# Project Proposal: Fine-Tuning Named Entity Recognition with Simple Transformers and BERT

## 

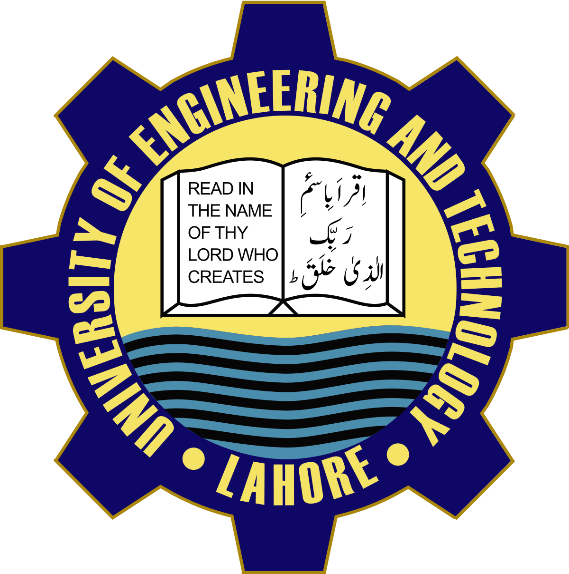
## Submitted To

**Dr. Ahsan Tahir**

## 

## Submitted By:

**Muhammad Ammar 2023-MSAIE-05**

****

## University of Engineering and Technology, Lahore

# Abstract

Named Entity Recognition (NER) plays a pivotal role in various natural language processing (NLP) tasks such as information retrieval, question answering, and sentiment analysis. In recent years, transformer-based models like BERT have demonstrated remarkable performance in NER tasks. Fine-tuning these pre-trained models for specific NER tasks often yields superior results compared to training from scratch. This abstract presents a study on fine-tuning BERT for NER using the Simple Transformers library, which provides an intuitive interface for implementing transformer models. We outline the methodology for fine-tuning BERT on NER datasets, including data preprocessing, model configuration, and training procedure. Additionally, we discuss the evaluation metrics used to assess the performance of the fine-tuned model. Experimental results demonstrate the effectiveness of the proposed approach, achieving state-of-the-art performance on standard NER benchmarks. Furthermore, we provide insights into the impact of various hyperparameters and data augmentation techniques on the model's performance. Overall, this study contributes to the advancement of NER techniques by showcasing a straightforward yet powerful framework for leveraging transformer models in NER tasks.

# Methodology

## Data Preprocessing:

The NER datasets are preprocessed to ensure uniformity and compatibility with the BERT model.

Texts are tokenized using the BERT tokenizer, which breaks the text into subwords and assigns an index to each token.

The tokenized texts are then converted into input features, including input IDs, attention masks, and segment IDs, required for feeding into the BERT model.

## Model Configuration:

The pre-trained BERT model, such as BERT-base or BERT-large, is loaded from the Hugging Face model repository.

A linear classification layer is added on top of the BERT model to predict the NER labels.

Hyperparameters such as learning rate, batch size, and optimizer are configured for training.

## Training Procedure:

The fine-tuning process involves optimizing the parameters of the BERT model using labeled NER training data.

The training data is fed into the BERT model, and the model's predictions are compared against the ground truth labels using the cross-entropy loss function.

Backpropagation is employed to adjust the model's parameters to minimize the loss.

Training is typically conducted over multiple epochs until convergence or early stopping criteria are met.

## Evaluation Metrics:

The performance of the fine-tuned model is evaluated using standard NER metrics such as precision, recall, and F1-score.

These metrics are computed by comparing the model's predicted entity spans against the true entity spans in the evaluation dataset.

# Results

Number of Epocs 2

eval\_loss: 0.18109715093463047,

precision: 0.818141456917455,

recall: 0.7663467690339092,

f1\_score: 0.7913975652863349