

Sairam Reddy L

LinkedIn: www.linkedin.com/in/sairam-0607

Hyderabad, Telangana • (+91) 9032226607 •

lsairamreddy53@gmail.com

OBJECTIVE:

To obtain a challenging and dynamic role in Machine Learning Engineering and Data Engineering within an innovative organization where I can apply more than two years of diverse experience in developing and deploying machine learning models and crafting robust data pipelines along with my proficiency in SQL and ADF, to drive value for the company. Dedicated to leveraging advanced algorithms and engineering principles, coupled with proficiency in Python, and TensorFlow to drive impactful insights and scalable solutions. With a passion for data and a deep understanding of data architecture and usage of Azure Data Factory. I am confident in my ability to design, develop, and optimize complex solutions that enable data-driven decision-making.

PROFILE SUMMARY:

- Demonstrated hands-on experience in developing and deploying machine learning models using Python libraries such as TensorFlow, Scikit-learn, or PyTorch, showcasing proficiency in model selection, training, and evaluation.
- Expertise in designing and implementing ETL processes using Azure Data Factory (ADF).
- Proficient in SQL with experience in writing complex queries
- Experience in working with large datasets, optimizing queries for performance employing advanced data manipulation techniques, and conducting exploratory data analysis (EDA) to derive meaningful insights, leveraging Pandas, NumPy, or equivalent libraries.
- Experience in architecting, implementing, and optimizing end-to-end data pipelines or similar technologies implementation, ensuring scalability, reliability, and efficiency in data processing.
- Strong analytical and problem-solving skills, with the ability to work in a fast-paced environment.
- Good communication and collaboration skills, with experience working in cross-functional teams.

EMPLOYMENT:

ERNST & YOUNG (EY)

ML/AI & DATA ENGINEER

Feb 2022 to till date

Project 1: Invoice Extraction(Unilever Client)

- Leveraged OpenAI and open-source LLM models for invoice data extraction, accommodating various regions and languages.
- Applied OpenAI family models for document classification and Named Entity Recognition (NER), enhancing information extraction processes.
- Developed financial chatbots using OpenAI assistance, facilitating dynamic and informative user interactions.
- Proficient in prompt engineering to optimize language model outputs for enhanced performance and accuracy.
- Deployed Mistral and LLaMa2 open-source models in Azure Machine Learning Studio, serving as API endpoints for tasks like text and intent classification, as well as sentimental analysis.

Project 2: Inventory Redeployment

- Creating ETL (Extract, Transform, Load) pipelines in ADF to collect and process data related to SKUs, inventory levels, demand forecasts, and transportation data.
- Knowledge of Azure Cognitive Services for integrating advanced AI capabilities such as natural language processing (NLP) and computer vision into data processing pipelines.
- Designed and executed Azure Databricks for solving Linear Programming optimization model for SKU types A & BC, both intra and inter-regions, to minimize transportation costs while maximizing revenue and demand satisfaction.
- Implemented continuous integration and continuous deployment (CI/CD) pipelines using Azure DevOps to ensure code quality and deployment reliability. Include unit tests within CI/CD pipelines for code quality checks.

Project 3: OCR(Kringle)Project

- Skilled machine learning with a strong background in driving process automation, Leveraging OCR technology, and optimizing business workflows. Proficient at bridging the gap between business needs and technical solutions.
- Orchestrated seamless integration of OCR tools, including ABBYY Vantage, Azure Logic Apps, and Python, optimizing data extraction from various document types
- Building efficient pipelines for receipt parsing. Proficient in extracting attributes and prices from JSON outputs generated by the Google Doc AI engine. Skilled in leveraging Python to process and organize the extracted data, ensuring seamless storage in dedicated tables. Developed Python scripts to process and manage extracted data, ensuring seamless integration with organizational workflows.
- Ingested data from data sources using a combination of SQL, Google Analytics APIs, and other APIs using Python to create data views to be used in BI tools.
- Familiarity with machine learning and AI concepts for implementing data-driven solutions.
- Proficiency in setting up and managing data orchestration workflows using Azure Data Factory for complex data integration and transformation tasks.

(INTERNATIONAL SCHOOL OF ENGINEERING) INSOFE

INTERN

Nov 2019- Jan 2019

Project: Churn Prediction team

- Developed analytical and modeling frameworks to predict customer churn in the telecommunication industry.
- Leveraged quantitative and qualitative features to accurately identify customers likely to exit contracts.
- Data Visualization techniques to present insights and patterns of churn and non-churn customers.
- Data Analysis and pre-processing to ensure high-quality input for predictive models, demonstrated expertise in data modeling, and creative predictive models that offer valuable insights into customer behaviors.
- Worked on Copy Data activity, Meta Data activity Web Activity, Webhook, lookup, Store procedure, Foreach, IF, and execute Pipeline activities.
- Designed and implemented Snowflake schemas, including logical and physical data models, to optimize performance and scalability.
- Configured and managed Snowflake security features, including access controls and user permissions, to ensure that data solutions are secure and compliant with industry standards.
- Collaborated with data engineers, data scientists, and business analysts to understand business requirements and translate them into technical solutions.
- Developed SQL scripts and queries to extract insights from Snowflake data, providing valuable business insights to stakeholders.
- Monitored and troubleshooted Snowflake data solutions, identifying and resolving issues quickly and effectively

EDUCATION:

- **PG Program Diploma course in Data Science & Big Data Analytics at INSOFE (2018-19)** – Researcher & Student.

I had done a Post Graduate Program diploma in Data Science at the International School of Engineering. I have learned the different automation techniques, and Well experience in Data Science, Statistical Modelling, Machine Learning, and Deep Learning. Well in Algorithms like Supervised, Unsupervised, and Recommendation engines. I will apply feature engine techniques to increase the performance of the algorithm, with the fine-tuning parameters to select the correct parameters.

- **Bachelor of Technology in Computer Science and Engineering** from CMR College of Engineering & Technology (CMRCET), JNTUH University (2013-17)

I was part of two major projects, File Transfer using Cryptography and Rent & Track

- **File transfer using cryptography:** It is used to secure the files, documents, audio, videos from source to destination with the help of a unique key, it will generate automatically to the destination side only. This project is developed using PHP with an AES 256 algorithm with only a private key.
- **Rent & track:** With this application, we can give equipment for a certain period with unique products or equipment id and will be tracked with a unique ID.

PROJECTS:

- **Real-time data processing pipeline:** Developed a real-time data processing pipeline utilizing Azure Functions and Azure Event Hubs to ingest, process, and analyze streaming data from IoT devices. Integrated Azure Databricks for efficient data transformation and Azure Cosmos DB for low-latency data storage. Implemented data quality checks and monitoring using Azure Monitor and Azure Log Analytics to ensure high data accuracy and reliability. Enabled real-time analytics and visualization of processed data using Power BI, providing actionable insights for business stakeholders.
- **Data Lake ETL Automation:** Designed and implemented an end-to-end data extraction, transformation, and loading (ETL) solution for a large-scale data lake using Azure Databricks and Azure Functions. Orchestrated complex data transformation workflows and optimized data processing pipelines to handle diverse data sources and formats. Leveraged Azure Functions for serverless data orchestration and task scheduling, ensuring seamless and efficient data processing. Integrated data quality checks and error handling mechanisms to maintain data integrity and accuracy throughout the ETL process. Implemented automated data pipeline monitoring and logging using Azure Log Analytics for real-time visibility and performance optimization.
- **Sales Data Transformation:** Implemented an automated data transformation pipeline using Azure Data Factory and Azure Databricks to process and aggregate daily sales data from multiple sources. Designed data transformation workflows to clean, standardize, and consolidate sales data into a unified format for analysis. Leveraged Azure Databricks for complex data manipulation and aggregation tasks, optimizing the processing of large datasets. Orchestrated the end-to-end data pipeline using Azure Data Factory, enabling regular and reliable data updates. Generated comprehensive reports and visualizations using Power BI, providing valuable insights into sales performance and trends for business stakeholders.
- **Crowd Sourced Mapping:** Integrating Open Street Map crowd-sourced data and Landsat time-series imagery for rapid land use/land cover (LULC) mapping, the training data automatically extracted from crowd-sourced Open Street Map. OSM and Normalized Difference Vegetation Index (NDVI) Landsat time-series data used for LULC classification. Training data extracted from OSM “land-use” and “natural” datasets. Noise-tolerant classification algorithms evaluated for dealing with data noise. SMOTE is used to mitigate the effects of class imbalance in the training data set

CERTIFICATIONS & TRAINING SKILLS:

- Data Science, Machine Learning, Deep Learning, Artificial Intelligence
- Chat-GTP,LLM Models(OpenAI,Mistral,Llama2)
- Azure (ADF, Logic Apps, Function App, Databricks)
- Power BI
- ABBYY Vantage