**An Integrated Multi-Task Model for Fake News Detection**

**Alternative Title:** Fake News using Bi-LSTM

**Aim**:

To enhance the assigning accuracy of former methods in fake news detection using advanced methods.

**Abstract:**

We are in the age of information, everytime we read a piece of information or watch the news on TV, we look for a reliable source. There are so many fake news spread allover the internet and social media. Fake news is misinformation or manipulated news that is spread across the social media with an intention to damage a person, agency and organization. The spread of misinformation in critical situations can cause disasters. Due to the dissemination of fake news, there is needfor computational methods to detect them. So, to prevent the harm that can be done using technology, we have implemented MachineLearning algorithms and techniques such as NLTK, LSTM. Our contribution is bifold. First, we must introduce the datasets which contain both fake and real news and conduct various experiments to organize fake news detector.

**Synopsis:**

The rise in popularity of social media sites like Facebook, Instagram, and Twitter, etc.., Fake news quickly spread across millions of users in a very short period of time. News with certain topics have high probabilities to be classified as fake news. Some authors have high probability to publish fake news and discussed in EDA part. We predict (based on topics).

**Existing System:**

Given the powerful automatic feature learning ability of deep learning, many works focus on detecting fake news by using machine Learning and deep neural networks based on RNN to capture the temporal and textual features from rumour posts in recent years.

**Problem Definition:** Fake news is one of the most serious problems in the modern age of the internet and social media. To detect a fake news with desire algorithm based on topic and also detecting the non fake news.

**Proposed System:**

We have proposed with the idea of using Bidirectional LSTM with GRU. So the bidirectional LSTM is capable of storing the previous sentence. So we can store the long titles and able to predict with the good accuracy.

**Advantage:**

Bi-LSTM are much better at handling long-term dependencies compared to RNN. This is due to their ability to remember information for extended periods of time. LSTMs are much less susceptible to the vanishing gradient problem.

**NLP, Bidirectional LSTM and GRU:**

We evaluated our detection scheme based on the title.

**Natural Language Processing:**

NLP is a field in machine learning with the ability of a computer to understand, analyze, manipulate, and potentially generate human language. It is the ability of a system to understand and process human languages. A computer system only understands the language of 0’s and 1’s, it does not understand human languages like English or Hindi. Natural Language Processing gave the computing system the ability to understand English or the Hindi language.

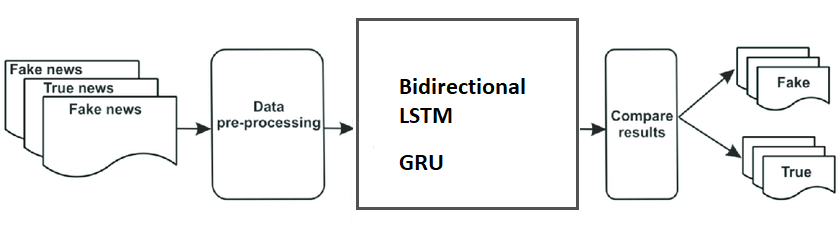
**Bidirectional LSTM:**

Bidirectional LSTM (BiLSTM) is a recurrent neural network used primarily on natural language processing. Unlike standard LSTM, the input flows in both directions, and it’s capable of utilizing information from both sides. It’s also a powerful tool for modeling the sequential dependencies between words and phrases in both directions of the sequence. In summary, BiLSTM adds one more LSTM layer, which reverses the direction of information flow. Briefly, it means that the input sequence flows backward in the additional LSTM layer. Then we combine the outputs from both LSTM layers in several ways, such as average, sum, multiplication, or concatenation.

**GRU:**

The Gated Recurrent Unit (GRU) is a type of Recurrent Neural Network (RNN) that, in certain cases, has advantages over long short term memory (LSTM). GRU uses less memory and is faster than LSTM, however, LSTM is more accurate when using datasets with longer sequences. Also, GRUs address the vanishing gradient problem (values used to update network weights) from which vanilla recurrent neural networks suffer. If the grading shrinks over time as it back propagates, it may become too small to affect learning, thus making the neural net untrainable.

**Architecture Diagram:**

­­­­

**Modules:**

* Data Preprocessing
* NLP, Bidirectional LSTM and GRU
* Web Framework

**Data Preprocessing:**

The first and foremost step in data processing is collecting the dataset. We have collected a dataset based on Twitter spam data. The dataset is a CSV file format data which consists of n number of Twitter spam Data. We need to select or extract the features from the collected dataset. Then the Data Cleaning should be initiated. Thus in this module data preprocessing will be completed.

**Hardware Requirements:**

* Hard Disk : 500GB and Above
* RAM : 4GB and Above
* Processor : I3 and Above

**Software Requirements:**

* Operating System : Windows 7 , 8, 10 (64 bit)
* Software : Python 3.7
* Tools : Anaconda (Jupyter Note Book IDE)
* Framework : Flask

**Conclusion:**The majority of activities in the twentyfirst century are done online. Newspapers, which were once preferred in hard copies, are now being replaced by applications such as Facebook, Twitter, and news articles. The growing issue of fake news further complicates matters by attempting to sway people's opinions and attitudes about the use of digital technology. When a person is misled by real news, one of two things can happen: people begin to believe that their assumptions about a specific topic are right. To fix this issue, we built our Fake News Detection system using NLP, Bi-LSTM and GRU.

**Future Work:**

System has the potential to provide an impulse to various emerging applications such as controlling the spread of fake news during elections, terrorism, natural calamities, crimes for the betterment of society. In the future, the efficiency and accuracy of the prototype can be enhanced to a certain level.