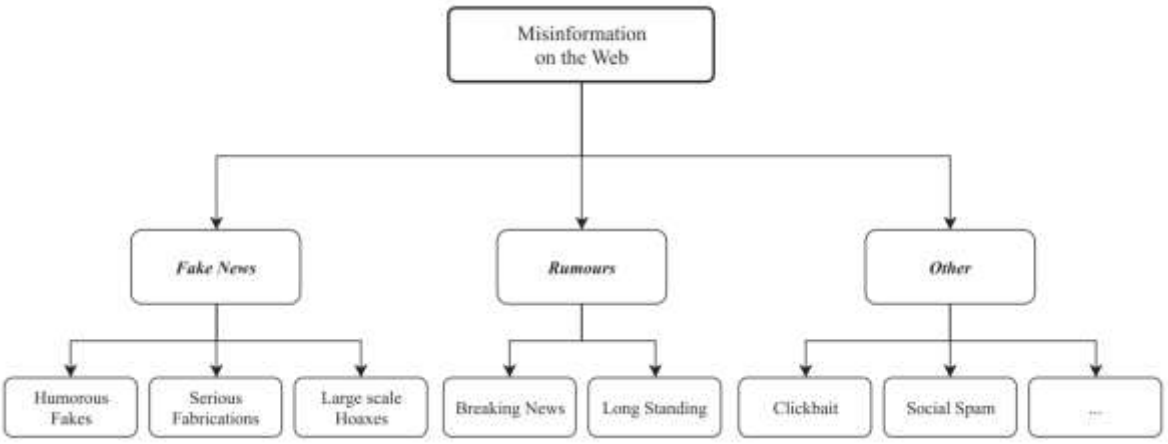



Figure 2.1 Type of Fake news (Bondielli A, 2019)

### 2.2.1 Humorous Fakes

Humorous Fakes are created to amuse audiences who are expected to detect the author's intended comedic tone (Bondielli A, 2019). Examples include satirical pieces artfully masquerading as real news, as found in outlets like The Onion and lercio, aiming to provoke laughter while ensuring readers are conscious of the content's fictional aspect.

### 2.2.2 Serious Fabrications

Serious fabrication entails the deliberate creation and dissemination of false or misleading information to deceive or sway others. This term refers to situations where the fabrication might have important effects, influencing how individuals view matters, make choices, or participate in conversations. It encompasses the fabrication of false narratives, deceptive material, or altered depictions, all motivated by a deliberate aim to influence views, convictions, or behaviors. The severity of serious fabrication stems from the possible damage it can inflict on individuals, communities, or society overall because of the spread of misleading information. For example, political or war propaganda represents a form of significant distortion that intentionally spreads information, frequently with a misleading or biased quality, intended to sway public opinion or advance a specific agenda.

### 2.2.3 Large Scale Hoax

Widespread hoaxes entail the distribution of misleading information, portrayed as credible news, frequently exceeding the limits of a conventional news article. These coordinated deceptions generally function on a broader level, frequently aiming to focus on notable public individuals or powerful concepts. The complex orchestration and implementation of these deceptions set them apart from mere misinformation, highlighting a purposeful attempt to sway public opinion or alter narratives on a larger scale.

## 2.3 Fake News Detection

As technology progresses, particularly in social media and other platforms for sharing ideas, anyone can freely express themselves and write about news. The simpler the access to news and the more straightforward the process of writing it, the more the digital world becomes inundated with information. With the overflow of information arises a danger known as fake news. False news consists of low-quality information containing intentionally misleading data, spread by individuals or bots that deliberately alter messages for gossip or political motives (Kumar P, et al. 2021). The phrase "fake news" has its origins in history, tracing back to the rise of mass media (Schudson, 2017). Nonetheless, it attracted considerable worldwide focus following the 2016 U.S. presidential election, during which the extensive spread of misinformation on social media drew the attention of a wider audience than conventional news readers, as pointed out by Kumar Amorin. False information is generated for different purposes, intending to sway public views for monetary or political advantage, or occasionally simply for entertainment. The rise of false information on social media adversely affects people by prompting them to justify detrimental actions (as pointed out by

Sadio, et al. 2021), undermining their basic rationales, and reducing faith in science, casting doubt on its capacity to provide trustworthy solutions (as observed in Eichengreen, et al. 2021). The previously mentioned factors highlight the need for developing fake news detectors.

## 2.4 Techniques to Detect Fake News

Multiple techniques have been employed to create fake news detection systems. A study conducted by Moorpani (2022) effectively utilized Naïve Bayes to develop a model capable of differentiating between genuine news and false news. While Patel was implementing a Super Vector Machine to detect fake news.

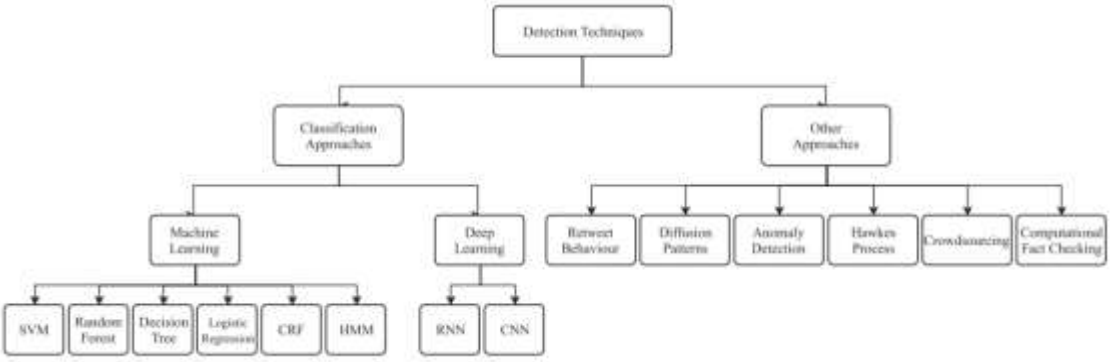



Figure 2.2 Fake news detection techniques (Bondielli A, et al. 2019)

### 2.4.1 Naïve Bayes

Naïve Bayes can be applied to identify fake news, as the algorithm is commonly utilized in machine learning tasks. Naïve Bayes is a probabilistic algorithm grounded in the concept of Bayesian probability. Naïve Bayes operates under the assumption that the characteristics used to represent an instance are conditionally independent when the class label is known.

$$P(h|d) = \frac{P(d|h) * P(h)}{P(d)} \quad (2.1)$$

The method for identifying false news involves extracting text from the sentence, which is then pre-processed and vectorized to enable data input into a machine learning algorithm. The algorithm demonstrates a 95% accuracy rate, making it effective for detecting fake news. The presumption of independent features, which often does not apply in various real-world situations, posed a constraint

for this algorithm. Nonetheless, the simplicity, efficiency, and effectiveness in specific NLP tasks contributed to Naïve Bayes' popularity.

### 2.4.2 Super Vector Machine

Super Vector Machine, or SVM, is a supervised model designed to identify a hyperplane that optimally divides data points from different classes. The main concept of SVM is to increase the margin between the classes, with the margin defined as the distance from the hyperplane to the closest data points of each class. The work conducted by Sudhakar (2022) on the use of SVM for detecting fake news achieved an accuracy rate of 94.93%. The Super Vector Machine has demonstrated itself to be among the most precise text classification algorithms. Nonetheless, misclassification may take place during the process. This is the reason performance measurement holds significance.

## 2.5 Word Embedding and Fake News

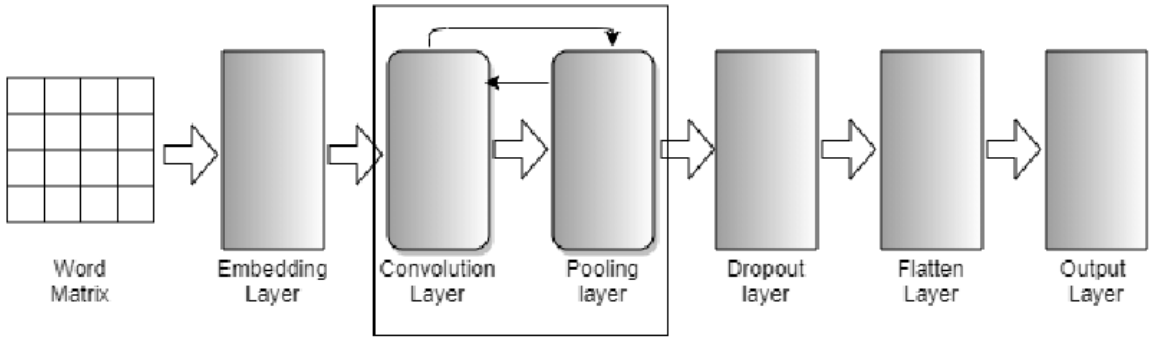
Numerous machine learning algorithms and nearly all deep learning frameworks are unable to handle string or text data in its raw form (NSS, 2023). Therefore, the plain text must be transformed into a numerical format. The conventional approaches such as word embedding and TF-IDF consider a word as distinct entities. More advanced word embeddings, such as Word2Vec, utilize Continuous Bag of Words and skip-gram techniques to establish connections between words. These methods depended on forecasts by allocating probabilities to words, showcasing their efficacy as innovative strategies in activities like word analogies and assessing word similarities. Word2Vec is not just one algorithm; it is a combination of CBOW and skip-gram methods derived from a shallow neural network.

On the other hand, GloVe, which stands for Global Vectors for Word Representation, is a model crafted to obtain vector space representations of words, with the goal of capturing subtle semantic and syntactic patterns using vector arithmetic. This model effectively leverages statistical data through training on worldwide word-word co-occurrence frequencies, demonstrating enhanced performance in tasks like word analogy, word similarity, and named entity recognition. In comparison to other models, GloVe excels in speed and efficiency, consistently outperforming word2vec across different metrics such as corpus, vocabulary, window size, and training duration. Its effectiveness is ascribed to its skill in identifying significant linear substructures, in line with recent improvements in log-bilinear prediction techniques such as word2vec. Showcasing its advanced features, GloVe achieves a remarkable 75% accuracy on the word analogy dataset (Pennington J, et al. 2014).

## 2.6 LSTM

Long Short-Term Memory (LSTM) is an architecture of Recurrent Neural Networks (RNN) alteration. RNNs are limited to remembering only recent data, while LSTMs are capable of managing data over an extended period. Additionally, an RNN model encounters a problem known as the disappearing gradient issue when dealing with lengthy sequence data; however, LSTM could circumvent this issue during training. This model is capable of remembering previous long-term sequential data and enables automatic regulation in the cell state for preservation helpful and eliminating unnecessary attributes. The three entrances that control characteristics the input gate, The forget gate, and the output gate are components of an LSTM model.

The input gate serves to enable new information to enter the cell state. The forget gate eliminates unnecessary data from the cell state. The output control, that governs the data extracted from the cell state, subsequently decides what will serve as the subsequent concealed state. An LSTM model might automatically store or erase. documented memory utilizing these gates. The physical structure of the suggested The CNN models utilized in this research are observed from




## 2.7 Discussion

A discussion section covers the implications, challenges and you get a sense of the results in this chapter. There is an immense variety and continual evolution of the techniques of fake news detection. High accuracy rates have also been reported for certain algorithms, such as Naïve Bayes and Support Vector Machines (SVMs), indicating their ability to accurately identify and classify fake news. But each approach has inherent limitations. Naïve Bayes, for example, presumes independence between features given the class (conditional independence), which may not be the case in reality, while SVM's efficacy is sensitive to how a suitable kernel and hyperparameters are chosen.

All this marks a significant advancement towards understanding linguistic patterns and semantics comparing to traditional approaches, such as bag-of-words and term frequency-inverse document frequency (TF-IDF), famous techniques for the text represented model like word embeddings

(Word2vec and GloVe), long short term memory (LSTM) networks. These approaches enhance the efficacy of fake news detection systems, especially when it comes to dealing with complex language constructs. Still, there are many aspects that need to be explored more with regards to data biases, setup of the tasks, the need for scaling up, handling multilinguality, the presence of rich context, etc.

Additionally, the chapter notes the wide reaching implications of fake news, which speaks to the need for a more comprehensive approach. However, it is clear that technical solutions need to work in conjunction with educational initiatives and policy frameworks to mitigate the effects of misinformation.

## 2.8 Summary

In this chapter, we presented a detailed overview of existing literature on fake news detection and profiling systems. It looked at different varieties of fakes, from jokey ones to far more serious fakes, and discussed their societal implications. Analyses were also conducted regarding techniques for detecting fake news, focusing on methods of machine learning and natural language processing, including Naïve Bayes, SVM, word embeddings, and long short-term memory networks.

Despite promising results, challenges were identified around model limitations, data diversity, and contextual understanding for the new technologies. These insights from this review will helpe facilitate an effective fake news detection framework, focusing on the incorporation of advanced algorithms with moral aspects and user-centric designs.