AI/ML Intern Challenge Submission Report

Title:	Advanced Sentence Transformation Classification

Introduction

This project aims to develop an NLP-based classification model that identifies different types of English sentence transformations, such as Active ↔ Passive, Direct ↔ Indirect Speech, and Positive ↔ Negative conversions. The goal was to build a robust, interpretable, and production-ready solution that demonstrates strong generalization across transformation types.

Objective

To create a model capable of automatically detecting the type of transformation between two sentences using advanced deep learning architectures and explainable AI tools.

Methodology

Dataset Preparation: A custom dataset of 200+ English sentence pairs was created, balanced across six transformation categories: Active ↔ Passive, Direct ↔ Indirect, and Positive ↔ Negative.

Data Preprocessing: Text cleaning, tokenization (BERT tokenizer), and label encoding were applied.

Model Architecture: Fine-tuned BERT (bert-base-uncased) with a dense Softmax classification head. Optimizer: AdamW. Loss: Cross-Entropy. Data split: 60:20:20.

Explainability: SHAP and attention visualization were used for interpretability.

Evaluation Metrics: Accuracy, Precision, Recall, F1-score, and Confusion Matrix for class-level insights.

Results

The fine-tuned BERT model achieved ~90% accuracy and 0.88 macro F1-score. It performed best on Active ↔ Passive and Positive ↔ Negative transformations, with minor challenges in Direct ↔ Indirect detection.

Error Analysis

Misclassifications occurred mainly in longer or syntactically complex sentences. Indirect speech was the most difficult due to variability in reported structures. Future improvement includes integrating syntactic parsing for better linguistic handling.

Conclusion

The transformer-based NLP model effectively classifies sentence transformation types and demonstrates explainability through SHAP and attention maps. This project validates the potential of interpretable AI in linguistic applications and can be extended for grammar correction or educational NLP tools.

Future Scope

Increase dataset size, experiment with RoBERTa/T5, and build a Streamlit-based demo for real-time transformation detection and visualization.