Grammar & Spell Fixer

PES2UG23CS450
Priyanka Satish

Abstract

Grammar correction is an essential task in Natural Language Processing, especially for improving the quality of written communication. This project presents a simple Grammar & Spell Fixer using a **Masked Language Model (MLM)** to correct grammatical errors in sentences. The approach works by masking an incorrect word in a sentence and using a pre-trained **BERT-based fill-mask model** to predict the most grammatically appropriate replacement based on context. The model accurately fills in the masked word with the correct grammatical form, resulting in a corrected sentence. This method is particularly effective for single-word grammatical errors such as verb tense, verb form, and preposition usage. Although the system requires manual masking and is limited to simple corrections, it demonstrates the practical application of masked language models for grammar correction tasks.

Model Used

This project uses a **Masked Language Model (MLM)** based on **BERT (bert-base-uncased)** from the Hugging Face Transformers library. BERT is a bidirectional transformer model trained to understand the context of a word by considering both the left and right sides of a sentence. For this task, the **fill-mask pipeline** is used, which predicts the most grammatically appropriate word to replace a masked token in a sentence.

The model takes a sentence containing a `[MASK]` token as input and generates multiple candidate words along with their confidence scores. The word with the highest probability is selected and inserted back into the sentence, resulting in a grammatically corrected output.

Example

Masked Sentence:

I am [MASK] to school

Model Output:

I am going to school