

Anmol Kumar Chauhan

Data Engineer

+91- 8923393212

✉ acrajpoot001@gmail.com

in Anmol-K-Chauhan

Passionate professional with a strong background in Data Engineering, Project Execution; Focused on leveraging expertise in Data Engineering Leadership to drive successful outcomes and contribute to organizational growth in targeted functional areas and industries

Core Competencies

Date Engineering

Solution Designing

Requirement Gathering & Analysis

Software Development Lifecycle

Process Excellence

Risk Management

Agile/ Scrum Methodology

CI/CD Methodologies

Big Data Technologies

Education

- Bachelor of Technology in Computer Science and Engineering from Dr. Ram Manohar Lohiya Avadh University in 2020

Profile Summary

- A results-oriented professional with over 4 years of hands-on experience in Big data engineering using Python. Proficient in architecting and implementing robust **ETL pipelines for Big Data utilizing Azure Data Factory**, contributing to seamless **data ingestion, warehousing**, and reporting, managing up to **30 gigabytes of data daily**.
- Committed to treating data as a strategic asset** and fostering a culture that prioritizes data quality, resulting in the creation of reusable data products pivotal in enhancing organizational-wide data capabilities, particularly in the railway domain.
- Excellence in delivering innovative, data-driven solutions** using **big data technologies** to consulting clients, resulting in enhanced business success and informed decision-making. Noteworthy achievements include analyzing and leveraging complex data sets to identify critical patterns and provide strong advisory support for optimal ML model implementation. Collaborative work with Data Engineering, ML, and DevOps teams, ensuring timely and efficient completion of projects driven by data insights leading to a **30% cost reduction** through meticulous cost optimization strategies.
- Proficient in designing and implementing scalable data lake and data warehouses**, integrating various **Azure services** for improved performance and cost efficiency while utilizing Kubernetes for resource deployment, with improved scalability, high availability and resource efficiency on Azure.
- Skilled in **Agile Project Management Methodologies** with skills in directing project teams and maintaining high-quality deliverables within time and budgetary parameters
- Proven capabilities in building relationship** with clients & partners for evolving strategic vision, driving change, building services roadmap, infusing new ideas, implementing emerging technologies for the long run and taking enterprise system performance & productivity to next level

Work Experience

29 Nov'2021-Present | Alstom Transport India Limited
As Data Engineer

Key Result Areas:

- Spearheaded the **design, development, deployment, and management** ensuring high availability, security, and scalability using AZURE.
- Collaborated seamlessly with **cross-functional teams** to gather project requirements, define clear objectives, and establish project timelines, streamlining development processes, and reducing time-to-market.
- Led and managed a **data engineering team**, developing and maintaining **scalable data pipelines** for analysing **big data** available at different sources in different format.
- Handled ingestion of **Big Data** over **30 GB data** per day.

Technical Skills

- Python
- Big Data tools
- PySpark
- Hive
- Hadoop
- Azure DevOps
- NiFi
- Databricks
- Azure Data Factory
- SQL and NoSQL
- Data Quality Checks
- PostgreSQL
- Flask
- Azure
- Data Analytics
- ETL
- Basic Machine Learning
- PowerBI
- Statistics

Personal Details

Date of Birth: 27th July 1996

Languages: English & Hindi

Address: Bengaluru, India

- Designed, developed critical software modules and implemented **cutting-edge development frameworks**, showcasing expertise in **system design**.
- Expertise in **processing signaling data** from Rail Signaling Systems from different locations, including Amsterdam, Santiago, Tel Aviv, Octys, Lyon, and Riyadh.
- Collaborated with cross-functional teams to ensure **data quality** and accuracy and optimize value from **analytics platforms**.
- Created and implemented **analytical algorithms** such as Ping Pong, Synchronization algorithms, and packet loss algorithms during handover to provide meaningful data insights.
- Developed **statistical models** and **KPI** to identify patterns in signaling data, and implemented algorithms for tasks such as **data normalization**, aggregation, and statistical analysis using **Python** and **SQL**.
- Designed and implemented an efficient orchestration system for an Azure-based **data lake**, managing end-to-end monitoring of data ingestion, warehousing, and reporting processes, and optimized data pipeline orchestration using **NiFi** and **Apache Airflow**.
- Spearheaded the migration of **big data** processing from local instances to a **distributed cluster environment**, delivering a **70% increase in data processing speed**.
- Collaborated in the development of multiple **Python Flask** services aligning with intricate business rules.
- Communicated project status updates to various project partners, including team members and executives by synthesizing a multitude of issues into a clear, concise, and focused message.
- Conducted comprehensive **code reviews**, provided constructive feedback to enforce best practices in **code quality**, maintainability, and consistency, and offered **mentorship** to junior developers.
- Troubleshoot, resolved complex technical issues, ensured **system stability**, and optimized **performance** consistently.
- Educated coached team on **Agile methodology** to achieve quality handover within time and cost guidelines and conducted collaborative efforts with **stakeholders** to gather information and requirements for project delivery, while **refactoring code** for improved optimization, performance, and efficiency.

1 Sep'2020-12 Nov'2021 | eAdvert Promotional Studio
As Data Analyst Intern

Key Result Areas:

- Analyse **large volumes** of data using statistical methods.
- Create visualizations to help stakeholders understand the meaning of data.
- Develop ETL pipelines (**Extract, Transform, Load**) or data mining workflows.
- **Extract** and **transform** data from multiple sources to generate insights.
- Ensure **data accuracy** and apply checks on **data quality**.
- Verify completeness of data sources.
- Clean and format data as necessary.
- Organize, store, and retrieve data efficiently.
- Design and produce regular reports and dashboards using **PowerBI** tool.
- Work to interpret data trends to provide actionable insights.



Achievements

- Third position in Alstom Statistics Competition for innovative statistical analysis and modeling approach.
- Invited to attend a Data Science workshop organized by Alstom in Saint Owen, France; showcasing strong critical thinking ability and passion for data-driven insights.
- Received recognition for providing excellent support to the Tel Aviv Project with radioscopy tool, leveraging Data Engineering skills to improve radio coverage.