Adwiteey Mauriya

https://www.amauriya.com
+44 7467080215
send@amauriya.com
Nationality: Portuguese

www.amauriya.com

ABOUT ME

I am an experienced data engineer who has a wide range of technical skills including prompt engineering. I am proficient in Apache Airflow, Snowflake, BigQuery, Hadoop, Spark, Dagster, Python, SQL, and MLflow, making me a valuable asset to any data-driven organization. With a deep understanding of these technologies, I can design, implement, and maintain complex data pipelines with ease. I am dedicated to delivering high-quality results, and has a proven track record of successfully implementing data projects from start to finish. I am an effective communicator, able to work well both independently and as part of a team. The combination of technical skills and experience make me an ideal candidate for any data engineering role.

EDUCATION

- **Ph.D.** 2015 - 2020 in **Technological Physics Engineering** at Instituto Superior Tecnico, Lisbon, Portugal. University of Oxford, U.K.
- **B.S.-M.S Dual degree,** 2009 - 2014 with **Physics Majors** at Indian Institute of Science Education Research, Pune, India.

TECHNICAL SKILLS

Advanced: Python, SQL, C++, Scala, Kotlin, Bash
Intermediate: C, JavaScript
Elementary: Ruby

SOFT SKILLS

Curious, life long learner, problem solver, team player.
Experienced in presenting results to expert and non-expert audiences in formal and informal settings.
Languages: English, Hindi, Portuguese

TECHNOLOGIES

Big data: Hadoop, PySpark, Airflow, Dagster, dbt, N8N
Devops: Kubernetes, Docker, EKS
Database: Snowflake, BigQuery, Postgres, Mysql, Clickhouse
ML: Scikit-learn, Keras, Tensorflow, MLflow, Langchain
BI: PowerBI, Tableau, SRSS, Apache Superset
Platforms: GCP, AWS, Supercomputers
Dev: Intellij, Vscode, Git, Jenkins, SVN
PM: Microsoft 365, Jira

KEY ACHIEVEMENTS

- Received competitive **KVPY** fellowship in 2009(for 5 years) to do BS-MS dual degree course, funded by Dept. of Science and Tech and Government of India.
- Received prestigious **FCT** fellowship (includes tuition fees + living expenses for 4 years) and Eurofusion research grants for participating in experiments and traveling to multiple laboratories in European Union.

WORK EXPERIENCE

Senior Data Engineer, Liquidity Capital

08-2022-present

I have designed and created a bank account integration solution to implement machine learning models. The models are integrated into the company's no-code platform and I successfully integrated with 50 banks worldwide. I have built investment committee automation report service using openai and prompt engineering. This service is deployed on company's app which is used investment managers. I also established a CI/CD process for the company's codebase and manage the production-level AWS Lambda and EC2 instances. Additionally, I was responsible for IT onboarding of new employees and contributed to the company's code review standard and offered input on the design of the backend infrastructure's resource management. The technologies used in the company include MySQL, N8N, Spark, Airflow, Superset, Step Functions, AWS Lambda, and Python.

Data Engineer, Tripadvisor.Inc 11-2020-07-2022

Maintained and monitored the datasets stored in multiple data lakes and warehouses (e.g. GCP, Snowflake and Hadoop clusters) of customer data platform and enterprise data teams by optimising the currently existing ETL and creating new if it is required. ETL saves cost of storage and processing of more than 10 millions USD annually by the entire team. The constructed datasets were in BigQuery, Hive and Snowflake which involved redesigning of databases, building data pipelines and optimising the underlying queries, accommodating the new data model, and liaising with stake holders. Improvements and performance was assessed in a form of dashboards and charts to have a bird eye view. I have created series of automated scripts to monitor ETL of various business segments in the company. I advocate the code review standards to have clean clear and consistent business data. I have provided support for onboarding of new data engineers in the team and the company.

Business Data Analyst (Contract), NetJets.Inc 10-2019-07-2020

Achieved 80% success in pilot training warehouse automation by building multiple reports and liaised with SMEs to identify discrepancies for troubleshooting and corrective action. We planned and organized workshop, producing 1.1 times the goal of qualified leads. Synthesized current business intelligence data to produce SRSS and Tableau reports and polished dashboards, highlighting findings and recommending changes using scrum framework and Kanban method. Developed flight school KPI metrics derived from raw company data to track improvements in gross margin of training department of the company.

Independent Project (Disruptions mitigation modeling in Tokamak using Neural Networks) 06-2019-12-2019

The purpose of this project was to create a tool/work-flow to predict the disruptions in Tokamak (Future Nuclear Fusion reactor) using neural networks by processing the measured experimental data. Tokamak has more than 200 probes and it produces few Terabytes of data with every experiment. Experimental data is stored in a noSQL database. Data is stored in binary format and can be accessed using MDSPlus server. I set up PySpark, Apache Spark and Hadoop in a workstation. Created a python library to create RDD from the NETCDF data. I used the ML library using PySpark and also run tensorflow on a Spark clustor.

Visiting Researcher at University of Oxford and Culham Center for Fusion Energy, U.K. 08-2016-09-2018

Developed numerical tools to study complex physical phenomena in nuclear fusion reactors. It required parallel programming in Fortran with MPI and Python. It is scaled on supercomputers which has CPUs and GPUs. Modeled the fusion experiments with an underlying theoretical model by developing the required numerical tools in Python. It required in-depth knowledge of mathematics, statistics, programming, and physics. Developed a Python framework to import fusion experiments data in multiphysics codes for various kinds of physics study. It also visualizes the output and fully modular.

Project Scientist at Institute for Plasma Research, India 06-2014-12-2014

I developed an algorithm and checked the time complexity(efficiency) of 1-D Grad-Shafranov equation to implement it in a parallel HPC code. Merged fix boundary equilibrium with free boundary method and patched it in a Fortran library which is integrated with a multiphysics code.

Linux System Administrator (Volunteer) at Soft matter simulation Laboratory, IISER Pune India 09-2012-04-2013

Provided IT support to a research group to resolve issues related with Linux operating systems, setting up scientific softwares, libraries and debugging High Performance Computing codes. I consistently coded multiple batch jobs in Bash and Python to automate tasks.