Syed Farhan Sr

farhan.syed.sr@gmail.com • +91 7019557070 • LinkedIn • GitHub • Portfolio

Summary

Recent MCA graduate with expertise in data modernization, ETL pipelines, and cloud platforms like GCP. Proven experience in migrating legacy systems using Hadoop and optimizing databases for real-time analytics. Collaborative team player with Agile/Scrum methodology exposure, eager to solve complex data challenges at IBM's Client Innovation Centers.

Education

Master of Computer Applications, Computer Science   Apr 2022 – Nov 2024
CMR College

Bachelor of Computer Applications, Computer Science   May 2018 – Sept 2021
SFGC College

Academic Projects

Sales Dashboard Modernization   *Academic Project — Tools: GCP, Hadoop, Apache Kafka*

- Migrated 10,000+ CSV records from on-prem SQL to GCP using Hadoop, reducing storage costs by 15%.
- Implemented CDC pipelines with Apache Kafka (self-taught via online labs), automating incremental data replication.
- Collaborated in 4 Agile sprints using Jira, delivering the project 2 days early.

Heart Disease Prediction API   *Academic Project — Tools: Python, Flask, Docker*

- Deployed a logistic regression model (85% accuracy) as a REST API using Flask, containerized with Docker.
- Improved model precision by 10% through grid search hyperparameter tuning.

Elasticsearch Library System   *Academic Project — Tools: Elasticsearch, MongoDB*

- Reduced search latency by 25% using Elasticsearch during a university library project.
- Automated borrowing reports with Python and stored logs in MongoDB for user analytics.

Skills

- **Data Engineering**: ETL/CDC Pipelines, Hadoop, GCP, Data Modeling
- **Programming**: Python, SQL, REST APIs, Flask
- **Databases**: MongoDB, Elasticsearch, MySQL
- **Machine Learning**: Scikit-Learn, Logistic Regression
- **Tools & Methods**: Jira, Git, Agile/Scrum
- **Currently Exploring**: Apache Airflow, AWS S3

Certifications

- Big Data Foundations, IBM   May 2025 (Completed)
- Machine Learning, Google   May 2025 (Completed)
- Hadoop Foundation, IBM   June 2025 (Completed)