Fake News Detection - Report 1

# 1. Introduction

This project focuses on building a machine learning model to detect fake news articles using BERT.

# 2. Dataset Description

The dataset consists of two files: 'fake.csv' and 'true.csv', which contain labeled fake and real news articles.

# 3. Methodology

Text data is preprocessed and tokenized using BERT tokenizer. A BERT-based model is fine-tuned for binary classification.

# 4. Model Training

The model is trained using Hugging Face Transformers and PyTorch. AdamW optimizer is used with a learning rate of 2e-5.

# 5. Results

Evaluation is done using classification metrics like precision, recall, F1-score, and accuracy.

# 6. Conclusion

BERT performs significantly well on the fake news detection task. Future improvements could include better text cleaning and using larger datasets.

# 7. Future Work

Try advanced architectures like RoBERTa, add explainability, or deploy the model in a web application.