## Project Overview

Objective: The goal of this project is to develop a machine learning-based system that provides grammar feedback to users in real time. By analyzing user-inputted sentences, the system will identify grammatical errors and suggest corrections. This feedback aims to help users improve their language skills, particularly focusing on grammar usage and sentence structure.

## Motivation

Grammar proficiency is an essential component of effective communication, particularly in written English. Many language learners struggle with applying correct grammar rules, and consistent feedback can be invaluable in reinforcing correct usage. This project seeks to create a supportive and efficient grammar feedback tool that provides targeted feedback to help users identify and correct their mistakes.

## Approach

The project is structured in the following phases:  
  
Data Collection: We compiled a dataset of English sentences that includes a range of common grammatical errors alongside their corrected versions. Each sentence pair is annotated with:  
- Error Type: Category of the grammatical error (e.g., verb form, subject-verb agreement, preposition).  
- Incorrect Sentence: The sentence containing the grammatical mistake.  
- Correct Sentence: The corrected version of the sentence.  
  
Data Preprocessing: The dataset was cleaned to remove any duplicates and missing values. We used tokenization to break down each sentence into individual words and identified the positions of errors within the sentence. Each error was categorized based on predefined rules to determine its severity, which allowed us to prioritize the feedback.

## Model Training

Model Selection: We experimented with language models like BERT and T5, as they are well-suited for text-based tasks such as error detection and correction.  
Training Process: The model was trained to detect and classify grammatical errors by learning patterns from labeled data. It was optimized to ensure high accuracy in predicting error types and suggesting the correct replacements.  
Evaluation: The model was evaluated on accuracy in error detection and correction, using a test dataset with unseen grammatical mistakes.

## User Interaction and Feedback Mechanism

Input: Users submit sentences for grammar evaluation.  
Processing: The model identifies grammatical errors, categorizes them by type and severity, and generates suggested corrections.  
Output: The system presents a detailed feedback report that includes the following:  
- Highlighted errors in the sentence.  
- Suggested corrections for each error.  
- A description of each error type to help users understand the rules better.

## Key Components and Features

Error Detection and Classification: The system identifies errors based on the trained model and classifies them into types, such as incorrect verb forms, missing prepositions, or subject-verb agreement issues.  
  
Feedback Generation: For each detected error, the system provides:  
- Explanation: A brief description of the error type.  
- Suggested Correction: Recommended changes to improve the sentence.  
- Error Positioning: Highlights the exact location of the error within the sentence for clarity.

## Exploratory Data Analysis (EDA)

To better understand our dataset, we conducted EDA, focusing on:  
  
Deleted and Added Words: Analyzed common words that were often deleted or added during correction, providing insights into typical grammar mistakes.  
  
Sentence Length Analysis: Compared the lengths of incorrect and correct sentences to examine if corrections affected sentence length significantly.  
  
Correlation Analysis: Assessed relationships between sentence structure and error types using a correlation matrix, which revealed patterns in how sentence length and complexity related to error frequency.

## Conclusion

The "Language Learning for Grammar Feedback" project demonstrates the potential of machine learning to offer real-time grammar assistance tailored to the needs of language learners. By leveraging a model trained on a variety of grammatical errors. Through targeted feedback, including explanations and suggested corrections, this system provides a practical approach to improving grammar skills.

Our exploration of the dataset through EDA provided valuable insights into the types of errors users are likely to make, common patterns in word usage, and the impact of corrections on sentence length. This project sets a foundation for a robust grammar feedback system, empowering language learners with instant, meaningful feedback to guide their learning journey.