MOHAMMAD RAEEZ

Data Engineer / Backend Engineer

LIVERPOOL 07767952576 mohammadraeez.mec@gmail.com

PROFESSIONAL SUMMARY

Results-driven and versatile Data Engineer/Backend Engineer with around four years of progressive experience in the dynamic intersection of data engineering, machine learning and software development. Diverse background spanning media and broadcasting, retail, automotive, and entrepreneurial ventures. Proven track record in leading teams, architecting solutions, implementing, and maintaining data pipelines, developing scalable backend systems, and driving cutting-edge projects. Holds a master's degree with distinction in Big Data and High-Performance Computing from the University of Liverpool. Deep thoughts, a drive for quality, a fervent enthusiasm for innovative ideas and leadership skills are my biggest strengths.

LINKS

LinkedIn GitHub

TOOLS, TECHNOLOGIES, AND PRACTICES

- **Languages:** Python**,** C/C++, SQL
- **Backend Development:** Django, Flask, Rest API, ORM, SQLAlchemy, ELK Stack, Celery, RabbitMQ, Redis, Prometheus, Grafana, Alation, Elasticsearch
- **Databases:** PostgreSQL, Microsoft SQL server, SQLite, MySQL, MongoDB
- **Data:** Databricks, ADLS, Airflow, Pyspark, SSIS, Terraform, Data Engineering, ETL, ELT, Data Cataloguing, Data Analysis, Data modelling etc.
- **Tools:** Jupyter, SSIS, SSMS, PgAdmin, Pandas, TensorFlow, PyTorch, Visual Studio
- **Cloud Services:** Azure, Synapse, Data Factory, AWS, Terraform
- **Version Control and CI/CD**: Git, GitHub, GitLab, GitHub Actions, GitLab CI/CD Pipelines
- **Containerization and Orchestration**: Docker
- **Distributed Computing**: Apache Spark, Hadoop, Celery, OpenMP, MPI

- **Machine Learning**: ML algorithms, Scikit Learn, PyTorch, TensorFlow, Computer Vision, Object Detection, OpenCV,
- **Practices**: Strategic Planning, Team Leading, backend architecture and deployment
- **Methodologies and Collaboration tools**: Agile, Scrum, Jira, Slack, MS Teams

EMPLOYMENT HISTORY

Tredence Analytics

Bengaluru,India

Senior Data Engineer

Sept 2021 – Aug 2022

➢ Cloud development and data migration Project:

- Contributed to a collaborative effort facilitating the migration from an on-premises APS system to Microsoft Azure. Conducted **Root Cause Analysis** (RCA) to identify and resolve data and count mismatches between the two systems. **Successfully implemented fixes**, ensuring seamless migration to the cloud **within the established timeline.**
- **Single handedly** developed **a highly appreciated Recon Framework in Spark and Databricks** which compares the data for count and data quality mismatch between the 2 databases (on prem APS and Azure Synapse, **around 400 large tables**) and provides a detailed report for it. This initiative resulted in a **significant reduction in manual hours** previously spent on data validation tasks for approximately 400 large tables.
- **TECHNOLOGY AND WORK RESPONSIBILITIES:** Microsoft Azure, Databricks, ADLS, Azure Synapse, Data Factory, SSMS, SSIS, Python, SQL, Visual Studio, RCA and fix, Development of recon framework, supervising the cloud systems.

➢ Data catalogue platform:

- Our team developed a **comprehensive data catalogue** (using Alation) for a prominent American discount department store chain, streamlining data management processes and enhancing business intelligence capabilities. This project involved **metadata collection, data dictionary creation, catalogue implementation for BI servers, and lineage information gathering**. Used Alation's Python APIs to seamlessly update and maintain all gathered information within the data catalogue tool, ensuring real-time accuracy and accessibility.

- Reduced data discovery time by 40% through the implementation of a centralized data catalogue.
- Enhanced cross-functional collaboration by providing clear lineage information, resulting in a significant decrease in data-related communication gaps.
- **TECHNOLOGY AND WORK RESPONSIBILITIES:** Python, Alation (Data Cataloguing tool), Python scripting, information collection and overall management.

TATA ELXSI

Trivandrum, India

Software Developer, ML Engineer, Data Engineer

Dec 2018 – May 2021

➢ CAN data analytics:

- Data analytics project focused on **processing and transforming vehicle CAN** (Controller Area Network) data to **identify critical signals** related to specified issues, with a particular focus on High Voltage (HV) Faults.
- Implemented advanced analytics techniques to prioritize and extract the most crucial signals from CAN data, resulting in a 30% reduction in signal noise and a more focused approach to issue identification. These identified signals were then passed on to a fault detection algorithm to precisely identify HV Fault issues within the CAN data.
- **TECHNOLOGY AND WORK RESPONSIBILITIES:** Python, ELK Stack, Can-DB++, Docker, Architecture and Deployment.

➢ Video Analytics Tool:

- Contributed to the development of AIVA, an **in-house intelligent video analytics platform** for content curation and object action event meta-tagging. The tool had different features like metadata extraction, actor recognition, emotion recognition and compliance moderation tasks such as explicit content and violence detection.
- Worked diligently as a backend engineer, **integrating disparate modules** from the ML team, to ensure cohesive functionality and optimal performance. **Implemented and optimized interactions with MongoDB** for efficient storage and retrieval of metadata, contributing to the overall responsiveness of the platform.
- **Utilised Flask to design and develop** the web application, creating a user-friendly interface for interacting with the AIVA platform. **Implemented RESTful APIs** in Flask,

facilitating communication between different components of the system and enabling external services to interact with AIVA programmatically.

- **Designed and implemented asynchronous task queues using Celery and RabbitMQ**, enabling the platform to break down video processing tasks into smaller units that could be **distributed** and processed in **parallel** across cluster.
- The AIVA platform not only streamlined content curation processes but also demonstrated a **30% reduction in processing time** through the efficient utilization of Celery for distributed video processing.
- **TECHNOLOGY AND WORK RESPONSIBILITIES:** Python, OpenCV, different ML libraries, MongoDB, Flask, Celery, RabbitMQ, Docker, Architecture and deployement.

➢ IPTV AAMP Player:

- As a C++ developer, contributed to the **development and maintenance** of an in-house lightweight video player, the 'Advanced Adaptive Micro Player (AAMP),' designed for the playback of IP-Linear, IP-VOD, and Cloud DVR content in RDK-V boxes used by one of the largest American telecommunications and media conglomerate.
- My responsibilities involved a comprehensive range of activities throughout the software development life cycle. Conducted thorough **root cause analyses** (RCAs) to pinpoint the underlying reasons for user reported issues involving playback, functionality, or performance concerns.
- Collaborated with team members to investigate complex issues and **determine optimal solutions.** Applied **necessary code changes** and enhancements to address identified issues. Conducted **rigorous testing** locally to ensure that implemented code changes met **quality standards** and did not introduce new issues. Utilized version control systems such as Git and Jerrit to manage code repositories and facilitate collaborative development.
- **TECHNOLOGY AND WORK RESPONSIBILITIES:** C/C++, Git, Jerrit, JIRA, Makefile, Eclipse IDE, GCC, Development and Bug Fixing.

PROJECTS

➢ Software Developer, University Dissertation Project

May 2023 – Sep 2023

- Developed a desktop application called **Vision Quest**, that harnesses **computer vision technology** to recognize products within images, with a specific emphasis on product identification. This application empowers users to capture an image of a product via a desktop app and precisely classify it into its respective category.
- **Integrated various pretrained models** into a unified framework. Compiled a bespoke dataset of Nike shoes by crawling images from the web, manually annotating them, and subsequently training a computer vision model using this dataset.
- My innovative approach to this project was **well-received by my supervisors**. Acknowledgment of my efforts culminated in the receipt of the **maximum grade**, attesting to the quality and impact of my work.
- **TECHNOLOGY AND WORK RESPONSIBILITIES:** Python, Django, PostgreSQL, OpenCV, Pytorch, TensorFlow, React.js, Next.js, Tailwind CSS.

➢ Team Lead, University Group Project **Jan 2023 – March 2023**

- **Lead the team** of 8 to successfully develop a **fitness tracking app** as part of university group project. **Effectively communicated** project goals, timelines, and expectations, ensuring a shared understanding among team members. Took the ownership of back-end development due to some unforeseen circumstances.
- My proficiency in **team management** and the implementation of a **strategic contingency plan** earned widespread appreciation. The ability to guide the team through challenges and ensure project completion underscored my effectiveness as a leader.
- **TECHNOLOGY AND WORK RESPONSIBILITIES:** Python, Django, PostgreSQL, Rest APIs, Docker, backend development, leadership, team management.

➢ Data Engineer, Personal Project **Oct 2023 – Nov 2023**

- Built an **efficient and automated data feed** using the Spotify API. It aimed to collect and organize user-specific music listening data on a daily basis. Created a **robust ETL** (Extract, Transform, Load) pipeline that runs daily and downloads and store user song data from Spotify into a local SQLite database.
- Built a **monitoring framework** of the Airflow cluster using StatsD metrics and visualising it using Prometheus and Grafana

- **TECHNOLOGY AND WORK RESPONSIBILITIES:** Python, Apache Airflow, SQLite, Prometheus, Grafana, SQLAlchemy, StatsD, Rest API, Json, ORM, data engineering.

➢ Co-Founder, ShapeRoute **Jan 2022 – Feb 2023**

- Co-founded and worked on an app in the fitness industry out of passion with a friend of mine. Undertook range of responsibilities managing and growing the app, leading a team of 5.
- **TECHNOLOGY AND WORK RESPONSIBILITIES:** Strategic Planning, Product Development, Database design and backend development, Project Management, Team Leadership

EDUCATION

MSc Big Data and High-Performance Computing, University of Liverpool, United Kingdom

Passed with Distinction

Sept 2022 - Sept 2023

Bachelor of Technology in Computer Science and Engineering, Model Engineering College, Kochi, India

CGPA – 7.75

Jun 2014 - Apr 2018

KEY ATTRIBUTES

⧫ Effective Leader ⧫ Problem Solver ⧫ Adaptable Professional ⧫ Curiosity-Driven
⧫ Active Communicator ⧫ Continuous Learner