Vinit Shah

Oracle Certified Gen AI & Data Science Professional
vinitshahofficial@gmail.com | +91 6353845489 | LinkedIn | GitHub | Kaggle

EDUCATION

Gujarat Technological University, SVIT, Vasad   2018 – 2022
Bachelors in Engineering in Information Technology   *CGPA: 8.33*

EXPERIENCE

AI/ML Developer   11/2024 – current
Aerocraft Engineering

– Developed an Intelligent Document Processing (IDP) system for detecting and classifying document components (e.g., text blocks, tables, images) with high accuracy, and deploying it via Docker.
– Implemented a robust data end to end pipeline for Object Detection with YOLOv11 compatible with current pipeline that increase 25% detection with 97% recall in FastAPI.

Research Engineer, NLP   09/2021 – 11/2024
AGS Health

– Collected and analyzed medical data, cleaned and pre-processed it for pattern recognition using MySQL, Python, and statistical methods to derive actionable insights and solutions.
– Led development of automation scripts, profiling tools, and database integration, improving system efficiency by 30% using MLflow.
– Designed and implemented medical code automation projects using machine learning and deep learning approaches, achieving 79% F1 score with integrated XGBoost and Logistic Regression models.
– Deployed efficient data cleaning pipelines on AWS SageMaker, ensuring data quality and integrity.
– Successfully conducted Proof of Concepts (POCs) for clients, achieving 90% recall in data analysis.
– Resolved memory challenges with binary serialization, achieving 97.25% file compression and saving over $25,000 annually.
– Utilized Agentic Retrieval-Augmented Generation (RAG) using LangChain and ChromaDB for information retrieval tasks.
– Fine-tuned LLMs (llama3, bioclinicalBERT, BioMistral) for information retrieval and classification, enhancing F1 scores in multiple projects.
– Developed solutions for extracting data from table structures in PDFs, scanned images, and digital formats.

AWARDS & ACHIEVEMENTS

- **Top position in Intellect AI Hackathon 2023:** Built a custom NER model using Faster-RCNN for invoice image entity detection.
- **5 Star Gold Badge in SQL:** Hackerrank achievement for advanced SQL proficiency.
- **5th Rank in Kaggle ML in Physics competition:** Implemented sklearn-based models to achieve high accuracy.
- **Star Performer Award at AGS Health:** Recognized in 2022 for continuous high performance.

CERTIFICATES

- Oracle Certified Generative AI Professional
- Oracle Certified Data Science Professional

TECHNICAL SKILLS

- **Libraries/Frameworks:** Pandas, NumPy, TensorFlow, PyTorch, LangChain, MLflow, OpenCV, NLTK, Hugging Face, Flask, FastAPI.
- **Tools:** Docker, AWS S3, SageMaker, MySQL, Git, VSCode.
- **Languages:** Python, Java, SQL, JavaScript.