LLM_Assigment1_Project

I.   Introduction

Sustainable agriculture is essential for achieving food security and addressing the challenges posed by climate change, population growth, and resource scarcity. Aligned with Sustainable Development Goal (SDG) 2: Zero Hunger, this project aims to promote sustainable agricultural practices that enhance productivity, ensure environmental sustainability, and improve livelihoods.

This project focuses on developing a sentiment analysis tool to gauge public opinion on sustainable agriculture. By leveraging the **"distilbert-base-uncased-finetuned-sst-2-english"** model, this tool analyzes sentiments expressed in various textual data related to sustainable agriculture. The insights gained from this analysis can help stakeholders understand public perceptions, identify concerns, and tailor communication strategies to promote sustainable agricultural practices.

II.   Sustainable Agriculture and SDG 2

Sustainable agriculture is a key component of **SDG 2: Zero Hunger,** which aims to end hunger, achieve food security, improve nutrition, and promote sustainable agriculture. Various studies have highlighted the importance of public awareness and support for sustainable agricultural practices, as they play a crucial role in ensuring the long-term viability of food systems.

III.   Model Selection

- Research and Setup: The **"distilbert-base-uncased-finetuned-sst-2-english"** model was selected for this project due to its strong performance in sentiment analysis tasks.

IV. Pre-trained Models for Sentiment Analysis

Pre-trained models like **DistilBERT** have revolutionized the field of Natural Language Processing (NLP) by offering efficient and accurate sentiment analysis capabilities. The **"distilbert-base-uncased-finetuned-sst-2-english"** model is specifically fine-tuned for sentiment analysis tasks and has been widely adopted due to its ability to handle nuanced language and context.

V. Data Collection

Dataset Creation: Unlike many projects that rely on publicly available datasets, the dataset used in this project was created by the me. The dataset consists of textual data related to agriculture, collected through surveys.

VI. Model Implementation

Sentiment Analysis Implementation: The sentiment analysis tool was implemented by feeding the preprocessed data into the loaded model. Each text entry was classified into three categories: Clay, sandy and salty.

VII. Loading the Model:

The "distilbert-base-uncased-finetuned-sst-2-english" model was imported from the Hugging Face model hub using the `'transformers'` library. The model was chosen for its capability to perform sentiment analysis efficiently on a wide range of textual data, making it ideal for analyzing diverse opinions on sustainable agriculture. The associated tokenizer was employed to preprocess the dataset, converting the text into a format suitable for analysis.

VIII.   Fine-Tuning

- Dataset Preparation for Fine-Tuning: Although the `"distilbert-base-uncased-finetuned-sst-2-english"` model is already fine-tuned for general sentiment analysis, additional fine-tuning was performed using the custom dataset created by the me.
- **Fine-Tuning Process:** The fine-tuning process involved running the model on the custom dataset for multiple epochs, with careful adjustments to the learning rate to optimize performance.

Reference

- Models used in the project are from Hugging Face