R Gopal Ram

7053242442 I rgopalram737@gmail.com I www.linkedin.com/in/rgopalram

Education

Puducherry Technological University I Puducherry
Electronics and Communications I 07/2024

Projects and Experience

PIMS I Puducherry
Data Engineer I 06/2024 – 12/2024

Project: Healthcare website for early prediction of Alzheimer's disease (Freelance)
Tools: Python, PySpark, SQL, MongoDB, AWS

- Implemented scalable **ETL pipelines** using **PySpark** and **SQL** queries to process patient medical data.
- Developed data integration workflows to merge structured and unstructured data from multiple sources into **MongoDB**, ensuring data consistency.
- Implemented **data quality** & validation checks in the **ETL pipelines** to maintain accuracy in patient records.
- Created automated data transformation process using **Python** scripting to standardize diverse medical datasets from different sources.
- Created **MongoDB** with **flexible schema** design to efficiently store and retrieve patient data.
- Built monitoring dashboards using **PySpark** and **MongoDB aggregation pipelines** to track KPIs.
- Deployed ETL workflows and data pipelines on **AWS** infrastructure, leveraging services like **EC2 and S3**.

L&T Precision Engineering and Systems I Bangalore
Project Intern I 01/2024 - 05/2024

Project: C-Driven Multilingual Offline Translator Using NLP
Tools: Python , PySpark, SQL, C-Language

- Created **ETL pipelines** using **PySpark** to process large multilingual datasets, reducing data preprocessing time while maintaining data quality for model training.
- Implemented **SQL** queries to manage and retrieve training data, creating a streamlined data flow architecture.
- Implemented **data cleaning and validation** procedures using **Python** and **Spark**, ensuring high-quality training data that improved model **accuracy by 80%** in final benchmarks.
- Collaborated with **embedded systems engineers** to optimize data processing architecture for resource-constrained environments while maintaining translation quality.
- Developed **Python** scripts for model evaluation metrics and performance monitoring to ensure translation accuracy benchmarks were met.
- Developed documentation for **data pipelines & model deployment**, facilitating knowledge transfer.
- Implemented memory-efficient data structures and algorithms in **C** for **embedded system** implementation, optimizing runtime performance while working within hardware constraints.

Skills

- **Programming Language:** Python, C++, JavaScript
- **Web Development:** HTML, CSS, ReactJS, NodeJS, Tailwind CSS, Flask
- **AI/ML:** PyTorch, Tensorflow, Scikit-Learn, Pandas, NumPy, SciPy, Data Science, Machine Learning, Deep Learning, NLP, Generative AI, HuggingFace, BERT, Transformers, CNN, RNN, LSTM, Classification, Regression
- **Cloud Tech Stack:** AWS (EC2, S3) and Azure
- **Databases:** MongoDB, SQL, NoSQL, PySpark, ETL, Data Warehouse
- **Others:** DSA, OOPs, Linux (Ubuntu), Windows OS, Shell Scripting, Docker, DevOps (Basics), GitHub

Certifications

- Introduction to Machine Learning I **IIT Madras**
- Development of soft skills I **IIT Kanpur**