SIDHARTHA MOHANTY

Contact: 9776296297 ⋄ Email: sidhartha.mohanty93@gmail.com ⋄ LinkedIn ⋄ Github ⋄ Medium

PROFILE SUMMARY

Machine Learning Engineer with 5 years of experience in building and deploying ML models. Skilled in Python, SQL, and end to end ML workflows, specializing in predictive analytics, NLP and AI driven solutions.

EXPERIENCE

Unit Manager, Bajaj Finserv, Pune, India   Dec-2021 - Present

- Developed and trained ML models for classification from scratch.
- Collaborated with data engineers to build data pipelines and preprocess large datasets for model training.
- Used Kubeflow and Mlflow to build machine learning pipelines to track versions, logs and manage models.
- Implemented NLP models from scratch for text classification.
- Integrated ML models into production environments using Docker and Github.
- Shared findings and insights through visualizations and presentations.
- Integrated CI/CD workflows with Azure ML.
- Created proof of concepts for new technologies in Generative AI.

Advisor, Concentrix, Bengaluru, India   Feb-2020 - Dec-2021

- Developed and deployed end-to-end ML models for predictive analytics, including regression and classification,
- Performed EDA using Matplotlib, Seaborn, NumPy, Pandas.
- Optimized performance using feature engineering, hyperparameter tuning, and model evaluation techniques.
- Verified data accuracy and completeness through data quality checks.
- Conducted model performance monitoring and optimization.
- Presented findings through data visualizations and reports.

PROJECTS

Customer support Chatbot: Created a RAG chatbot using company's policies documents and frequently asked questions to answer customer queries. Used Langchain, OpenAI GPT-4o mini and Azure vector index to implement the project.

Intent based AI agent: Developed an agentic workflow using few-shot prompting to accurately identify user intent and route queries to specialized agents for resolution. Integrated LLAMA and Mixtral models via Groq for high-speed inference, and utilized GPT-4o-mini from OpenAI for lightweight reasoning. Built the backend with FastAPI and designed the frontend using HTML and Javascript. Github

PDF RAG Chatbot: Created a chatbot to answer questions based on a provided PDF content using Langchain, GPT-4o mini, FAISS vector store. Github

OCR based document processing: Leveraged Pixtral-12B for high-accuracy text extraction from images. Applied few-shot prompt engineering to convert raw text into structured JSON, enabling seamless transformation into business-ready tables.

Lead to disbursement prediction: Built a model to predict whether a lead will get converted to a personal loan customer by considering loan history and demographics. Used PCA for dimensionality reduction and Random forest for training the model.

Chat with your video bot: Built a RAG app to answer questions taking YouTube video content as knowledge base. Used OpenAI's Whisper, GPT-3.5, Pinecone to achieve this. Medium story

SKILLS

Programming Languages	Python, SQL, Pyspark
Technical Skills	Predictive modeling, EDA, ANN, NLP, Hyperparameters tuning, Deployment, Generative AI
Tools	Databricks, SQL Server, LLMs, Langchain, Docker, TensorFlow, Keras, PyTorch, Kubeflow, Mlflow, Scikit-learn, Pinecone, ChromaDB, Vector Databases, Graph Databases, Gemini, FastAPI, LLAMA
Working Environments	Jupyter Notebook, Google Colab, Spyder, VSCode, Databricks, CursorAI

CERTIFICATIONS

Knowledge Graphs for RAG Link

Generative AI for Everyone Link

Prompt Engineering for ChatGPT Link

GPT Vision: Seeing the World through Generative AI Link

Querying Microsoft SQL Server with Transact-SQL Link

EDUCATION

Bachelor Of Technology, National Institute of Science and Technology