MUHAMMED FARIS

Senior Data Engineer

Kerala • 8111863658 • muhammedfaris.um@gmail.com • LinkedIn • GitHub

SUMMARY

Data Engineer with 5 years of expertise, including 3+ years as a Data Engineer and 1.5 years as a Backend Developer. Proficient in designing and optimizing large-scale data pipelines, ETL processes, and data warehousing solutions.

SKILLS

- **Programming Languages:** Python, PySpark
- **Databases**: PostgreSQL, MySQL, Redis
- **Cloud Services**: AWS, Databricks
- **Orchestration and Messaging**: Airflow, Kafka, RabbitMQ
- **Dev Tools**: Docker, Linux, GIT
- **Soft Skills:** Teamwork, Critical thinking, Problem-solving, Research and Development
- **Languages:** English, Malayalam

WORK EXPERIENCE

Senior Data Engineer, Chi-Square labs   **Nov 2021 - Present**

- Expert in maintaining and optimizing *ETL pipelines*, ensuring the seamless flow and transformation of large-scale data across multiple systems.
- Proficient in utilizing technologies such as *Python, Airflow, Docker, and PostgreSQL* for efficient data processing, orchestration, and storage.
- Proficient in extracting and transforming structured and unstructured data using Python libraries such as *BeautifulSoup, Playwright, PDFMiner, pdfplumber, Camelot, JSON, Pandas, and ocrmypdf*.
- Hands-on experience deploying and managing data solutions using AWS services like *S3, Lambda, ECS, and EC2*.
- Worked with *Databricks* for big data processing, collaborative analytics, and notebook-based development.
- Skilled in *troubleshooting* and resolving issues related to data processing, ensuring minimal disruption to business operations, and maintaining high data availability.
- Implemented and managed communication between different services using *FastAPI, Kafka*, and *RabbitMQ*, ensuring reliable and real-time data exchange across distributed systems.

Software Engineer, Xanthron E-Solutions   **Jul 2020 - Oct 2021**

- Collaborated within a team dedicated to building applications in the education domain.
- Designed and implemented efficient, scalable RESTful APIs using *Flask*, contributing to seamless backend functionality.
- Utilized a tech stack encompassing *Python, Flask, and MySQL* for robust backend development.
- Collaborated with front-end teams to design and refine APIs, ensuring seamless integration.

ACHIEVEMENTS AND CERTIFICATIONS

- **Databricks Fundamentals Accreditation [E-VP9791]**
- Achieved **HIPAA** (Health Insurance Portability and Accountability Act) certification.
- Identified and resolved a *memory leak* in the Python library *pdfplumber*, improving the efficiency and stability of the document processing pipeline.
- *Automated key processes*, reducing manpower requirements and operational costs, increasing overall efficiency and resource allocation.

EDUCATION

Bachelor of Science in Computer science   **2017 - 2020**

University of Calicut

- Gained a strong foundation in software development, algorithms, and data structures.