

**Project Report**

**Fake News Detection Using Transformer Models**

Course: CS4045 - Deep Learning For Perception

Instructor: Ms. Sumaiyah Zahid

Group Members:

Mahnoor Sohail (21k-3834)

Hadia Syed (21k-3924)

**1. Objective**

The primary goal of this project is to detect and classify news articles as real or fake by analyzing their headlines and body text using advanced transformer-based Natural Language Processing (NLP) techniques. The project aims to achieve high accuracy and generalizability by comparing the performance of pretrained models (BERT, BART) and a custom-trained transformer built from scratch.

**2. Problem Statement**

In today’s digital landscape, fake news spreads rapidly, influencing public opinion, politics, and public health. Manual detection is ineffective at scale. Therefore, there is a strong need for an automated and intelligent solution capable of analyzing textual data and classifying articles as either genuine or fabricated.

**3. Methodology**

1. **Dataset**

* Source: Kaggle (Real and Fake News Dataset)
* Format: Two CSV files: True.csv, Fake.csv
* Columns: title, text, subject, date
* Preprocessing:
  + Combined title and text as the main content
  + Removed empty or irrelevant text entries
  + Labeled: Real → 1, Fake → 0

1. **Models Used**
2. **BERT (Pretrained)**
   * bert-base-uncased fine-tuned on the dataset
   * Uses WordPiece tokenizer and segment embeddings
3. **BART (Pretrained)**
   * Sequence-to-sequence model repurposed for classification
   * Trained using the encoder output and final decoder logits
4. **Custom Transformer**
   * Built from scratch using PyTorch
   * 4-layer encoder architecture, self-attention, and classification head
   * Tokenized with BertTokenizer for consistency
5. **Training Setup**

* Platform: Google Colab (T4 GPU runtime)
* Loss Function: CrossEntropyLoss
* Optimizer: AdamW
* Scheduler: Linear learning rate decay
* Batch Size: 16
* Epochs: 3–5
* Evaluation Metrics: Accuracy, Precision, Recall, F1-Score

**4. Results**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Model | Accuracy | Precision (avg) | Recall (avg) | F1 Score (avg) |
| BERT | 96.26% | 97% (weighted) | 96% | 96% |
| BART |  |  |  |  |
| Custom Transformer | 99.86% | 100% | 100% | 100% |

**5. Conclusion**

This project demonstrates that transformer-based architectures are highly effective in text classification tasks like fake news detection. Pretrained models like BERT give excellent results with minimal training, while a custom transformer offers flexibility and performance trade-offs. The findings suggest that such models can be integrated into real-world fact-checking systems to combat misinformation online.

**6. References**

* What are Transformers (YouTube [Link](https://www.youtube.com/watch?v=ZXiruGOCn9s))
* Transformer models and BERT model (https://www.youtube.com/watch?v=t45S\_MwAcOw)
* Hugging Face Transformers Documentation: <https://huggingface.co/transformers/>
* Hugging Face BERT Documentation: https://huggingface.co/docs/transformers/en/model\_doc/bert
* Dataset: https://www.kaggle.com/clmentbisaillon/fake-and-real-news-dataset