Mukesh Reddy Mavurapu

mukeshreddymavurapu@gmail.com | LinkedIn | GitHub | 920-570-8976 | Houston, TX

SUMMARY

- **4.5+ years of experience** in AI/ML, data engineering, and software development, specializing in fine-tuning and deploying advanced machine learning models, and integrating generative AI solutions using RAG methodologies and LangChain for semantic search optimization.
- Proficient in **cloud platforms** like AWS, Azure, and Databricks, with hands-on expertise in deploying scalable applications and machine learning solutions using **TensorFlow**, **H2O Driverless AI**, and **.NET**.
- Skilled in **AutoMLOps** and leveraging **AWS** for seamless workflow automation, model deployment, and integration, improving business efficiency with **Python**, **R**, and **C#** across diverse applications.

SKILLS

Programming Languages: Python, SQL, R, C#, VBA, JavaScript
Database: SQL Server, PostgreSQL, MySQL, Aurora, DynamoDB
Cloud Technologies: AWS, Azure, CMS Cloud, H2O Driverless AI
Frameworks: TensorFlow, Keras, H2O, .NET MVC, ASP.NET
Big Data Ecosystem: Hadoop, Spark, Databricks, Snowflake, Apache Airflow, AWS Glue, Tibco EBX
Methodologies: SDLC, Agile, Waterfall
Version Control: GitLab, GitHub
BI Tools: Tableau, Power BI, Alteryx
IDE's: Visual Studio, Jupyter Lab, PyCharm
DevOps: Docker, AWS Codepipeline
NLP: LLaMA, BERT, GPT3, SpaCy

EDUCATION:

Master of Science in Engineering Data Science | University of Houston, Houston, TX — **Dec 2024**
Bachelor of Technology in Electrical and Electronics | National Institute of Technology Calicut — **May 2021**

EXPERIENCE:

Measurement and Evaluation Center, USA | Research Assistant — **Jan 2023 – Current**

- Deployed and maintained .NET web applications with ASPX webforms, enhancing user experience and managing the full application lifecycle.
- Fine-tuned the LLaMA3 70B model for comment summarization, improving ROUGE scores by 20% through RAG and prompt engineering.
- Deployed LLaMA3.2 on Azure Kubernetes Service (AKS), exposing it as a REST API via Azure API Management for seamless .NET integration.
- Enabled real-time summarization using Azure Functions and MS SQL Server, processing over 46,000 evaluations with Azure Blob Storage and Application Insights.
- Managed MS-SQL databases and streamlined over 50 PL/SQL procedures resulting 30% reduction in execution time.
- Automated PDF report generation for 46,000+ student course evaluations using VBA and Python.
- Developed end-to-end BI dashboards in Tableau, visualizing three decades of institutional data for over 10,000 stakeholders.

Accenture, India | Data Science Analyst — **Jun 2021- Dec 2022**

- Designed AWS Lambda pipelines for model outputs and risk profiling, improving decision accuracy and reducing processing time.
- Extracted and processed customer data from HDFC systems to PostgreSQL on the cloud, enabling advanced data analysis and model training with over 1,000 features using Python and SQL.
- Optimized service request categorization, reducing AWS Lambda execution time by 50% and boosting throughput.
- Deployed predictive ML models using H2O Driverless AI to categorize loan applicants based on financial data.
- Built Hadoop and Spark data pipelines for scalable genomic data processing with fault tolerance.
- Automated CI/CD pipelines in AWS CodePipeline, streamlining metadata migration for Novartis data system with quick deployment of changes
- Orchestrated metadata migration from Snowflake to TIBCO EBX using Databricks and AWS Glue, optimizing ETL.

Mainavi Solutions, India | Data Analyst Intern — **Jun 2020 – Apr 2021**

- Captured real-time IoT sensor data using Amazon EC2 and Kinesis Data Streams to process equipment data.
- Developed and deployed XGBoost models to predict equipment failures, reducing downtime by 25% and optimizing maintenance schedules.
- Created interactive QuickSight dashboards connected to Aurora for real-time monitoring of equipment performance and maintenance.
- Collaborated with cross-functional teams to meet business goals and explore new technologies.
- Monitored and fine-tuned model performance post-deployment to improve prediction accuracy.

PROJECTS:

- **Store-Item-Demand-Forecasting-Challenge**: Developed a time series forecasting model using XGBRegressor, with moving average for noise reduction, predicting item demand. [code]
- **Topic Modelling**: Applied LDA model for newspaper content analysis, refining a corpus and identifying key features for insightful visualizations. [code]
- **Page Block Classification Ensemble Model**: Built an ensemble of 10 ML models for web page section classification, optimizing hyperparameters for improved accuracy. [code]
- **Intelligent ECG Acquisition System for Chronic Heart Disorder Detection**: Developed an event-driven process for real-time ECG data acquisition, achieving a high F-measure of 0.88.
- **Generative AI and NLP for Healthcare**: Implemented NLP workflows using GPT-3 and LLaMA3 models for healthcare claims data analysis, improving decision-making accuracy.
- **Diabetes Progression Prediction**: Built a predictive model with multiple linear regression in R to assess diabetes progression, achieving high predictive accuracy. [code]
- **Taxi Management System**: Developed a full-stack platform using React, Node.js, and MySQL for real-time taxi tracking and payment integration. [code]
- **Course Evaluation Summarization**: Fine-tuned LLaMA3 models for summarizing 10,000+ course evaluations, automating report generation with AWS and Python. [code]