

FAKE NEWS DETECTION

By
Prajwal Gutti



PROBLEM STATEMENT

Design a AI model to detect fake news by analysing the text. The goal is to help people trust the information they read by news. This system will reduce the spread of fake news on social platforms.

DEMONSTRATION

The screenshot shows a Jupyter Notebook interface with a dark theme. The top bar displays tabs for 'You are signed in as becuser', 'Minimalist Template - Presenta...', 'Home', and 'Deep learning BiLSTM model'. The main area has a toolbar with file operations like 'File', 'Edit', 'View', 'Run', 'Kernel', 'Settings', 'Help', and a 'Code' dropdown. A status bar at the bottom indicates 'Trusted' and 'JupyterLab'.

The notebook content consists of two code cells:

```
#Training accuracy of the BiLSTM model
train_loss, train_accuracy = model.evaluate(X_train, y_train, verbose=0)
print(f"Training Accuracy: {train_accuracy * 100:.2f}%")
#Testing accuracy of the BiLSTM model
test_loss, test_accuracy = model.evaluate(X_test, y_test, verbose=0)
print(f"Testing Accuracy: {test_accuracy * 100:.2f}%")

Training Accuracy: 99.65%
Testing Accuracy: 99.51%
```

```
[49]: import nltk
import re
import string
from nltk.corpus import stopwords
nltk.download('punkt')
nltk.download('stopwords')
nltk.download('wordnet')
from nltk.tokenize import word_tokenize
from nltk.stem import WordNetLemmatizer, SnowballStemmer
from sklearn.feature_extraction.text import TfidfVectorizer
from rich.console import Console

# Initialize the console object
console = Console()

# Initialize the lemmatizer and stemmer
lemmatizer = WordNetLemmatizer()
stemmer = SnowballStemmer("english")

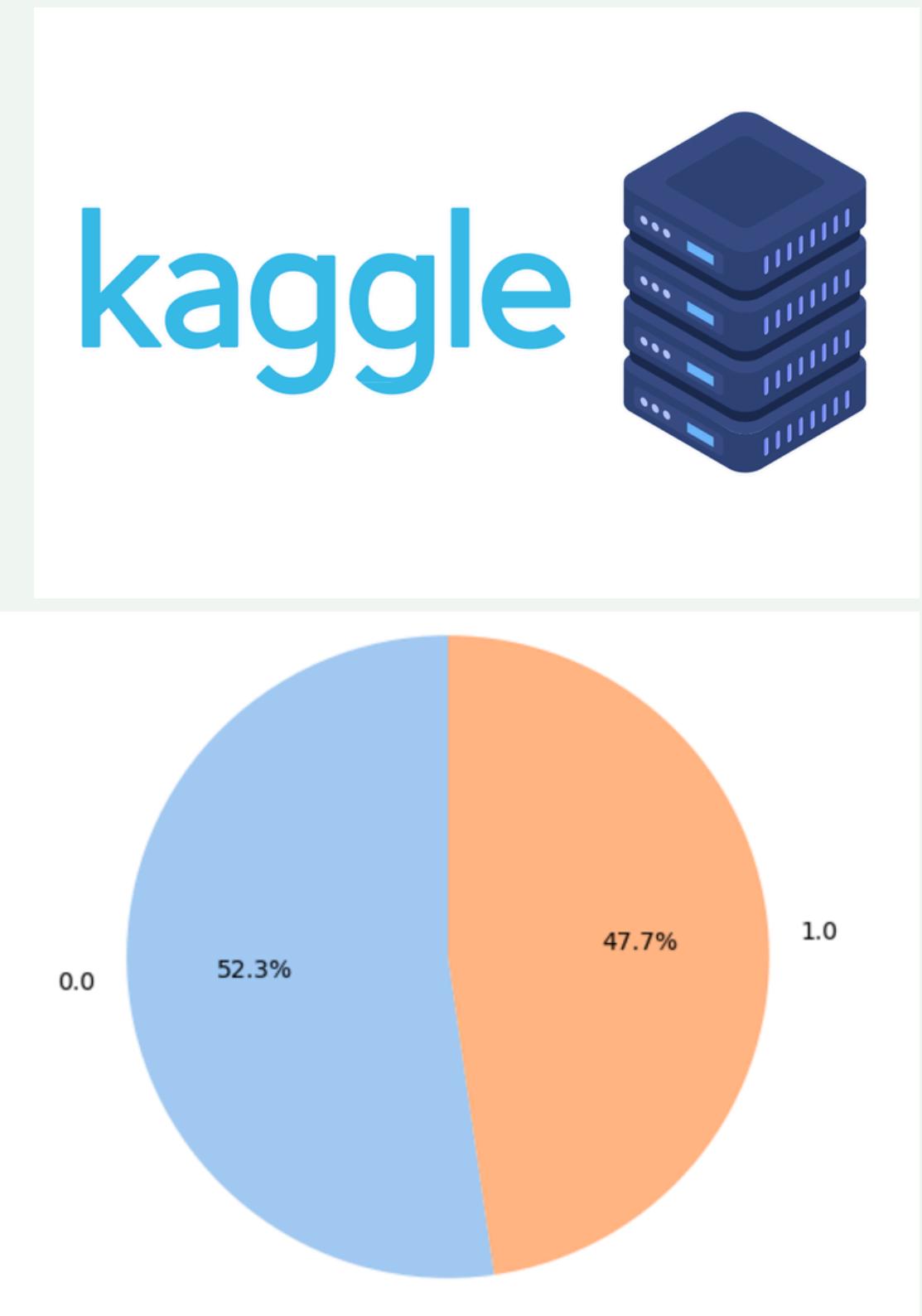
# Function to process text
def process_text(text):
    # Step 1: Convert text to lowercase and remove special characters
    text = re.sub(r'[^\\w\\s]', '', text.lower()) # Remove special characters and convert to Lowercase
    text = re.sub(r'\\s+', ' ', text).strip() # Replace multiple spaces with a single space

    # Tokenize the clean text
    tokens = word_tokenize(text)

    # Step 2: Remove stopwords
    stop_words = set(stopwords.words('english'))
```

DATA COLLECTION

Collected data for fake news detection from Kaggle, which includes a dataset of 39,105 entries. The dataset contains three columns: the title of the news, the text of the news, and a label that indicates whether the news is fake or real.



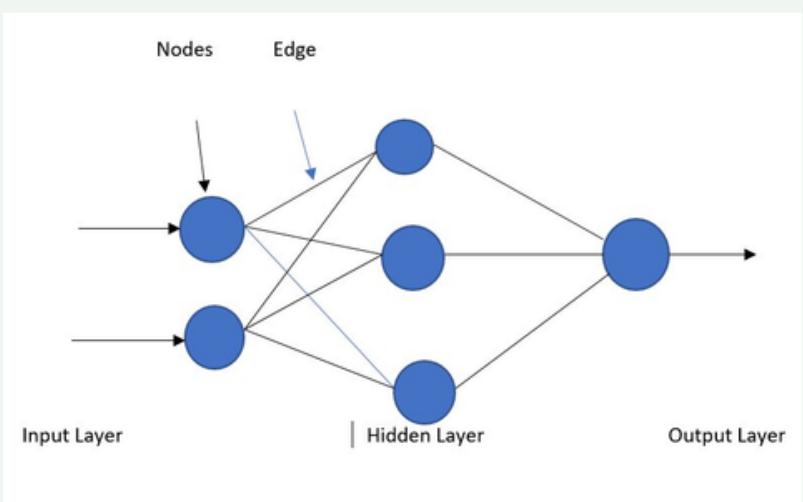
DATA PREPROCESSING

This included converting text to lowercase, removing special characters, tokenizing, and removing stopwords. Lemmatization and stemming were used to simplify words. These steps help the model focus on meaningful patterns in the text.

MODELS USED

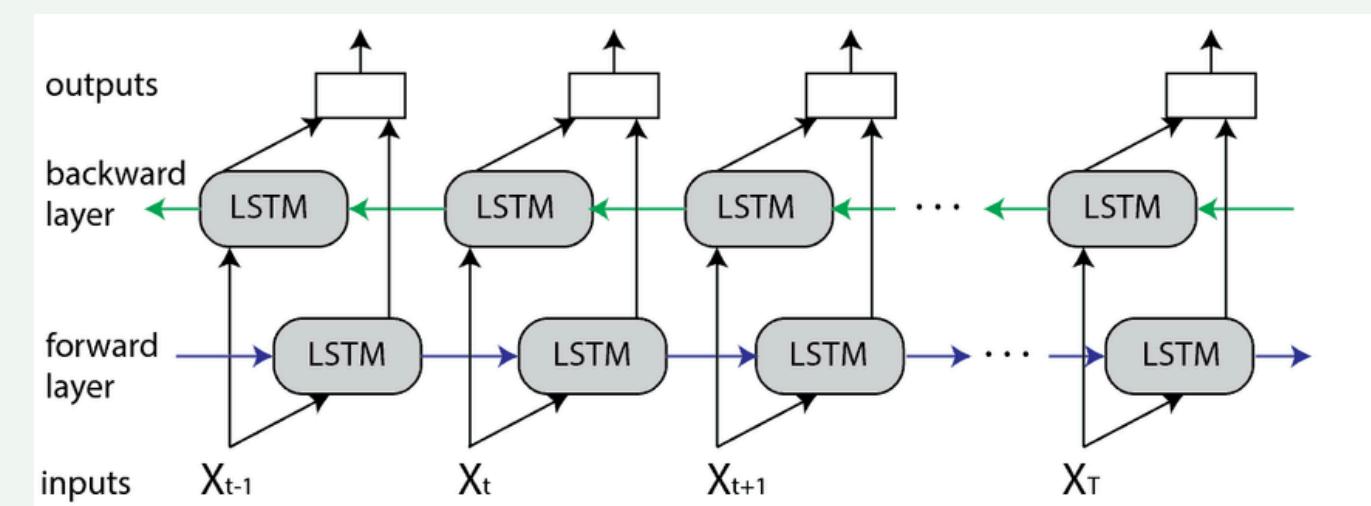
MLP Classifier

The MLP model is a simple neural network consisting of an input layer, two hidden layers (64 and 32 neurons), and an output layer that classifies news as fake or real based on extracted features.



BiLSTM

The Bidirectional LSTM (BiLSTM) model consists of an input layer, two bidirectional LSTM layers, and an output layer that classifies news as fake or real news.



HYPERPARAMETER TUNNING

BiLSTM

The BiLSTM model is compiled with binary cross-entropy loss for binary classification and the Adam optimizer for efficient training. During training, the model is tested on the validation dataset to monitor performance, with a batch size of 32 and 5 epochs for the number of training iterations. The training process is verbose, providing detailed progress updates.

PERFORMANCE METRICS

BiLSTM

The BiLSTM model achieved a training accuracy of **99.96%**, demonstrating a strong fit to the training data. The testing accuracy of **99.08%** indicates excellent generalization to unseen data, with a minimal drop, showcasing the model's robust performance on the test set.

Precision and Recall is around **97.25%**



CONCLUSION

The implementation of a BiLSTM based model for fake news detection has demonstrated promising results, achieving a high training accuracy of 99.96% and a testing accuracy of 99.08%. This reflects the model's ability to effectively learn and generalize from the dataset, showcasing its robustness in distinguishing between fake and real news

THANK YOU

