Siddharth Shukla

+91-7521018246 shuklasiddharth418@gmail.com linkedin github portfolio

Education

Graphic Era Hill University 2022 - 2026
Bachelor of Technology in Computer Science (GPA: 8.71 / 10) *Dehradun, Uttrakhand*

Prabhat Senior Secondary Public School 2020 - 2021
Class 12th - Percentage: 95.2% *Kanpur, Uttar Pradesh*

Prabhat Senior Secondary Public School 2018 - 2019
Class 10th - Percentage: 94.8% *Kanpur, Uttar Pradesh*

Projects

Airline-Passenger-Refferal-Prediction | *Machine Learning, Scikit-learn, Classification, Gradio*

- Built and deployed a classification model to predict whether a passenger will refer the airline to others, addressing a critical business need to improve customer satisfaction and retention.
- Validated model robustness through cross-validation, confirming stable performance with only a 0.02% difference in misclassifications compared to a single train-test split.
- Generated a model accuracy of **94%** using **Logistic Regression** and a precision score of **94%** for the positive class.

EduChat | *Python, Flask, LangChain, Prompt Engineering*

- Engineered a full-stack **Retrieval-Augmented Generation (RAG)** system using **LangChain** and a **FAISS** vector store to enable accurate, context-aware chat over uploaded PDF documents.
- Developed a AI Essay Grading module using the **Gemini 2.5 Flash Chat Model**, leveraging dynamic system prompting to evaluate essays against user-defined rubrics
- Demonstrated robust Software Implementation by building a production-ready backend, handling API Key security and successfully debugging complex cloud errors (400/404 Model Access, 429 Quota Exceeded).

Advanced-Zomato-Restaurant-Clustering-and-Sentiment-Analysis | *Clustering, NLP, Scikit-learn, Sentiment Analysis*

- Developed a **clustering** model for Zomato restaurants based on reviews and features, allowing the identification of different restaurant segments and assisting in strategic marketing decisions.
- Faced challenges in **text preprocessin**g for sentiment analysis, conducted feature engineering for clustering and reduced dimensionality by **90%** using **PCA** for better visualization.
- Achieved an **ROC AUC** score of **0.921** with an **XGBoost classifier**, and obtained **86%** accuracy for sentiment classification.

ChatBot-Using-Deep-Learning | *Keras, TensorFlow, Sequence Modeling , LSTM*

- Built and fine-tuned an LSTM-based chatbot to handle natural language queries, enhancing user interaction and automating customer support responses.
- Implemented the model with TensorFlow and Keras, utilizing pre-trained embeddings to improve contextual understanding of responses, leading to a 25% improvement in response relevance.
- Achieved a validation accuracy of 90% demonstrating a well trained language model for conversational AI.

Skills

Programming Languages: Java, C, C++, Python, SQL
Artificial Intelligence: Scikit-learn, NLP, Recommender Systems, Hypothesis Testing, TensorFlow, Keras, Neural Networks, LangChain, RAG, GenAI.
Tools: Git, VSCode, Jupyter Notebook, Github, Hugging Faces.
Concepts: Operating System, DBMS, Data Structure and Algorithm, Computer Networks, OOPs, REST APIs.
Soft-Skills : Communication, Analytical thinking, Teamwork, Adaptability, Time management

Achievements

- 700+ problems solved on LeetCode, with a focus on data structures and algorithms: My LeetCode Profile
- Engineered a predictive pricing model using a Decision Tree Regressor for the Kaggle '**Backpack Prediction Challenge**', achieving a Mean Absolute Error of 33.66 and ranking in the top 61% (2066/3,383) of global competitors.