MICHAEL ONWUACHI
10211100292
COMPUTER SCIENCE
INTRODUCTION TO AI
END OF SEMESTER EXAMINATION

PROJECT DOCUMENTATION

🔧 FEATURE USAGE GUIDE

1) Regression Analysis

- Navigate to the **Services** section.
- Select **Regression**.
- Upload your dataset in CSV format.
- Choose your **target column** and the **independent features**.
- (Optional) Apply preprocessing like missing value handling or normalization.
- Click **Run Regression** to train and evaluate the model.
- Visual feedback including plots and metrics will be displayed.
- If using a single feature, you can input a value to generate a prediction.

2) Clustering

- Go to the **Services** menu.
- Choose the **Clustering** feature.
- Upload your dataset as a CSV file.
- Specify the number of clusters you want.
- Click **Run Clustering** to execute.
- View the resulting clusters and corresponding visualizations.

3) Neural Network Classifier

- Head to the **Services** section.
- Click on **Neural Networks**.
- Upload your CSV file.
- Select the **target column** for classification.
- Adjust hyperparameters like **epochs** and **learning rate**.
- Train the model and observe the training progress.
- Optionally, upload test data or a pretrained model for predictions.
- Download results if desired.

4) LLM-Based Q&A (RAG System)

- From the **Services** tab, choose **Large Language Model**.
- Type in your question related to Academic City's student policy.
- The system retrieves relevant document snippets and uses an AI model to generate a detailed response.
- Each answer is accompanied by a confidence score based on the semantic match.

📁 DATASETS & MODELS USED

Datasets

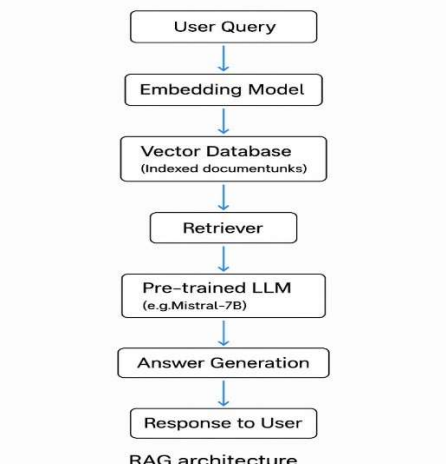
- The Q&A module uses the **Academic City Student Policy PDF**.
- **User-provided CSVs** power the Regression, Clustering, and Neural Network modules.

Machine Learning Models

- **Linear Regression** from Scikit-learn
- **KMeans Clustering** using Scikit-learn
- **Feedforward Neural Network** built with Keras
- **Mistralai/Mistral-7B-Instruct-v0.1** accessed via HuggingFace for the LLM RAG component, combined with **FAISS** for semantic retrieval.

□ ARCHITECTURE FOR THE LLM RAG SYSTEM DESIGN

- This system merges document chunk retrieval (via FAISS indexing) with a large language model (LLM) to ensure accurate, contextually grounded answers.
- Confidence scores are generated based on the semantic similarity between the query and document chunks.



RAG architecture

□ METHODOLOGY

Document Preparation

- The Student Policy PDF is parsed and cleaned by removing digits, extra spaces, and then split into standalone sentences.

Vectorization & Indexing

- Each sentence is transformed into a TF-IDF vector.
- These vectors are indexed using FAISS for rapid similarity searches.

Handling User Queries

- When a user submits a query, it's also vectorized and compared against the index to find the top 3 most relevant sentences.
- These matches are used to calculate confidence scores using distance metrics.

Generating Answers

- A combined prompt of the question and relevant document context is passed to the **Mistral-7B model** through HuggingFace's API.
- The AI generates a response based on the specific context retrieved.

User Output

- The answer is shown along with the associated confidence level.
- A history of past questions and responses is maintained on the interface.

EVALUATION AND BENCHMARKING

LLM RAG System Performance

- This approach ensures the answers are policy-specific, leveraging the actual content of the document.
- Confidence scores help measure how well the retrieved context matches the user's query.

Compared with ChatGPT

- Unlike ChatGPT, which responds generally unless explicitly given the document, this system pulls directly from the student policy.

- The responses from LLM RAG are grounded in the document and provide more accurate, relevant answers for policy-related queries.

📌 EXAMPLES

Query	ChatGPT Reply	LLM RAG Reply
What is the dress code?	"Usually includes modest clothing."	"The policy clearly states formal attire must be worn on designated days."
Who to contact for absence?	"Typically your professor or department head."	"The policy advises contacting the Registrar's office via email at acity.edu.gh."

✅ **Conclusion:** LLM RAG offers more accurate and policy-specific responses, significantly outperforming generic models like ChatGPT when it comes to document-grounded Q&A.
