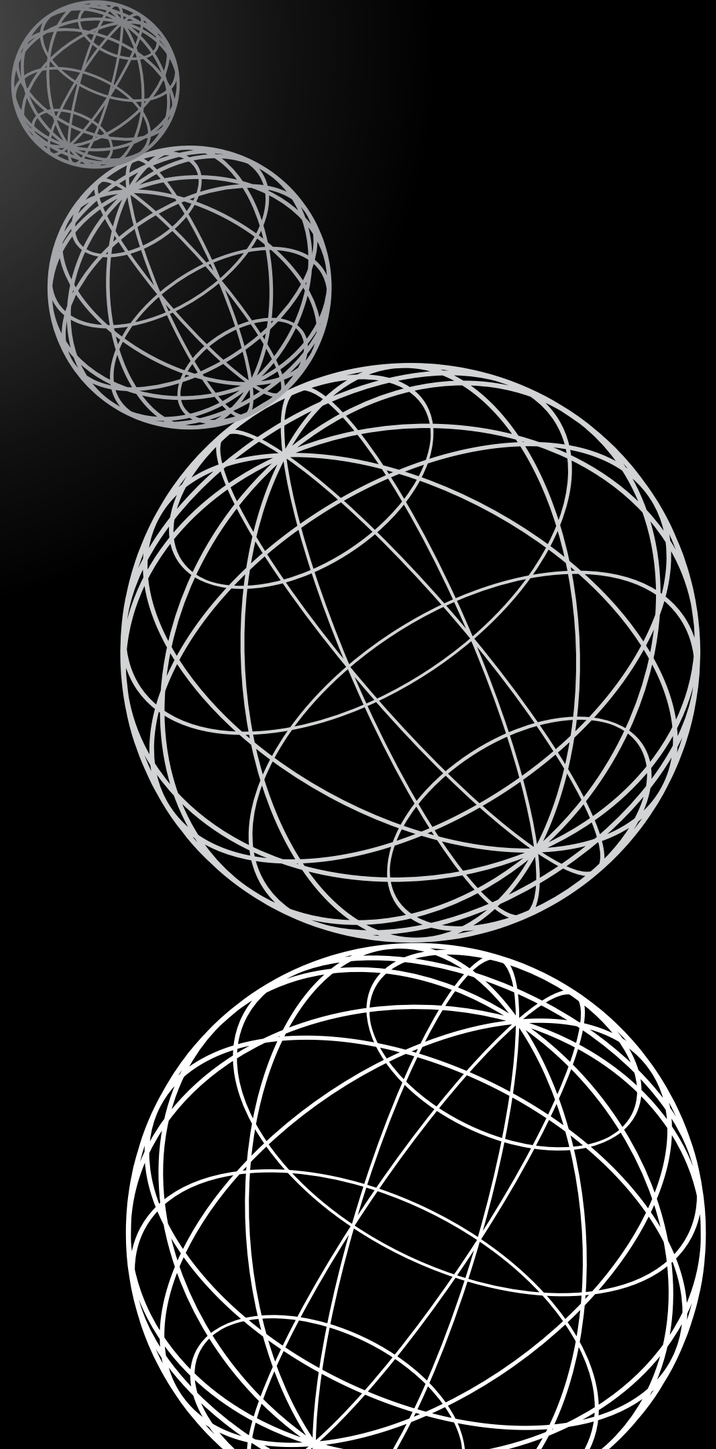
COMPARATIVE STUDY OF LSTM VS. BERT FOR FAKE NEWS DETECTION

BASED ON ISOT FAKE NEWS DATASET

CS22B2029-MARAM MOUNISH
CS22B2024-DAMPETLA HARSHA
CS22B2032-YATHIN



PROBLEM STATEMENT

Fake news is a rising threat in digital communication, influencing public opinion and spreading misinformation. Detecting fake news automatically is essential for information credibility.

Goal: Compare traditional RNN-based (LSTM) and transformer-based (BERT) models for detecting fake news.

MODEL ARCHITECTURE

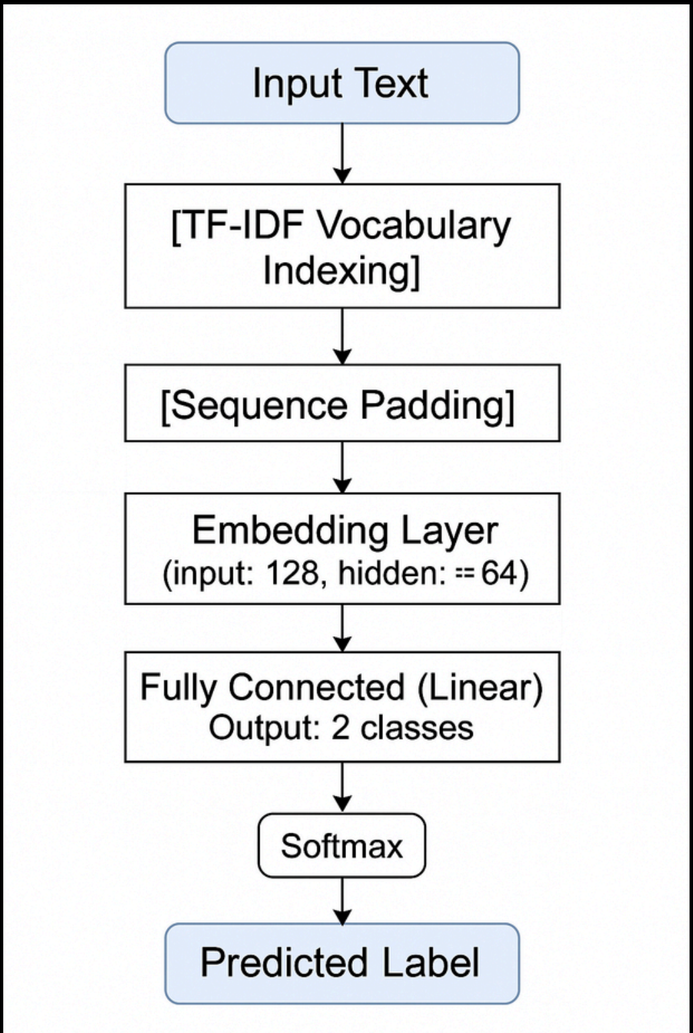
LSTM:

- Recurrent neural network that captures sequential patterns.
- Handles time dependencies well but limited in understanding long-term context.

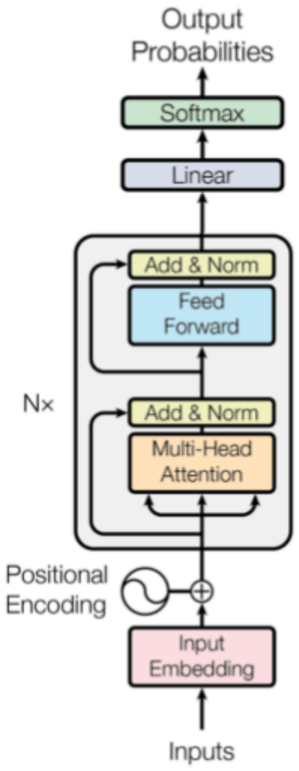
BERT:

- Transformer-based pre-trained model.
- Uses attention mechanism to capture context in both directions.

LSTM



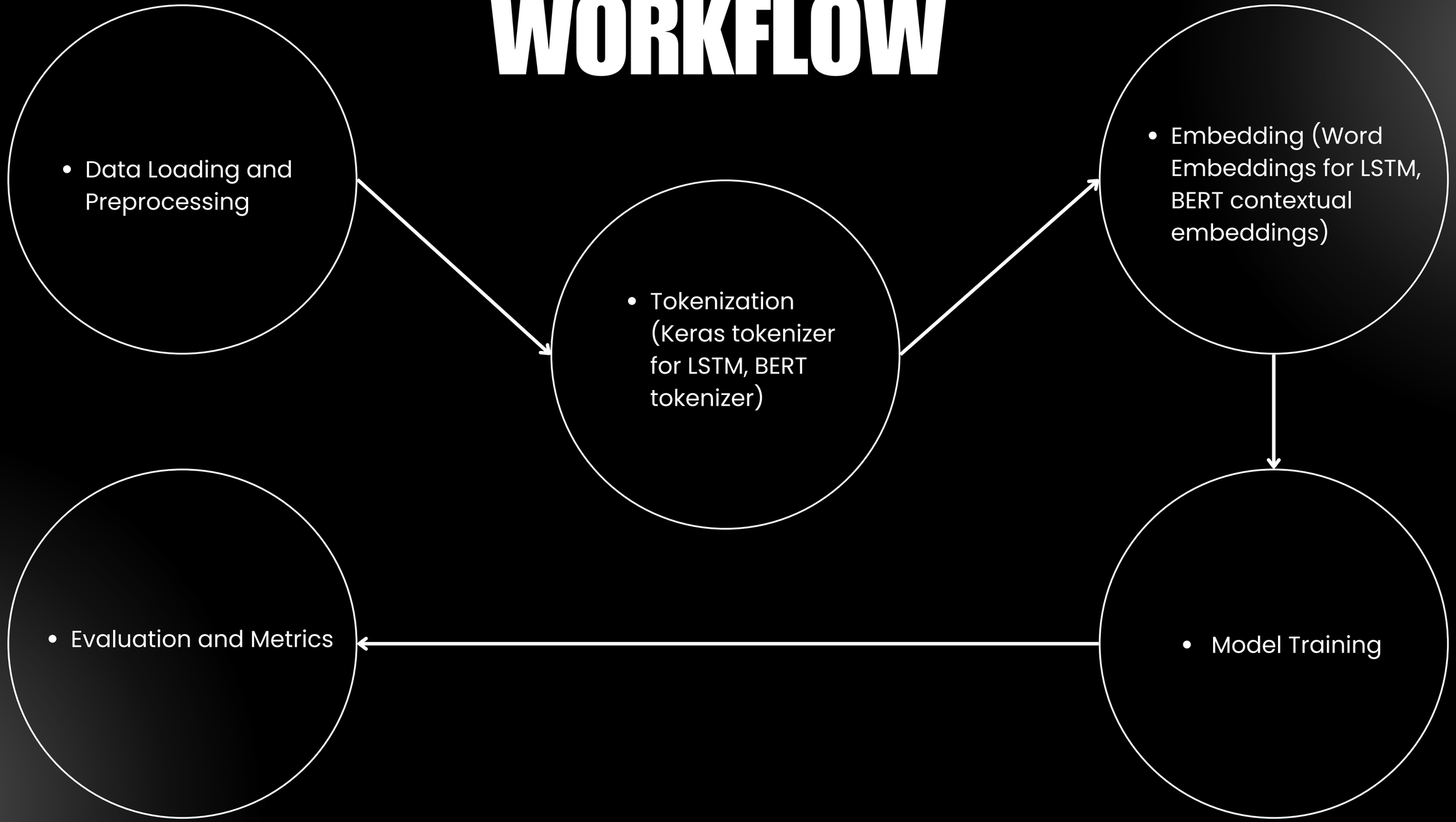
BERT



BERT Base

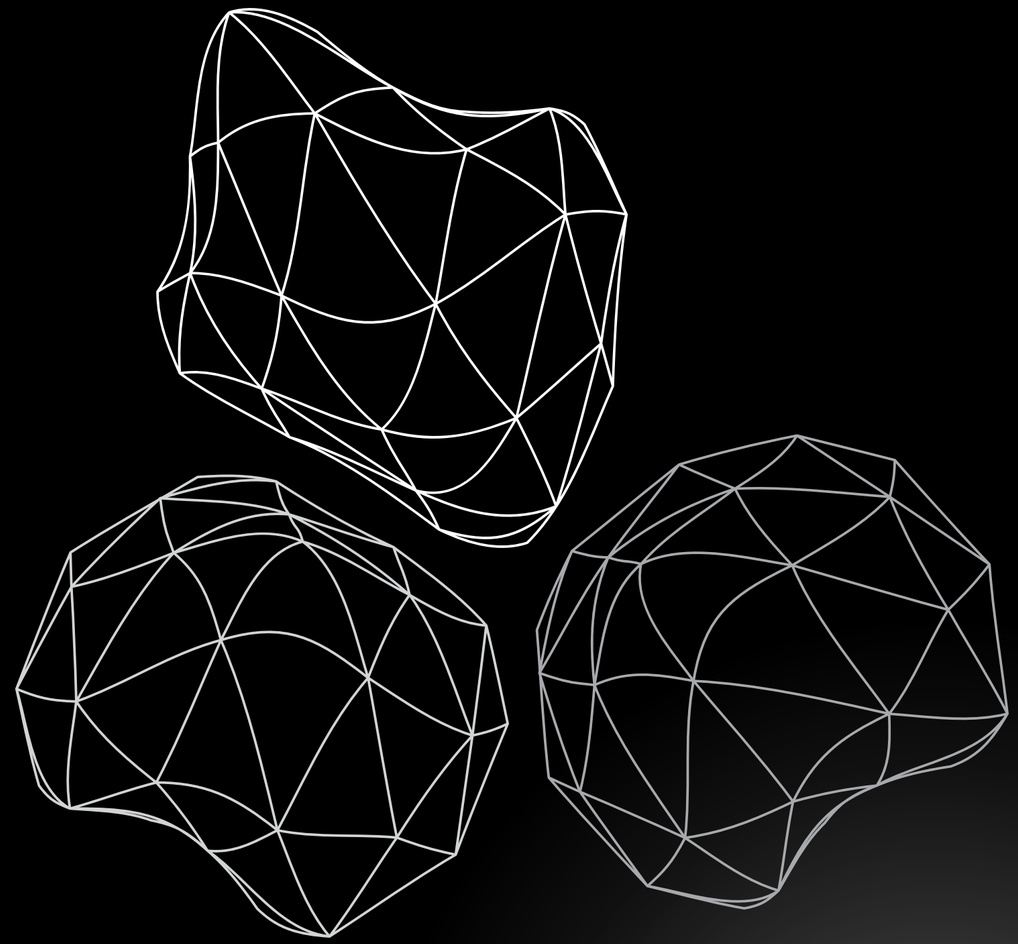
Number of embedding dimensions, d_model	768
Number of encoder blocks, N:	12
Number of attention heads per encoder block:	12
Size of hidden layer in feedforward network:	3072
Total parameters:	110 million

WORKFLOW


• Data Loading and Preprocessing
• Tokenization (Keras tokenizer for LSTM, BERT tokenizer)
• Embedding (Word Embeddings for LSTM, BERT contextual embeddings)
• Model Training
• Evaluation and Metrics

DATASET

- ***Dataset Used*** : ISOT Fake News Dataset
- ***Classes***: Real News, Fake News
- ***Size***:
  - Real: 21,417 samples
  - Fake: 23,481 samples
  - Split: Train(80%)/Test(20%)



METHODOLOGY

BERT

- Using HuggingFace Transformers
- Pre-trained bert-base-uncased model
- Fine-tuning with added classification head
- Tokenized inputs with special tokens [CLS], [SEP]
- Training with AdamW optimizer, learning rate scheduler

LSTM

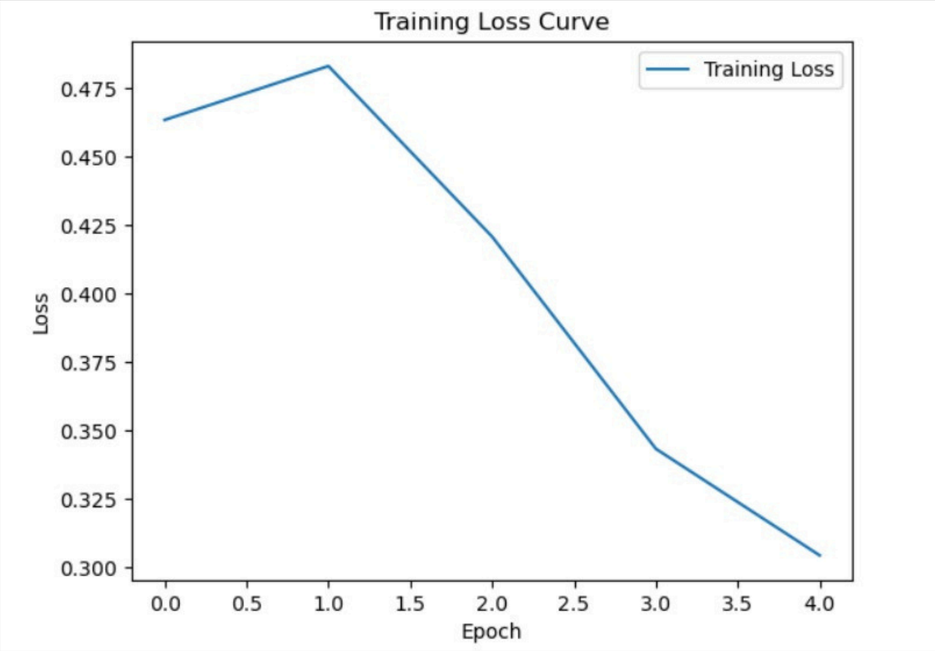
- Preprocessing with Keras tokenizer
- Embedding layer with pre-trained or random weights
- LSTM layer with dropout
- Dense output layer (sigmoid for binary classification)

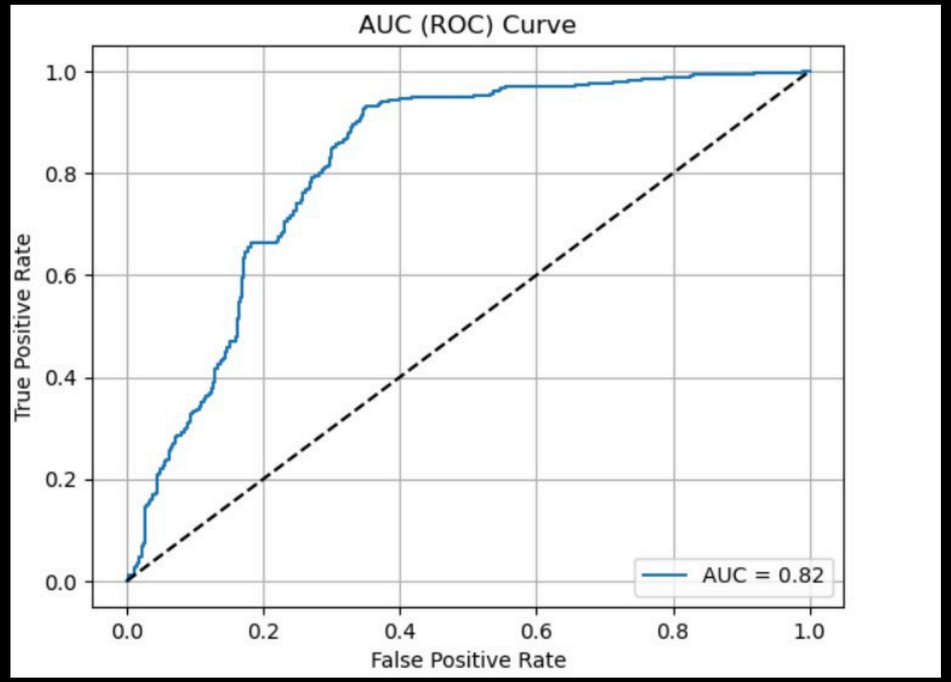
METRICS

Metric	LSTM Model	BERT Model
Accuracy	96.98%	99.98%
AUC	0.9945	0.999
Precision	0.9735	1.00
EER	0.0289	0.00

Visualization

LSTM

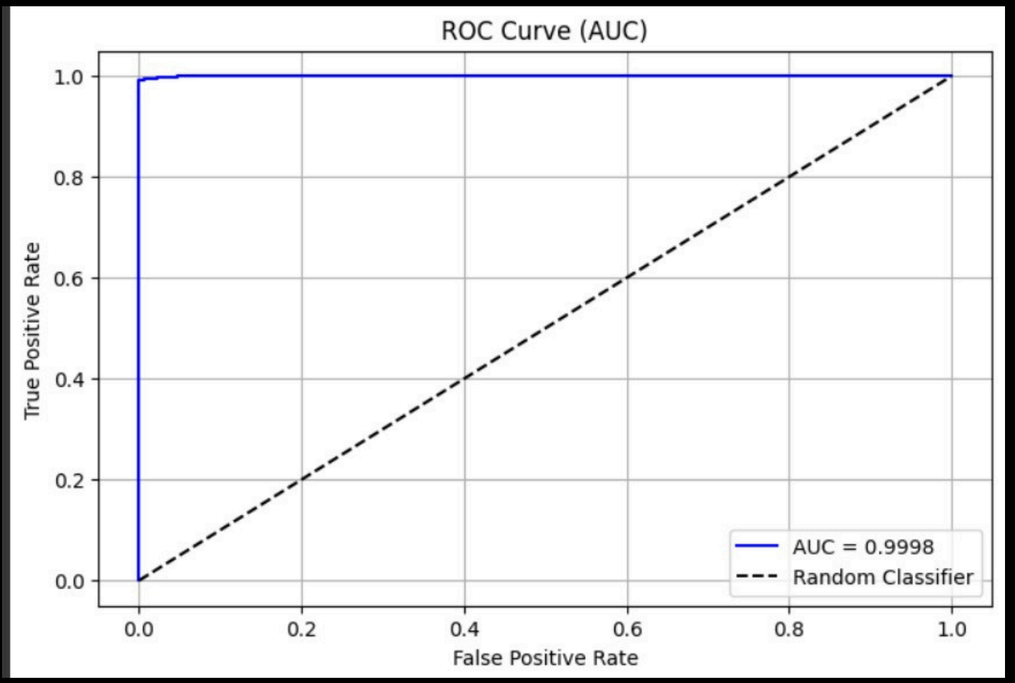


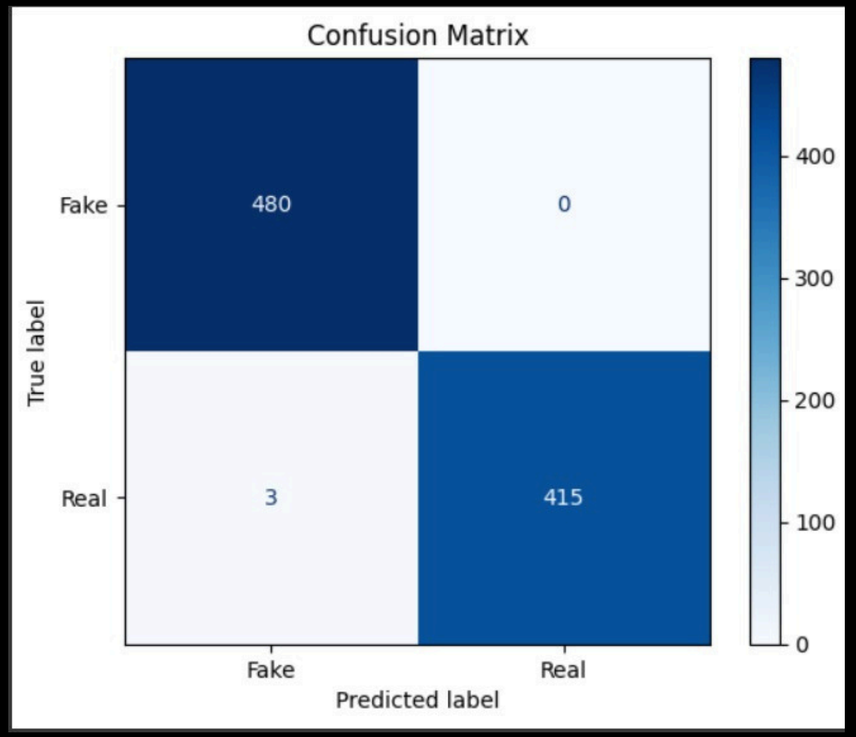



```
Epoch 1, Loss: 0.4811
Epoch 2, Loss: 0.1034
Epoch 3, Loss: 0.0630
Epoch 4, Loss: 0.0237
Epoch 5, Loss: 0.0076

Test Set Performance:
Accuracy: 0.9698
AUC: 0.9945
Precision: 0.9735
EER: 0.0289
```

Visualization

BERT





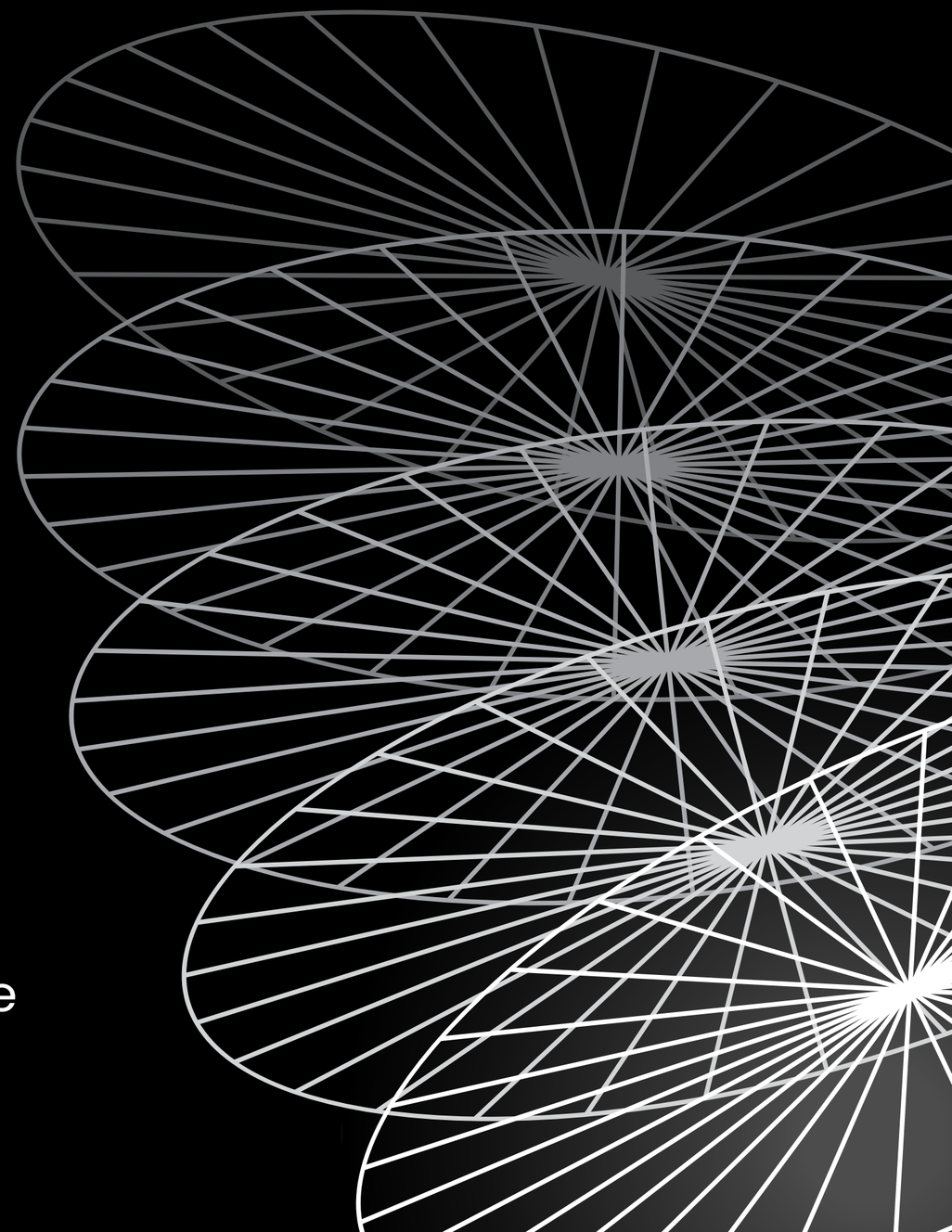

```
Confusion Matrix:

[[4669    0]
 [   1 4310]]
Accuracy Score: 0.9998886414253898
AUC Score: 0.9999768979214538
Precision: 1.0
EER: 0.0
EER Threshold: 0.9999515
```

ANALYSIS

- LSTM Pros
  - Lightweight, fast to train
  - Suitable for simple sequential tasks

- BERT Pros:
  - Much better context understanding
  - State-of-the-art in many NLP tasks

- Cons:
  - LSTM: Poor long-term dependency capture
  - BERT: Heavy computation and memory



CONCLUSION

- Implemented fake news detection using deep learning models: BERT and LSTM.
- Trained models on a labeled dataset of True.csv and Fake.csv news articles.
- BERT outperformed LSTM in accuracy, precision, and recall due to its deep contextual understanding.
- Demonstrated the potential of NLP models in combating misinformation.
- Results show feasibility of using such models in real-world news verification systems.

THANK YOU