Mahikshith

+91-6303423352 | [Linkedin](#) | [Github](#) | [LLM & AI](#) | [Scaler](#) | [Tableau](#) | mahikshith97@gmail.com

SKILLS

- Python, SQL, NumPy, Pandas, Matplotlib, Seaborn, Scikit-learn, TensorFlow, Keras, Data Analysis, Stats and Probability, Tableau, Hypothesis testing, Machine Learning , NLP, Time series forecasting, Flask, Docker, AWS, EC2 , ECS , SageMaker, LLM , RAG, MLflow, Langchain, Web scraping, ETL, CI / CD pipeline

EXPERIENCE

Data Science Associate   Mar 2021 - Jul 2022
Amazon

- Reduced Order Defect Rate (**ODR**) by **0.27%** and brand infringement claims by 9.46% monthly, enhancing operational efficiency through **Amazon QuickSight** and **SQL** analytics.
- Developed and implemented **ETL pipelines** from scratch using **Python**, increasing data accuracy and timeliness by **7%** through automated ingestion and transformation processes.
- Streamlined operational efficiency with **CI/CD pipelines**, resulting in faster deployment cycles and improved data handling.
- Accelerated model development for large datasets with **Amazon S3** and **AWS SageMaker**, achieving a **3%** reduction in model response time in production.

PROJECTS

AWS-Deployed Predictive Loan Risk classification Model
https://github.com/mahikshith/End-to-End-implementation-of-Loan-tap-classification

- Conducted exploratory data analysis using **Python** to reveal key insights and trends that informed strategic business decisions.
- Executed feature engineering to enhance model performance, boosting predictive **accuracy** by **12.5%** for creditworthiness assessments using a **Logistic regression model**.
- Delivered actionable insights and recommendations from **Data analysis** and Feature engineering , driving improved business strategies and outcomes.
- Developed an interactive demo using **Streamlit** to effectively showcase model capabilities to stakeholders.
- Performed API testing in **Postman** with **Flask API** to ensure robust functionality and reliability for end-users.
- Created and published a **Docker** image and Established an **AWS EC2** instance to host the application, ensuring reliable access and performance for users.

Fine-tuned Transformer-based Text Summarization with NLP & ETL Pipeline implentation
https://github.com/mahikshith/Transformer-Text-Summarizer-Fine-tuning-with-ETL-pipeline-and-Deployment

- Engineered and deployed an **end-to-end NLP text summarization** system by **fine-tuning** Google's Pegasus **transformer model** (568M parameters) on 16k conversations, boosting **12 % increase in ROUGE** scores for enhanced dialogue summarization accuracy.
- Architected a robust MLOps **ETL pipeline** that automated data preprocessing and model training, reducing manual intervention by 70% and reducing iteration time from days to hours
- Implemented memory-efficient training strategies including **gradient checkpointing** and mixed-precision (**FP16**), reducing GPU memory fragmentation by 20% with **PyTorch** for stable training on conversations up to **1024 tokens**.
- Designed a comprehensive monitoring and **logging system** that improved issue detection speed by 65%, reducing debug time from hours to minutes in production.
- **Containerized** and deployed a **FastAPI**-based endpoint service using **Docker**, ensuring consistent performance and optimal latency across multiple environments.

LLM powered E-commerce review Chatbot: AWS Deployment and API testing
https://github.com/mahikshith/E-commerce-review-chatbot-end_to_end-AWS-deployment

- **Web scraped** product information, reviews, and ratings from an e-commerce site using **Selenium**, gathering over 10,000 data points for analysis.
- Executed exploratory data analysis and feature engineering, improving **embedding accuracy by 23.4%** for the **Gemini 1.5 Pro LLM**.
- Created and stored embeddings in **Astra DB**, enhancing data accessibility and boosting **retrieval speed by 15% through cosine similarity**,
- Developed a **retriever** and constructed a **Langchain** workflow, optimizing data processing and reducing retrieval time by 30%.
- Showcased application functionality using **Streamlit**, ensuring robust performance through local testing with **Flask** API and **Postman**.
- **Deployed** the review **chatbot** application on **AWS EC2**, enhancing security and monitoring performance through comprehensive public **API testing**.

EDUCATION

Scaler woolf neovarsity   2022 - 2024
MS in Computer Science | 3.9 GPA
Specalization in Artificial Intelligence and Machine learning