BERT and Domain-Specific Transformers for Misinformation Detection:

This project focuses on developing a deep learning pipeline for fake news detection using BERT-based transformer models. The dataset used is from Constraint@AAAI-2021, containing 10,600 COVID-19-related English social media posts labeled as real or fake.

Approach:

The workflow begins with preprocessing the text data, including tokenization, normalization, and handling of noisy social media text. Four transformer-based models were fine-tuned: bert-base-uncased, covid-twitter-bert, twhin-bert-base, and socbert. Fine-tuning was conducted using the Hugging Face AutoModelForSequenceClassification.

Hyperparameter Optimization:

Key parameters such as learning rate, batch size, and number of epochs were optimized through a combination of manual experiments and Optuna-based tuning. The goal was to maximize classification performance while avoiding overfitting. Evaluation: Models were evaluated on a held-out test set using accuracy, precision, recall, and F1-score (both micro and macro averages). Confusion matrices provided deeper insights into classification errors. Among the tested models, covid-twitter-bert achieved the best results with 88% accuracy and an F1-score of 0.85, outperforming the baseline bert-base-uncased model.

Findings:

- Domain-specific BERT models such as covid-twitter-bert performed better than generic BERT in handling pandemic-related misinformation.
- Hyperparameter tuning significantly improved performance over default configurations.
- The approach demonstrates that transformer-based architectures are effective for misinformation detection tasks, especially when fine-tuned on domain-specific text.

This project showcases the application of modern NLP techniques to real-world problems like misinformation detection and highlights the impact of domain adaptation in transformer-based models.