TITLE: FAKE NEWS DETECTION USING NLP

INNOVATION TECHNIQUES DOCUMENTATION

TEAM MEMBERS :-	REG NO
1. M. BHARANI DHARAN [LEADER]	731221104302
2. S. ARULMURUGAN.	731221104004
3. S. VASANTH KUMAR.	731221104038
4. S. SURYA.	731221104036
5. V. SRINIVAS.	731221104307

AGENDA:-

1. **INTRODUCTION**
2. **ENSEMBLE METHOD**
3. **DATA COLLECTION**
4. **DATA PREPROCESSING**
5. **WORD EMBEDDING**
6. **MODEL ARCHITECTURES**
7. **TRAINING**

INTRODUCTION:

Overview of the project and the importance of AI in “Fake News Detection”.

Emphasize the focus on improving accuracy and robustness through innovative techniques.

ENSEMBLE METHOD:

Neural Network Ensembles: [**LSTM**]

Ensemble methods can also be applied to neural networks, including LSTM(Long Short-Term Memory) networks, to improve their performance.

Ensembling can be done directly within neural networks by creating ensemble layers.In this approach, multiple LSTM subnetworks are combined within a larger neural network architecture.

MODEL SELECTIONS:

Choose a variety of NLP models for your ensemble. Some common choices include:

- Logistic Regression
- Random Forest
- Gradient Boosting (e.g., XGBoost, LightGBM)
- Support Vector Machines (SVM)
- Neural Networks (e.g., LSTM, CNN, BERT

DATA COLLECTION:

- ➔ Fake news detection using LSTM (Long Short-Term Memory) networks involves collecting a dataset of news articles labelled as real or fake.
- ➔ Data collection is a critical step in building an effective fake news detection model.

DATA PREPROCESSING:

Tokenization:➔ Split the text into words or subwords.

Text Cleaning:➔ Remove special characters, punctuation, and unnecessary whitespaces.

Stop Word Removal:➔Eliminate common stop words that don't carry significant meaning.

Padding:➔Ensure that all sequences have the same length by padding shorter sequences.

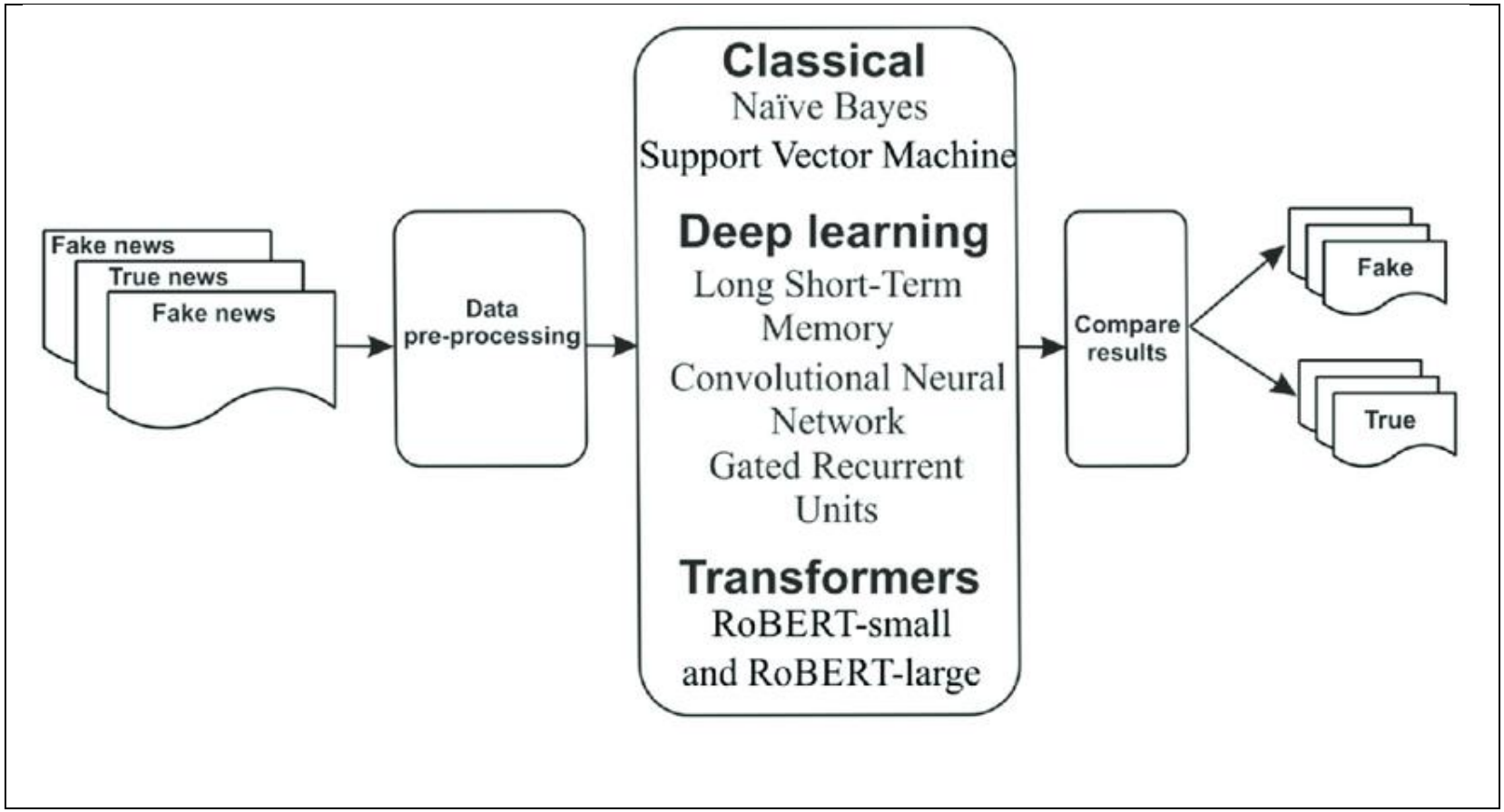
WORD EMBEDDINGS:

- ➔ Convert words into numerical vectors using pre-trained word embeddings.
- ➔ It like Word2Vec, GloVe, or FastText. These embeddings capture semantic relationships between words.

NLP AS A VALUABLE TOOL:-

NLP techniques play a pivotal role in identifying fake news by analyzing and understanding the content of textual information. NLP enables the extraction of valuable features and patterns from text data

MODEL ARCHITECTURE:



Build an LSTM-based neural network for text classification. A simple model might look like this using libraries like TensorFlow or PyTorch.

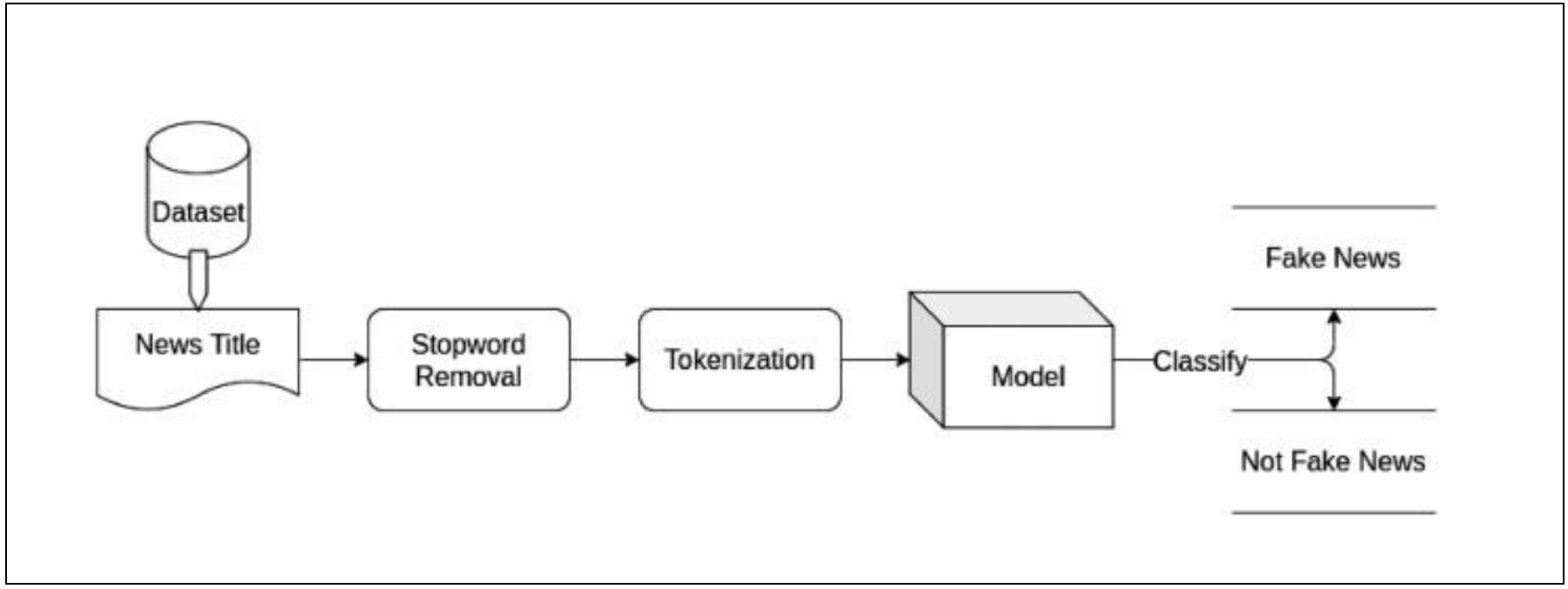
Experiment with different architectures, including variations of LSTMs, bidirectional LSTMs, and adding dropout layers to prevent overfitting.

TRAINING:

- ➔ Split your dataset into training, validation, and test sets.
- ➔ Train your LSTM model on the training data using binary cross-entropy loss and an optimizer like Adam.
- ➔ Monitor the validation accuracy to avoid overfitting.

EVALUATION:

Evaluate your model on the test dataset using metrics such as accuracy, precision, recall, F1-score, and ROC AUC.



THANK YOU!