Sai Dhiren Musaloji

musalojidhiren@gmail.com | +1(862)-423-8830 | LinkedIn | Git-Hub | Portfolio

Education

New Jersey Institute of Technology - MS in Data Science GPA:3.85/4

Mahatma Gandhi Institute of Technology – B. Tech in ECE GPA: 7/10

Certification

Microsoft Azure Data Science Associate, Microsoft Azure AI Engineer Associate 2022

Technologies

Python, R, SQL, Java, Bash|· scikit-learn, TensorFlow, Keras, PyTorch, XGBoost, MLlib |· Hugging Face, BERT, LSTM, GNN, ARIMA| Pandas, NumPy, Statsmodels, Spark, Hadoop, MapReduce, AWS EMR |· Tableau, Power BI, Jupyter, Matplotlib, Seaborn, Plotly | AWS (EC2, S3, SQS, Rekognition), GCP, Azure |· Docker, Kubernetes, Hadoop, Spark, Oozie, GitHub Actions, CI/CD, MLflow | ·MySQL, PostgreSQL, SQL Server | · ETL pipelines, modular architecture, scalable cloud-native systems |

Experience

AI Engineer Intern - Tech Mahindra, Makers Lab Pune, Maharashtra (Oct 2023 – Jan 2024)

Large Language Model Development

- Collaborated with 5+ AI researchers and engineers to architect NLP pipelines for low-resource languages, implementing SentencePiece and vocabulary strategies to process code-mixed input
- Engineered hybrid regex classifiers to filter toxic content from 1TB+ conversational corpus, eliminating 20% of unwanted data
- Delivered weekly sprint presentations to 4 senior architects and documented internal workflows for reproducibility and deployment

Projects

Traffic Forecasting Project – Multimodal Ensemble Spatiotemporal Modeling - *(ARIMA, LSTM, GNN, t-SNE, PCA, PEMS-BAY dataset)*

- Constructed a composite forecasting system combining **ARIMA, LSTM, and GNN** to capture temporal and topological traffic patterns.
- Applied **t-SNE** and **PCA** to reduce the dimensionality of a **325-sensor dataset**, improving model convergence in **25 fewer** epochs.

Deep Learning - Advanced Neural Network Architectures - *(PyTorch, GANs, Diffusion Models, BERT, Reinforcement Learning)*

- Built and fine-tuned **DNN and GNN models** for classification and forecasting, increasing model accuracy through parameter tuning and cross-validation**.**
- Created image synthesis systems using **GANs and diffusion models** with distributed TensorFlow for faster training.
- Applied BERT fine-tuning and **reinforcement learning (Policy Gradient, Actor-Critic)** to address NLP and decision-making tasks.

Database Management Web Application - *(Flask, MySQL, HTML/CSS/JS, CI/CD)*

- Built a scalable Flask-MySQL web application for **core banking operations** with **real-time transaction handling**.
- Integrated analytics dashboard with **fund flow heatmaps** and **anomaly detection**, containerized for **horizontal scaling** and CI/CD deployment.

AWS SQS Image & Text Recognition Pipeline – Event-Driven Distributed System - *(AWS EC2, SQS, S3, Rekognition, Python, Boto3)*

- Architected an event-driven **processing pipeline** using **AWS EC2, SQS, S3, and Rekognition** for asynchronous image and text recognition**.**
- Configured **stateless compute nodes** for **parallel object detection** and **OCR**, enabling elastic scaling and high throughput.
- Ensured system resilience with retry logic, visibility timeout tuning, and idempotent handling, achieving 99.9% uptime.

Wine Quality Prediction – Parallel Machine Learning on AWS - *(AWS EC2, SPARK, EMR, DOCKER, Python, Wine Dataset)*

- Designed and deployed a **parallelized ML training pipeline** using **Apache Spark** on a 5-node AWS EMR cluster**,** reducing model training time by 60% through streamlined data partitioning, in-memory caching, and executor tuning.
- Containerized the **trained model** using **Docker** with reproducible dependency management and deployed it on a compute-optimized **EC2 instance** with custom startup scripts for **automated bootstrapping** and inference serving.
- Engineered the **pipeline** for high availability and performance, incorporating **fault-tolerant EMR configurations**, **autoscaling policies, and Spark resource monitoring** to ensure 99.9% operational uptime and 92% cluster resource utilization.