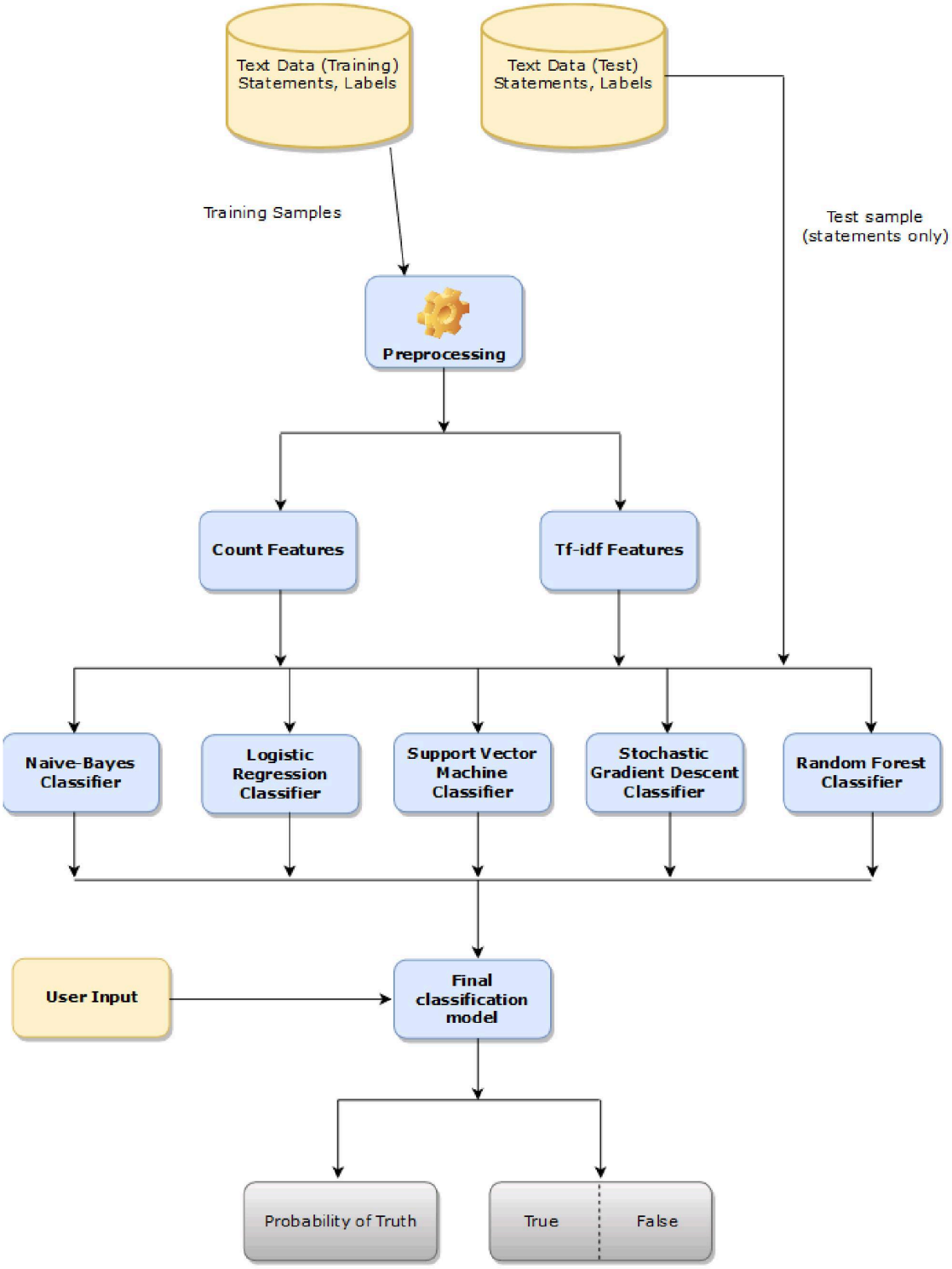
FAKE NEWS DETECTION

Team members:

THIRUMALAI B	821021104049
TAMILSELVAN M	821021104048
NITHISH KUMAR S	821021104035
KISHORE G M	821021104027

Fake News Detection Using NLP

The following figure illustrates the flow or the process or the steps involved in the fake news detection.



Problem Definition:

Fake news detection using natural language processing (NLP) is a challenging and evolving field. The goal is to develop models and techniques that can automatically distinguish between genuine and misleading news articles or information. Here's a general outline of how you can approach building a fake news detection system using NLP

Model Selection:

Choose appropriate machine learning or deep **learning models for classification. Common choices include:**

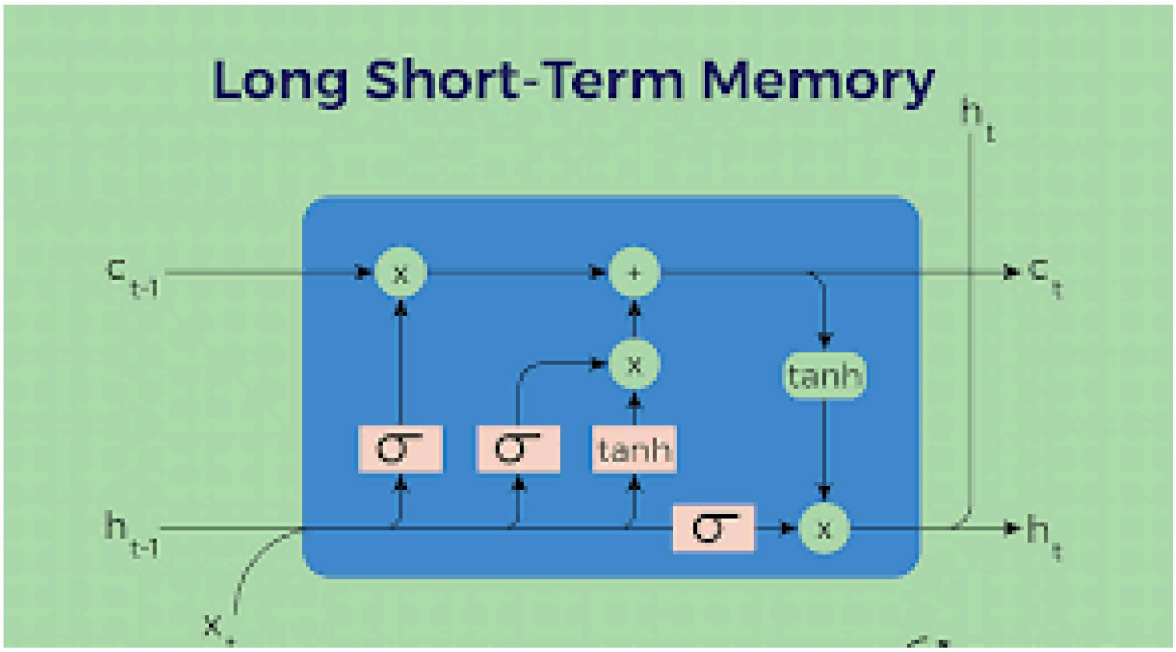
- **Logistic Regression**
- **Naive Bayes**
- **Support Vector Machines (SVM)**
- **Random Forest**
- **Neural networks (e.g., LSTM, CNN)**

LSTM:

Long short-term memory (LSTM) network is a recurrent neural network (RNN), aimed to deal with the vanishing gradient problem present in traditional RNNs. Its relative insensitivity to gap length is its advantage over other RNNs, hidden Markov models and other sequence learning methods.

What is LSTM model used for?

LSTMs are predominantly used to learn, process, and classify sequential data because these networks can learn long-term dependencies between time steps of data. Common LSTM applications include sentiment analysis, language modeling, speech recognition, and video analysis.



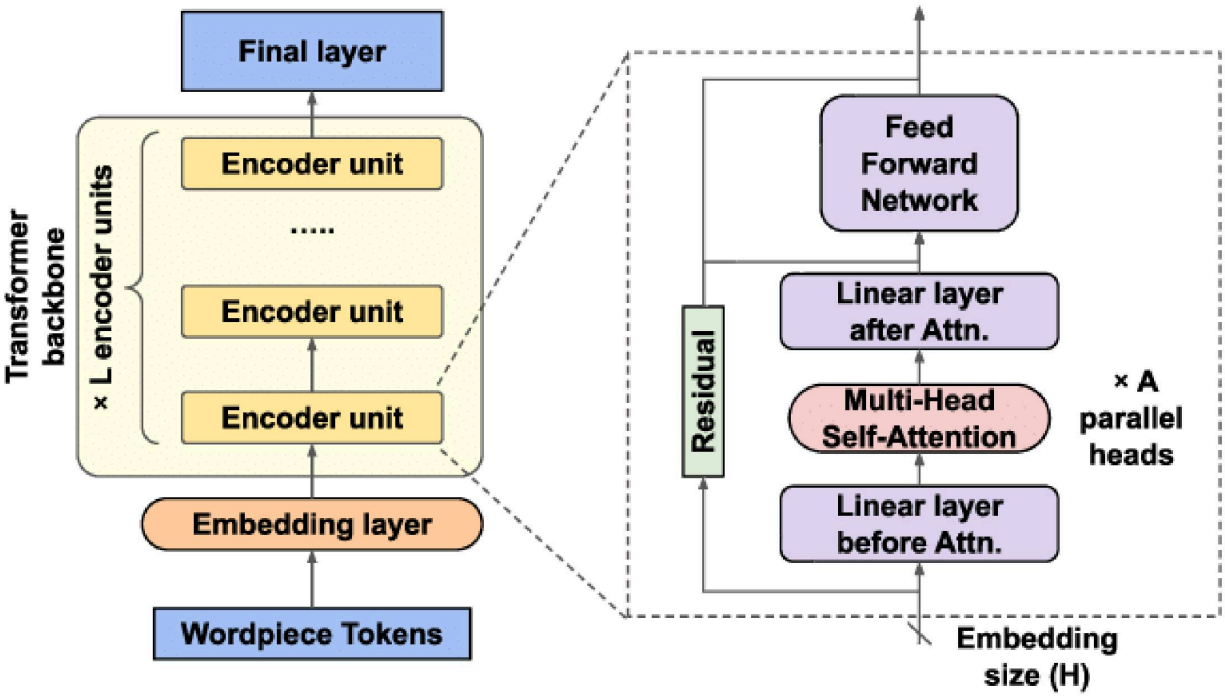
What is a BERT?

BERT, short for Bidirectional Encoder Representations from Transformers, is a machine learning (ML) framework for natural language processing.

What is BERT used for?

BERT is designed to help computers understand the meaning of ambiguous language in text by using surrounding text to establish context. The BERT framework was pre-trained using text from Wikipedia and can be fine-tuned with question and answer datasets.

https://www.kaggle.com/code/clmentbisaillon/classifying-fake-news-with-bert



ACCURACY

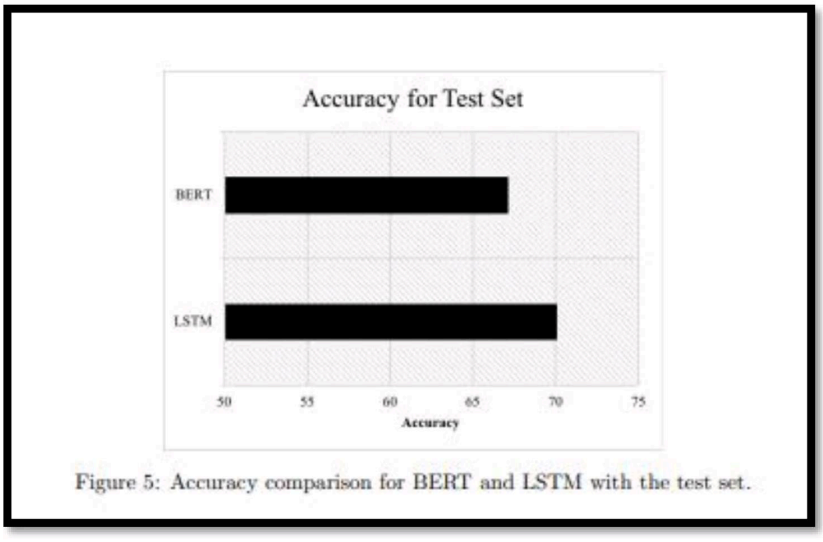


Figure 5: Accuracy comparison for BERT and LSTM with the test set.

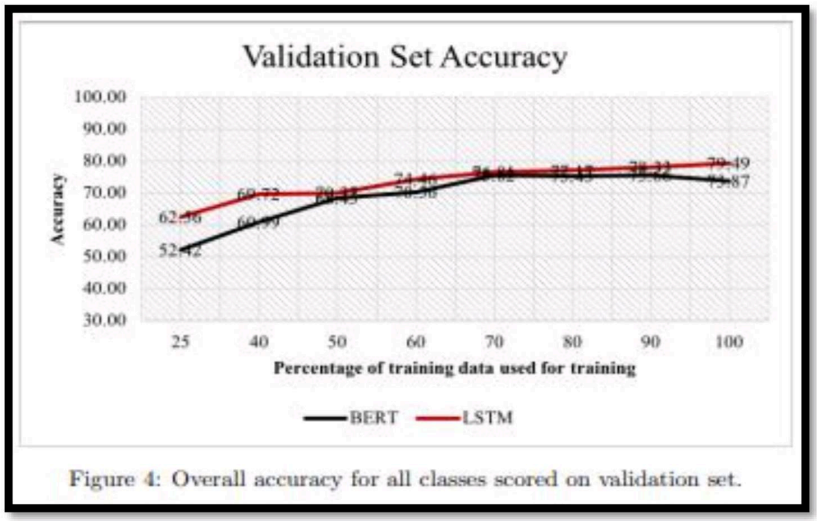


Figure 4: Overall accuracy for all classes scored on validation set.

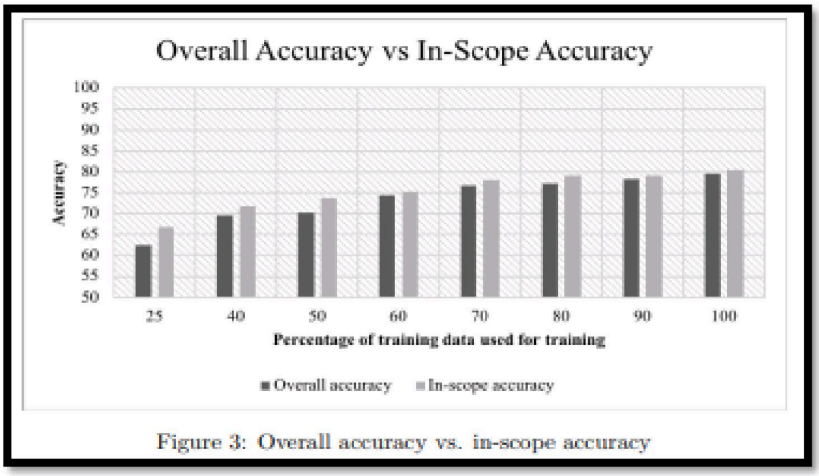


Figure 3: Overall accuracy vs. in-scope accuracy

conclusion:

The final model classifier is derived from combining the other classifiers such as random forest, support vector machine classifier, stochastic gradient descent, naive bayes classifier, logistic regression classifier. When the user give input to this final model classifier, this model classifies the probability of the input or to determine whether the given input news is true or fake.