Table of Contents

**Objective ................................................................................................................ 1**

**Steps to be Followed................................................................................................ 1**

**Dataset Used............................................................................................................2**

**Chatbot Stages in Action ......................................................................................…..2**

**Screenshots …………………………………………………………………………………………………….3**

**Objective of the Project** – We are supposed to build an AI System using Large Language Model (LLM). We have to build an AI based System that can answer the query from the documents which can be in any format. Also, we are to use Openai API Key to generate response from the system.

**Steps To Be Followed** – The Project should implement all the 3 layers of LLM

1. The Embedding Layer – The pdf document needs to be processed, cleaned and in proper format and chunked for the embeddings. We can choose any of the modelling strategies, either Openai or any other model
2. The Search Layer – we need to design at least 3 queries against which we will test our system. We need to embed the queries and search Chromadb vector database against each of the queries. Implementing a cache memory is also mandatory. Also we need to implement re- ranking block
3. The Generation Layer – the final prompt we design will be the major component. Prompt should be exhaustive in its instructions and relevant information is correctly passed to the prompt. Few – shot prompting can also be provided to the prompt to get better expected responses.

**Dataset Used** – I have taken the Crop Recommendation Data from the Kaggle platform which have 8 features related to crops, their growing climatic conditions which includes area, temperature, wind, soil type, etc.

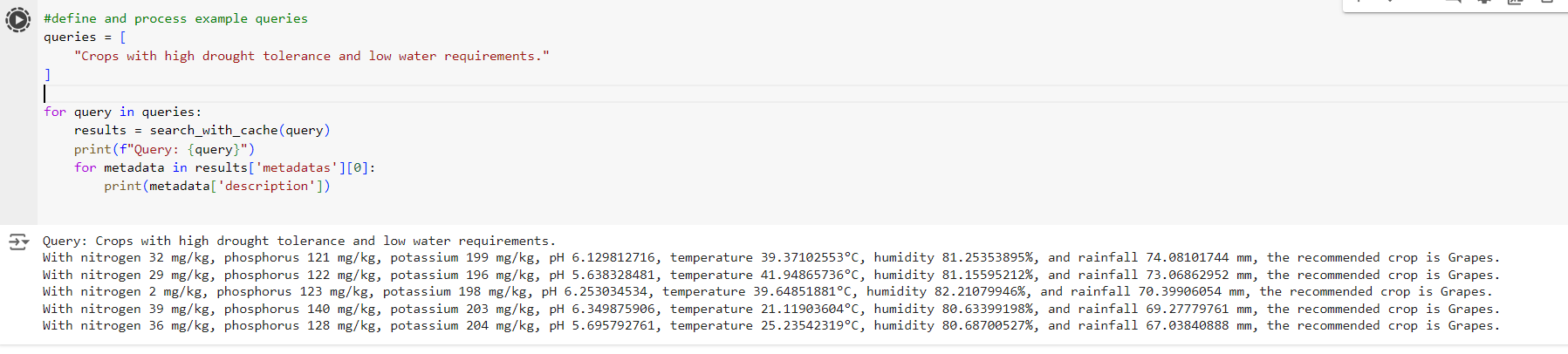
Dataset was available in csv format so I changed its format from csv to pdf, then chunked it for embeddings.

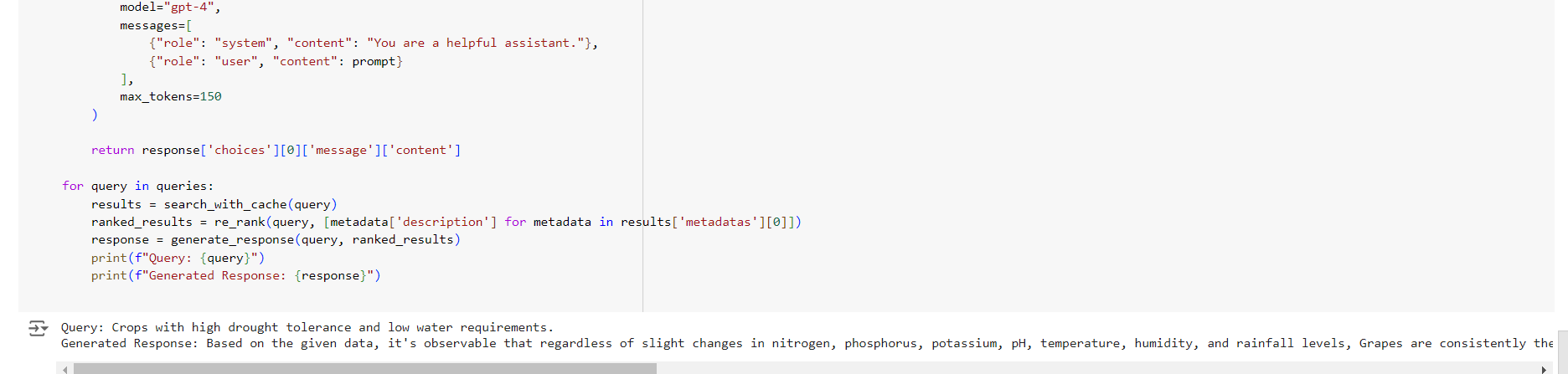
**Chatbot Stages in Action** – following stages were performed under the project

1. Installed all the necessary libraries with updated and supported versions
2. Fetched data and read it raw
3. Converted data into required format
4. Generated embeddings
5. Chromadb setup and added unique ids
6. Added embeddings and description to chromadb
7. Initialized cache for searching system
8. Processed some example queries
9. Initialized re- ranking function
10. Generated code for the response and got the final answer

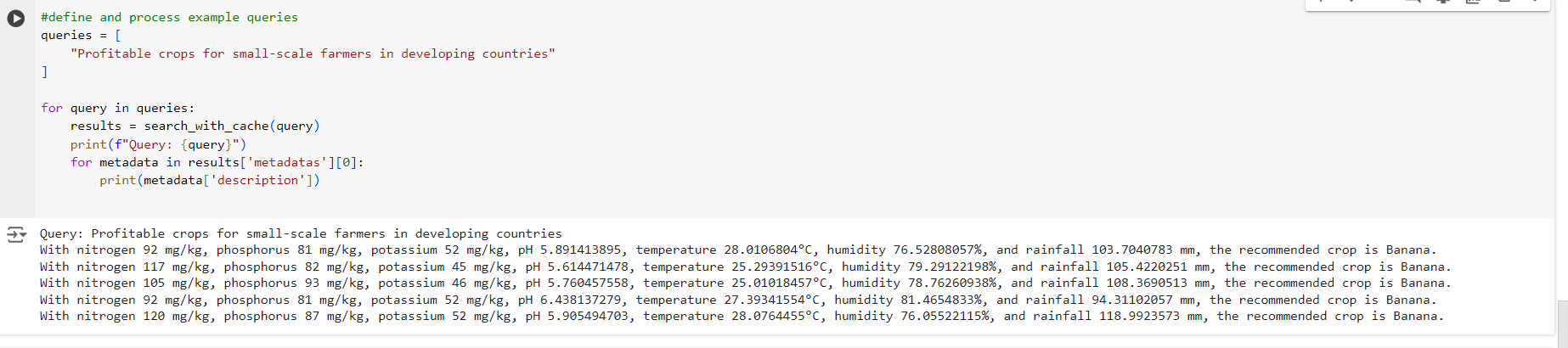
**6 Screenshots for queries**

1. Query 1 – Crops with high drought tolerance and low water requirements
2. Response Generated - Grapes





1. Query 2 - Profitable crops for small scale farmers in developing countries
2. Response Generated - Banana





1. Query 3 – High demand crops in local market with premium prices
2. Response Generated - Maize

