

DETECTING FAKE NEWS WITH LSTM: A DEEP LEARNING APPROACH

Tarun Kumar

ABSTRACT

Fake news has become a significant concern in today's digital age, as it can easily mislead and deceive people. One of the fundamental steps to combat fake news is through fake news detection. In recent years, research has focused on developing effective techniques for detecting fake news using Natural Language Processing. It further affects their emotions and cultural values, creating a tense among people, and instigating anger against fellow people or the government. This further changes the order in society causing hate speeches, riots, strikes, or even fatal accidents that affect innocent lives.

When situations like these affect society, it's wise that we utilize the necessary technologies in aid to eliminate this kind of issue. One such tech that's disrupting human language processing is Natural Language Processing (NLP) which processes information like text, audio, etc. between languages and text, and audio formats. Leveraging this technology to process the text and analyze the motive and sentiment could shed light on the reliability of the trueness of the text. As a result of this, we could develop a fake news detection system using NLP.

To develop such a system, we have to train the machine learning model with an existing dataset. Based on the knowledge gained from this training, the model would be able to identify a pattern in the test data we give. Here, training the model plays an important role as it affects the accuracy of the model in giving the final output. This model can be evaluated by testing it with real and fake news articles. We can develop a better understanding of how to detect fake news from the result generated by the model and we can fine-tune it in order to attain a higher accuracy level. These developed accuracy models can be used by others for various purposes in NLP to build better adapting models that make the life of humans easier.

Keywords: Fake News, Natural Language Processing, Machine Learning, detection.

II. LITERATURE REVIEW

[1]. The authors dive deeper into the increased traffic in social media leading to spread of misinformation which becomes popular and trending. This further raises a doubt against any other statements made even if it is true. Since online content is diverse, monitoring it is almost impossible which is the main reason for the spread of fake news. They further explain how they use state of the art detection methods to identify features of the users, content, and context that indicate misinformation. This is the basic structure of instructions for carrying out research on building a Fake news detection system. It would have been helpful if the authors had mention about the approaches they used to build this system.

[2]. The authors of this paper dive into Multi Domain Fake News Detection(MFND) and its variation from Single Domain Fake News Detection (SFND) which provides a much more compromised performance compared to MFND. Considering variations of data distributions in different domains due to domain shift, they propose a specialized design for the MFND model by using domain gate to put together multiple representations extracted from various experts. The three fold approach of this project involves the following steps: 1) creating an exclusive dataset - Weibo21, with news from 9 different domains, 2) proposal of a system that uses domain gate to include multiple features extracted from a mixture of multiple domains and 3) systematically demonstrating the effectiveness of the system by performance metrics in terms of multiple methods implemented using the Weibo21 dataset. The MFND system corroborates the use of Multi Domain Fake News Detection for a much better precision and an improved accuracy rate.

[3]. The authors are inspired by the research work by Kaliyar et al., (2020) and take the approach of deep neural networks in this research work. They explore a variety of neural networks and their functional and architectural structure and decide on which is best for the system of fake news detection. They finalize the leverage of Mc-DNN elaborated as multi-channel deep neural networks for processing news headlines and articles from different sources for fake or true news classification. Using this they achieve 99.23% accuracy on ISOT fake news dataset all thanks to

the approach of integrating several models in convolutional neural networks and deep neural networks by utilizing the maximum potential of deep learning methodologies. This is a deep analysis and a proper guide for building a DNN fake news detection system.

[4]. The authors use a deep convolutional neural network that learns to discriminate between fake and true news rather than relying on the existing hand-crafted models. Their FNDNet uses deep CNN to detect fake news features using multiple hidden layers in the deep neural network to extract several features in every layer. They also compare their model with several baseline models to improve the performance metrics of the model and have achieved an enormous 98.6% accuracy using benchmark datasets. These results are validated using various evaluation parameters like false positive, true negative, precision, Wilcoxon, recall, F1, and accuracy. These results exhibit the variation in performance of deep neural network models by a huge margin. This too is a great contribution towards the utilization of neural networks in fake news detection.

[5]. The authors indicate how the new methods that analyze the sequence of text in a unidirectional approach to detect fake news from different sources such as news headlines, articles and contextual information from social media. And contrast out their approach of using a bidirectional approach using the BERT (Bidirectional Encoder Representations from Transformers) combined with different parallel blocks of single layer deep Convolutional neural networks(CNN) to handle ambiguity that other models encounter. This is done using different kernel sizes and filters with BERT to form their own system - FakeBERT with a whooping 98.9% accuracy rate.

[6]. The authors explore the potential effects of fake news spread and the problems that might arise afterwards. Using 2D TF-IDF features instead of existing pre-trained models. They have also built a de facto dataset with labels for ‘aggressivity’ to detect and classify twitter comments as verbal aggression or not. In terms of algorithm, they used CNN with 128 filters on the convolution layer. This also involves usage of mean pooling instead of max pooling and two layered multilayer perceptron to classify tweets. This could be fine tuned for a fake news dataset in order to use it for fake news detection for its architecture.

[7]. The authors explore multiple methods that analyze both text and image data and their lack of comprehensive representation of relevant features. They further suggest the use of a multimodal feature vector with a higher information density called EFND system which combines contextual, social context and visual data from news articles and social media. They further use a multimodal factorized bilinear pooling method and a multilayer perceptron over a shared representation for fake news classification. They evaluate EFND using a standard dataset called FakeNewsNet. With EFND, they outperform industry standard machine learning and deep learning models to achieve an accuracy of 98.8% with PolitiFact dataset and a 99% accuracy with GossipCop dataset.

[8]. In this research paper, the authors explore how low cost, easy access to the internet and social media has led to an increased spread of low-quality fake news. Furthermore, they consider this a threat to society which results in decrement in the trust quotient for media and the bewildering condition of the audience. To tackle this, they suggest the use of their novel framework called Message Credibility (MCred) which uses the benefits of local and global text semantics to detect fake news. They have combined the potential of Bidirectional Encoder Representations from Transformers (BERT) for global text semantics by using the relationship between words in sentences and Convolutional Neural Networks (CNN) for local text semantics by using N-gram features to demonstrate an experimental result which is 1.10% more accurate than existing models on the popular Kaggle dataset.

[9]. In this paper, the authors suggest a combination of two techniques namely Chi-square and principal component analysis (PCA) to reduce the number of feature dimensions. They employ it with a hybrid neural network architecture of Convolutional neural networks(CNN) and long short-term memory(LSTM) on a multi-class dataset from Fake News Challenges(FNC) website. With this, they are able to gain an accuracy edge of 0.04 and a 0.20 edge in F1 score and with PCA an exorbitant 0.978 accuracy than Chi-square and other existing methods.

[10]. The authors explore a new hybrid model with a Weighted fine tuned BERT extraction model which extracts text using BERT and weighted attention to focus on specific parts of a sentence or document. Here, different weights are assigned to different parts of input based on

their importance to the desired task. After extracting the text features from BERT, words are embedded with weights. The features including weight values are represented as embedded vectors. Following this, it is passed through a BiLSTM (Bidirectional LSTM) which follows sequence processing where two inputs are taken opposite directions to each other so as to increase the amount of information available to each network. These methods are unique and could be an asset in building a fake news detection system when implemented with a precise fine tuning.

III. INTRODUCTION

3.1. FAKE NEWS

Any misinformation spreads as news for a variety of reasons including making money, influencing people's opinion, sowing discord, or harming a person or a group of people. Though it is difficult to identify fake news, there are several things that can expose the originality of a news article such as spelling, grammatical errors, lack of trusted sources and the extent of unbelievability. A study conducted by Oxford University in 2018 showed that fake news is more prominent than real news in social media. It was found in a 2020 study by the Pew Research Center that about 64% of Americans believed that the spread of fake news was a major problem in the US. In 2021, the Center for Media and Public Affairs found that most of the news on facebook was fake. A 2022 study by the University of California in Berkeley showed how fake news could affect the beliefs and behavior of people.

3.2. THREATS OF FAKE NEWS

Reliability is a tough job when it comes to the news we hear. It's become a lot more difficult to even find the mere difference between actual news and fake news. It's a problem in today's society when we have got a lot of sources and cannot even ensure their reliability. The effect of fake news significantly affects the lives of people as it misleads them about important events and issues. It further affects their emotions and cultural values, creating a tense among people, and instigating anger against fellow people or the government. This further changes the order in society causing hate speeches, riots, strikes, or even fatal accidents that affect innocent lives. Fake news refers to deliberately false or misleading information that is crafted to deceive people and manipulate public opinion. Its impact on society is substantial, as it spreads

misinformation and can fuel animosity towards individuals or specific groups, affecting their values, emotions, and cultural beliefs. They cause both psychological, mental or even physical harm to people who the news is targeted against. This causes the people to get exploited in many ways which leads to loose hope in life, people and government.

3.3. IMPACTS OF FAKE NEWS

3.3.1. Priyanka Reddy's Suicide In 2018

In December 2018, Priyanka Reddy, a veterinarian, was gang-raped and killed in Hyderabad, India. It created a huge spread of violence, with many people sharing misleading information, false rumors, and videos on social media. These rumors falsely claimed that Reddy was a drug addict and that she had been killed by a Muslim man. This led to an increase in communal violence, with many people attacking Muslims.

3.3.2. The 2019 Lynching Of Tabrez Ansari

In June 2019, a mob in Jharkhand, India killed Tabrez Ansari, a Muslim man because he was falsely accused of theft and of carrying beef, which is considered sacred by Hindus. The main cause of this incident was because People shared many false rumors and videos on social media falsely claiming that Ansari was a Muslim terrorist and that he had been caught carrying beef. This led to the mob forming and attacking Ansari, who was later found dead.

3.3.3. The Delhi Riots Of 2020

In February 2020, the death of a Hindu man in a mistaken identity case caused widespread riots in Delhi, India, following false news which played a significant role in the spread of violence, with many people sharing false rumors and misinformation on social media. These rumors falsely claimed that Muslims were attacking Hindus and that the government was not doing anything to protect them. This led to an increase in communal violence, with many people being injured or killed.

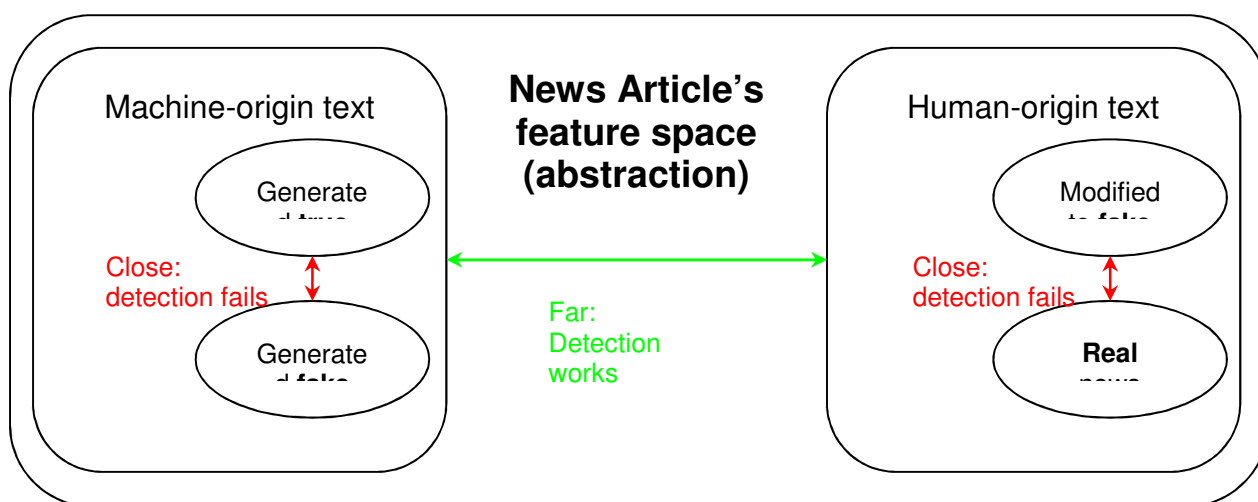
These are just representative samples from recent times but the effects of fake news on society goes to even older than the 5th century AD, the Tamil literature Silappathikaram by Ilanko Adikal depicts the wrath of a woman whose husband was executed based on a false

accusation that he stole the royal jewelry. So this has been a serious issue throughout history and we need a solution to this issue.

3.4. SEEKING SOLUTION WITH TECHNOLOGY

When situations like these affect society, it's wise that we utilize the necessary technologies in aid to eliminate this kind of issue. Detecting fake news can be approached in various ways. One method involves employing human fact-checkers, but this approach could be more scalable due to the sheer volume of online news articles. Another viable approach is utilizing a human-invented technology. One such tech that's disrupting human language processing is Natural Language Processing (NLP) which processes information like text, audio, etc. between languages and text, and audio formats. Leveraging this technology to process the text and analyze the motive and sentiment could shed light on the reliability of the trueness of the text. As a result of this, we could develop a fake news detection system using NLP.

To develop such a system, we have to train the machine learning model with an existing dataset. Based on the knowledge gained from this training, the model would be able to identify a pattern in the test data we give. Once trained, these algorithms can classify new articles as either authentic or fabricated. Additionally, an emerging approach with potential lies in applying natural language processing (NLP). NLP encompasses the processing of human languages and methods of communication between humans and computers. It can extract textual features that aid in identifying fake news. For instance, spotting certain keywords or phrases commonly linked to false information. It can also be applied to assess the emotional tone of a written piece,



aiding in the identification of potential deception within the text.

Fig 3.1 Fake news detection using NLP

Here, training the model plays an important role as it affects the accuracy of the model in giving the final output. This model can be evaluated by testing it with real and fake news articles. We can develop a better understanding of how to detect fake news from the result generated by the model and we can fine-tune it in order to attain a higher accuracy level. These developed accuracy models can be used by others for various purposes in NLP to build better adapting models that make the life of humans easier. These features make NLP a prominent tech in handling text and linguistic data both in small and large scale. This is the reason why I have employed NLP in my project that uses the following algorithm:

3.5. LONG SHORT-TERM MEMORY (LSTM)

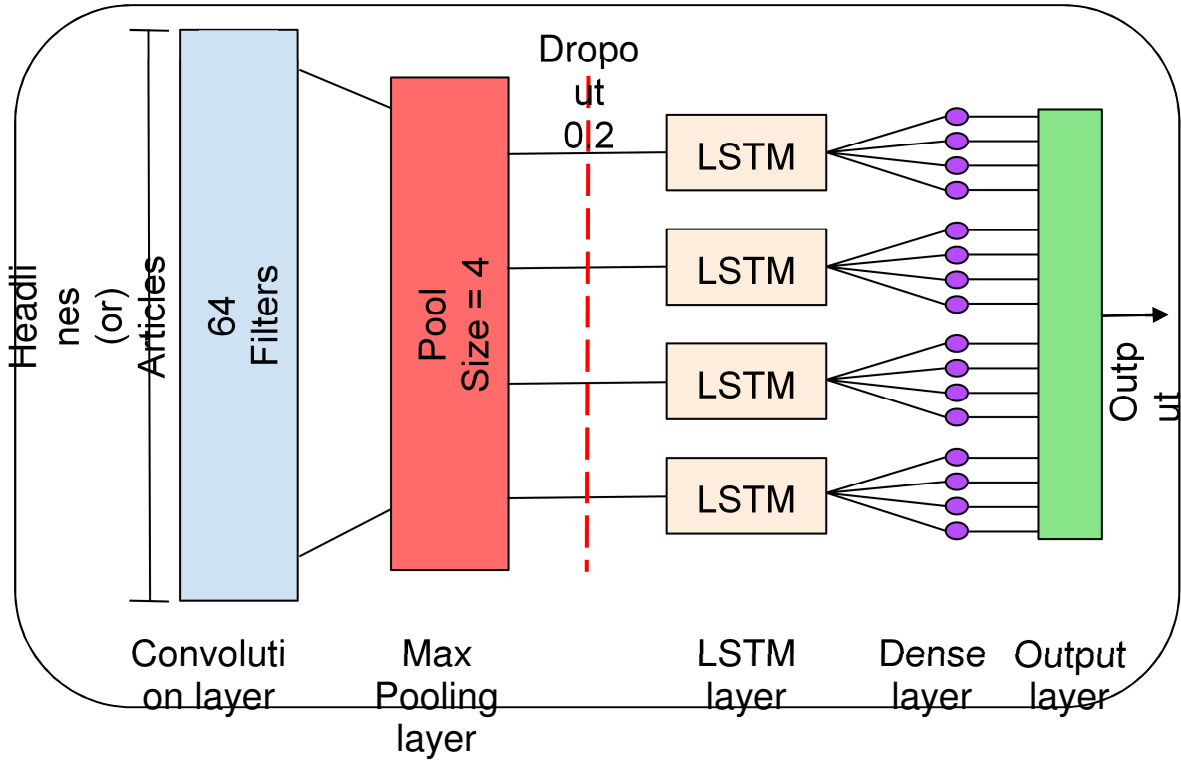


Fig 3.2 LSTM layers for Fake News Detection

I propose a fake news detection system that uses NLP and Long short-term memory (LSTM). LSTM is a type of artificial recurrent neural network (RNN) that can be used to classify text into different categories. I use LSTM to learn a model that can distinguish between fake and real news articles. Our system works by first extracting features from the text of a news article. These features can include the presence of certain keywords or phrases, the sentiment of the text, and the structure of the text. The features are then used to train a LSTM model. I evaluate our system on a dataset of real and fake news articles. In the future, I plan to improve our system by using more sophisticated NLP techniques. I also plan to evaluate our system on a larger dataset of news articles.

LSTM is capable of performing tasks that require sequential processing of data which is also used for performing various tasks such as machine translation, speech recognition and text generation. It has an advantage in fake news detection that it can learn long-term dependencies in text. This feature is used to identify the telltale patterns of deception like inconsistency in the context and the usage of sensational language which is generally found in fake news. Apart from this, LSTMs can be trained on large amounts of data which would come in handy when the fake news is constantly evolving and can be used to adapt to the changing trends by learning new patterns to stay ahead of the curve.

LSTM is capable of processing an entire data sequence apart from single units like images. This unique feature proves to be helpful while applying it in speech recognition, machine translation and various other NLP based applications. The major difference between LSTM and RNN is that when a task is given to modify a part of a data, RNN applies a function to completely change the data. Whereas, LSTM makes only small changes to the data by simple addition or multiplication operations that flow through the states of the cell. This helps LSTM to selectively forget and remember things which is an edge compared to RNN.

IV. FAKE NEWS DETECTION - PROCESS FLOW

The proposed solution of Fake news detection using NLP uses the following processes in order to classify the news as fake or not:

- Data Gathering

- Data Preprocessing
- Splitting Data
- Feature Selection
- Model Training
- Model Evaluation

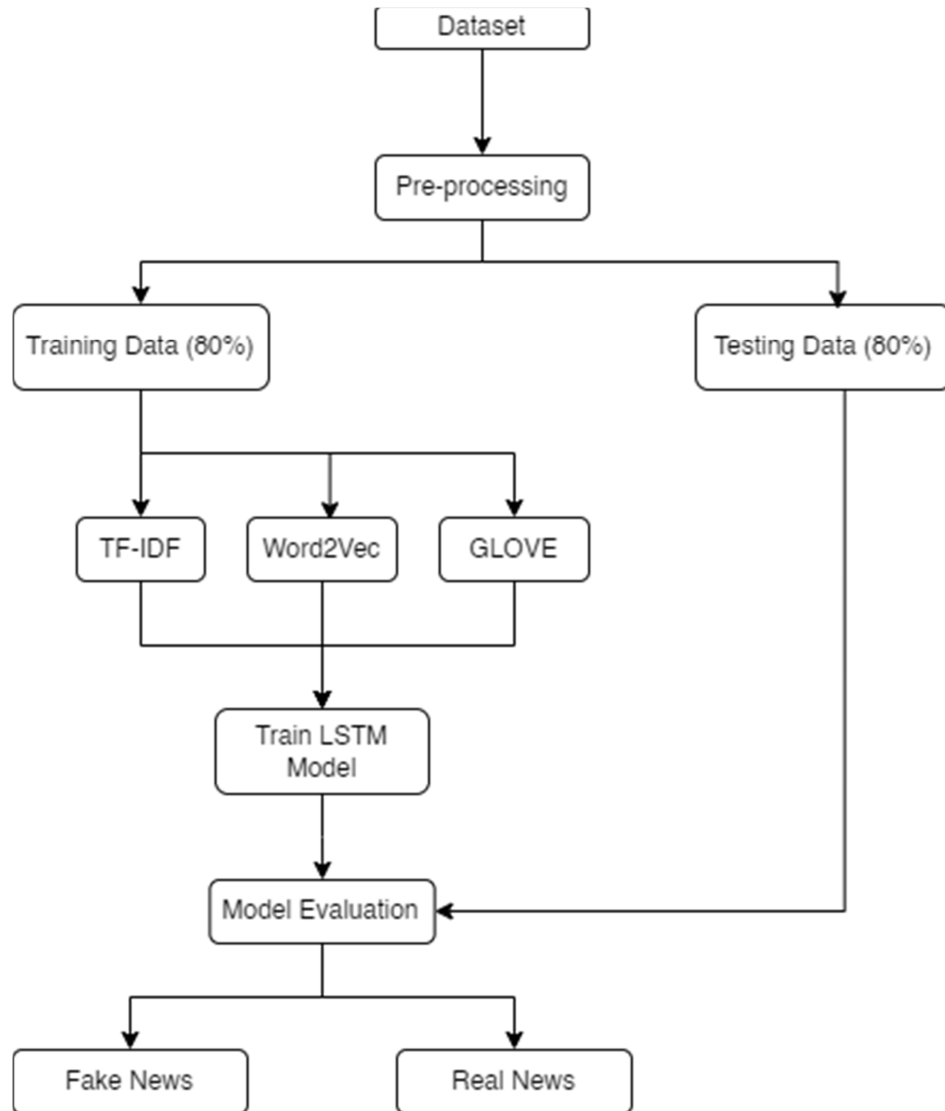


Fig 4.1 - Fake News Detection - Process flow

4.1 DATASET

Dataset is a combination of real and fake news with the following attributes: id, title, author and text taken from kaggle. It contains 18285 articles for training and testing models.

4.2 IMPLEMENTATION AND PREPROCESSING

To put the data into usable form, we need to remove null values, and decide the vocabulary size of 5000 words and apply Natural Language Toolkit (NLTK) to remove the stop words and punctuations from the data. Every word is verified for stopword, if it is stop word, it is removed and if not, a word tag is given from postage and this is added to the document.

4.3 TOKENIZING AND EMBEDDING

To tokenize the words, text is appended to a list and the list is named as documents and all the words in this list are given as output.

For embedding, the input cannot be given in text format so in order to convert it to a numeric form, we use Onehot representation. Onehot representation replaces each word in the dataset to the sentence according to the vocabulary size and index. Furthermore, feature extraction is applied to the dataset.

4.4 MODEL

The Output data from word embedding is fed into the model which is a sequential model with features like vocabulary size, number of features and length of sentence. Following this, LSTM is used with 100 neurons for each layer, followed by a dense layer with sigmoid activation function. To obtain one final output, binary cross entropy is used to calculate the loss, adam optimizer is used for adaptive estimation and a dropout layer is used as intermediary in order to prevent overfitting. After all these are finely tuned, the training and testing process starts.

4.5 LONG SHORT TERM MEMORY (LSTM)

LSTM units form the building block of layers of recurrent neural networks (RNN). The cell of an LSTM unit has an input gate, an output and a forget gate[11]. The cell remembers values over a large time interval to facilitate the relation function between the word at the start of the document and the output of word at the end of document. This cannot be done using the traditional neural networks which is a major shortcoming.

4.6 CLASSIFICATION

The preprocessed testing data is predicted and if the prediction value is > 0.5, it is classified as 1 which is real news and if it is 0, then the news is fake. The calculation for accuracy is derived using the formula;

$$\text{Accuracy} = (\text{TP} + \text{TN}) / \text{Total}$$

Where, TP - True Positive and TN - True Negative.

If it is TN, then prediction is negative and test case is also negative

If it is TP, then prediction is positive and test case is also positive

If it is FN, then prediction is negative but test case is positive

If it is FP, then prediction is positive but test case is negative

4.7 RESULT

The classification accuracy for real news and false news has only a marginal difference but the classification accuracy of fake news is slightly deviated from the plot. The usage of confusion matrix and classification report is to visualize these details and measure the accuracy of each model. The more the data the more improved accuracy could be attained. Here, the given dataset comprises 18285 news articles of which 80% i.e. 14628 is used for training and 20% i.e. 3657 is used for testing.

N = 3657	Predicted:NO	Predicted:YES
Actual:NO	TN = 1900	FP = 182
Actual:YES	FN = 145	TP = 1430

Also, using a dataset with greater length of news articles would improve the performance of the model to a notable extent. Removing stop words and applying stemming will reduce the word to its root form for a better and efficient processing. LSTM usage in fake news detection proves to be an efficient way with a 99% accuracy. But it is also wise to compare the performance of other models that are leading in the industry like, Naive Bayes algorithm, Decision tree algorithm, BERT classification, RNN, etc.

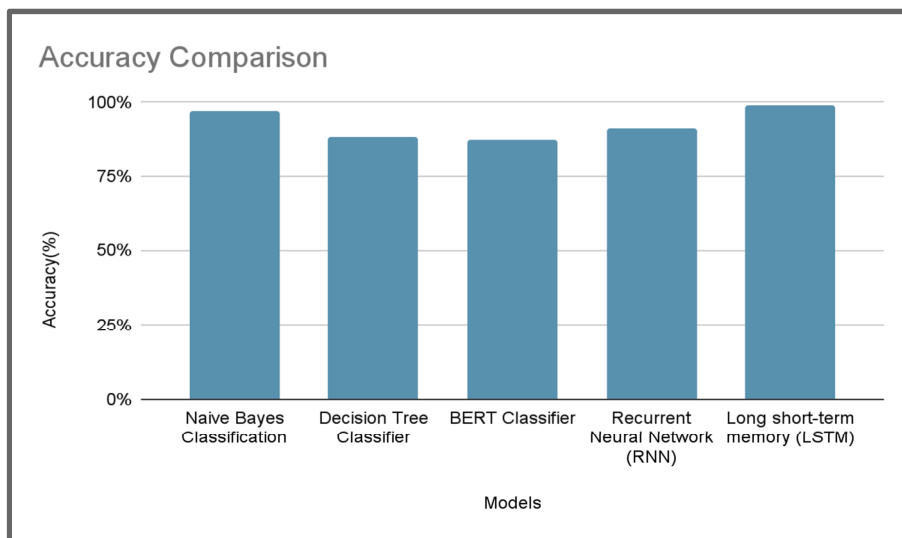


Fig 4.2 - Accuracy comparison of various models

So, from the above approaches and outputs we can see that LSTM outperformed all the other models and proved to be the best model in terms of accuracy. The ability to remember the long-distance dependencies for a long term is the main feature that provides this edge to LSTM.

V. CONCLUSION & SUGGESTION FOR FUTURE WORK

To conclude, fake news is an issue that is becoming relevant and widespread at an alarming rate. After searching for a viable solution for this issue, I found that the LSTM is an efficient way to classify news as reliable or not even when the resources are limited. A solution that is able to function even with less computational power is not only simple but also provides a better interpretation of the solution that comes in hand while explaining the solution. Also, with its tendency to remember the long-distance dependencies in dataset is a major positive as it has specialized memory cells for this. Furthermore, using NLP techniques to the solution makes it faster and easier to interpret textual information into categorical and numerical data to process it, fit it into the model, and arrive at a conclusion about whether the news is reliable or not.

Suggestions for future research:

Using pre-trained word embeddings, Incorporating attention mechanisms, Exploring other deep learning architectures, Incorporating additional features and Developing explainable and interpretable models are some possible fields of interest for future research.

VI. REFERENCES

1. Zhang, Xichen & Ghorbani, Ali. (2019). An overview of online fake news: Characterization, detection, and discussion. *Information Processing & Management*. 57. 10.1016/j.ipm.2019.03.004.
2. Nan, Qiong & Cao, Juan & Zhu, Yongchun & Wang, Yanyan & Li, Jintao. (2021). MDFEND: Multi-domain Fake News Detection. 3343-3347. 10.1145/3459637.3482139.
3. Tembhurne, J. V., Almin, M. M., & Diwan, T. (2022). Mc-DNN: Fake News Detection Using Multi-Channel Deep Neural Networks. *International Journal on Semantic Web and Information Systems (IJSWIS)*, 18(1), 1-20.
4. Kaliyar, Rohit & Goswami, Anurag & Narang, Pratik & Sinha, Soumendu. (2020). FNDNet- A Deep Convolutional Neural Network for Fake News Detection. *Cognitive Systems Research*. 61. 10.1016/j.cogsys.2019.12.005.
5. Kaliyar, R.K., Goswami, A. & Narang, P. FakeBERT (2021): Fake news detection in social media with a BERT-based deep learning approach. *Multimed Tools Appl* 80, 11765–11788.
6. Junyi, Chen & Yan, Shankai & Wong, Ka-Chun. (2020). Verbal aggression detection on Twitter comments: convolutional neural network for short-text sentiment analysis. *Neural Computing and Applications*. 32. 10.1007/s00521-018-3442-0.
7. Nadeem, M.I., Ahmed, K., Li, D., Zheng, Z., Alkahtani, H.K., Mostafa, S.M., Mamyrbayev, O.J., & Abdel Hameed, H. (2022). EFND: A Semantic, Visual, and Socially Augmented Deep Framework for Extreme Fake News Detection. *Sustainability*.
8. Verma, P.K., Agrawal, P., Madaan, V., & Prodan, R. (2022). MCred: multi-modal message credibility for fake news detection using BERT and CNN. *Journal of Ambient Intelligence and Humanized Computing*, 1 - 13.

9. Aljrees, T., Cheng, X., Ahmed, M.M., Umer, M., Majeed, R., Alnowaiser, K., Abuzinadah, N., & Ashraf, I. (2023). Fake news stance detection using selective features and FakeNET. PLOS ONE, 18.
10. Viji, D., & Revathy, S. (2022). A hybrid approach of Weighted Fine-Tuned BERT extraction with deep Siamese Bi – LSTM model for semantic text similarity identification. Multimedia Tools and Applications, 81, 6131 - 6157.
11. S.Ananth, Dr.K.Radha, Dr.S.Prema, K.Niranjan International Journal of Innovative Research in Computer and Communication Engineering “Fake News Detection using Convolution Neural Network in Deep Learning”.
12. Granik, Mykhailo, andVolodymyr Mesyura. "Fake news detection using naive Bayes classifier." Electrical and Computer Engineering (UKRCON), 2017 IEEE First Ukraine Conference on. IEEE, 2017.
13. Fabula AI(UK), USI Lugano (Switzerland), Imperial college of London-Fake News Detection on Social Media using Geometric deep learning.