# **Project Title: Tense Classification Using Natural Language Processing**

### **1. Introduction**

In the realm of Natural Language Processing (NLP), one critical aspect is understanding and processing the tense of sentences. Accurate tense classification can significantly impact various applications such as text analysis, machine translation, and information retrieval. This project aims to develop a model that can classify sentences into three tense categories: past, present, and future.

### **2. Objective**

The primary objective of this project is to design and implement a model that accurately classifies text into one of the three tense categories:

* **Past Tense:** Sentences describing events that have already occurred.
* **Present Tense:** Sentences describing events happening currently or regularly.
* **Future Tense:** Sentences describing events that will occur in the future.

### **3. Problem Statement**

Understanding the tense of a sentence is crucial in various NLP tasks. The model needs to accurately determine whether a sentence is referring to past, present, or future events. This task poses challenges due to the complexity and subtleties of natural language, such as different ways tenses can be expressed and the context in which they appear.

### **4. Data Description**

#### **4.1 Dataset**

The dataset consists of sentences labeled with their respective tenses. Each entry in the dataset includes:

* **Sentence:** The text to be classified.
* **Tense:** The correct tense label (Past, Present, Future).

The dataset is split into training and testing sets to evaluate the model's performance.

#### **4.2 Data Preparation**

The dataset is processed and prepared for model training through the following steps:

1. **Label Encoding:** Convert textual labels (Past, Present, and Future) into numerical format for model training.
2. **Tokenization:** Use a BERT tokenizer to convert sentences into a format suitable for the model.

### **5. Methodology**

#### **5.1 Model Selection**

For this task, a pre-trained BERT (Bidirectional Encoder Representations from Transformers) model will be fine-tuned. BERT is a state-of-the-art transformer-based model known for its effectiveness in various NLP tasks.

#### **5.2 Model Architecture**

* **Pre-trained Model:** BERT-base-uncased.
* **Fine-Tuning:** The model will be fine-tuned on the tense classification task with a classification head added on top of BERT.

#### **5.3 Training Procedure**

1. **Data Tokenization:** Use the BERT tokenizer to convert sentences into tokenized input suitable for the model.
2. **Training:** Fine-tune the BERT model using the training dataset.
3. **Evaluation:** Assess the model's performance on the test dataset using metrics such as accuracy\_score and Matthews correlation coefficient.

### **6. Expected Outcomes**

The model is expected to:

* **Accurately classify sentences** into past, present, or future tense.

### **7. Challenges and Considerations**

* **Complexity of Language:** Sentences can have multiple tense indicators, and some may not fit neatly into one category.
* **Contextual Understanding:** The model must understand the context in which tenses are used to make accurate classifications.
* **Diverse Data:** The dataset should include diverse sentence structures and contexts to train a robust model.

### **8. Future Work**

* **Expand Dataset:** Include more diverse sentences and languages to improve model robustness.
* **Refine Model:** Explore additional model architectures or techniques to enhance classification accuracy.
* **Deployment:** Develop an application or API for real-world usage of the tense classification model.

### **9. Conclusion**

This project aims to leverage NLP techniques to classify sentences into past, present, or future tenses effectively. By addressing the challenges and employing advanced models like BERT, the project seeks to improve understanding and processing of tense in natural language.