SUBHAJYOTI MAITY

Tech Consultant

@ subhajyotimaity123@gmail.com in https://www.linkedin.com/in/subhajyoti-maity-84577a25b/ https://github.com/subhajyotirkmveri/ (+91) 7319453140

BACKGROUND

- Tech consultant
  01/09/2024 - present — QJ Consulting Pvt Ltd.
- Tech consultant Intership
  01/06/2024 - 31/08/2024 — QJ Consulting Pvt Ltd.

EDUCATION

MSc in Big Data Analytics
Ramakrishna Mission Vivekananda Educational and Research Institute
Aug 2022 – July 2024

B.Sc - Mathematics Honours
Under Vidyasagar University
Jul 2019 – July 2022

EXPERTISE

Generative Artificial Intelligence | EDA | Feature Engineering | Machine Learning Modelling | NLP | Documentation | Web Scrapping | Time Series Modeling | Power Automation | Real time data updation in Power Bi | Semantic Model

TECHNICAL SKILLS

Python | R | SQL | Power Bi | Power Automation | Real time data Updation | Pyspark | pytorch | tensorflow | Scikit-Learn

CERTIFICATION

- NPTEL Data Mining
- NPTEL DataBase Management System
- SQL certification from HackerRank

PROFESSIONAL EXPERIENCE

Devlope an application that extract data from various invoice formats

LLM | Gemini Pro

06/2024- — QJ Consulting Pvt Ltd.

- The application will enable users to upload multiple files whether in image or PDF format.
- Sends prompt text that include a comprehensive prompt for the Google Gemini model to ensure accurate extraction of invoice details such as Invoice No, Supplier Name, Bill To Details, Product Name, Quantity, Unit, Per Unit Price, and Amount and image data to the Google Gemini model and retrieves responses.
- Identify and extract relevant substrings from the model's response.
- Convert the structured text into a pandas DataFrame, ensuring correct parsing and formatting of invoice data and combine the data from all processed files into a single DataFrame for easy review and download

Develope an application that translates natural language question into SQL queries

LLM | Gemini Pro

06/2024- — QJ Consulting Pvt Ltd.

- Created a sqlite database from three Excel sheets related to purchase order details, integrating them into a cohesive data structure.
- A comprehensive prompt template is used to guide the AI model in generating accurate SQL queries based on the input questions
- Executes the generated SQL query on the that database and displays the results in the Streamlit app
- Appends each interaction (user question, generated SQL query, and response) to a CSV file for logging and future reference

ACADEMIC PROJECTS

Developed a question answering app specifically for PDF documents

LLM | LLama-2 | Mistral 7B

01/2024- 05/2024

- Create an object for generating vector embeddings of text chunks using pre-train sentence tranformer model like all-MiniLM-L6-v2.
- I utilize the FAISS (Facebook AI Similarity Search) library to efficiently search for embeddings of multimedia documents that exhibit similarity to one another.
- Perform Nearest Neighbourhood Search to retrieve top k relevant chunk of the user input.
- LLM model are used for generating responses in a conversational setting.
- Develop a Streamlit web application that return the answer to a question relevant to the PDF documents and include the source document page.

Object detection using the YOLOv8 model on custom datasets

Computer Vision | YOLOv8

01/2024- 05/2024

- Leveraged data augmentation techniques such as rotation, flipping, and scaling to increase the diversity of training data and improve model generalization.
- Utilized YOLOv8 (You Only Look Once version 8) object detection model for accurate and real-time detection of butterflies and squirrels in images.
- The evaluation metrics used to assess the model's performance such as mAP (mean Average Precision), IoU (Intersection over Union).
- Develop a Streamlit web application for detecting images containing butterflies and squirrels .

Spam or Non-Spam Message Detection

Natural Language Processing

04/2023-07/2023

- Conducted initial preprocessing of the text corpus, which included both spam and non-spam messages.
- Employed the **Bag of Words (BoW) approach to transform and structure the text data for further analysis and also use Tf-Idf Model for enhanced feature extraction.**

Home Loan Approval Prediction

Machine Learning

01/2023-04/2023

- Explored and implemented diverse Machine Learning classification models (including KNN and Naive Bayes,Logistic Regression,Decision Tree) for predicting Home Loan Approval.
- Among those models, the Decision Tree model demonstrated superior accuracy in predicting Home Loan Approval status.

ACHIEVEMENT

Inspire Scholarship (SHE)