Email : akostuv@gmail.com

Koustav Mallick

Software Developer

Mobile : +1 (260) 348-7097
GitHub: **StaticOwl**
LinkedIn: **mallkou**

SKILLS

Languages: Scala, Bash, Java 8, Python
Framework: Apache Spark, Hadoop, Map Reduce, Maven, Hive, Apache AirFlow
Software Development: Oracle, Agile, GCP, Git, JIRA

EDUCATION

M.S || Computer Science — Fort Wayne, IN | 2023-Present
Purdue University
Relevant Coursework: Deep Learning, DB Designs, Machine Learning

B.E || Electronics and Tele-Communication Engineering — Kolkata, INDIA | 2019
Jadavpur University
Relevant Coursework: Data Structures and Algorithms; Operating Systems; DBMS

WORK EXPERIENCE

Deutsche Bank | Associate (Big Data) — Pune, INDIA | Jul 2019 - Aug, 2023

- Designed and Developed Scala-Spark common ETL framework, optimizing daily data ingestion and processing **from 3 days to 1 hour** with support for **GCP and Cloudera environments**. Also **simplified** the framework to ingest from a **single accumulated feed** instead of **40+ different sources** using **Spark, Scala, Maven and Hive**.
- Engineered Scala-Spark-based **Independent Decision Making Framework**, reducing processing time **from 7 daily-machine-hours to 3-5 minutes**, compatible with GCP and Cloudera using **Scala, Spark, Apache AirFlow, Shell, BigQuery, Hive, Impala, Beeline**.
- Created **AirFlow DAG Code Generator Utility** from simple JSON input, accessible across the organization, with support for **generating multiple DAGs with similar template, config recognition and parsing**, incorporation of custom commands, which reduced **approx. 40 man hours monthly** in creation of Airflow DAGs using **Python, Jinja, AirFlow**.
- Designed a **Regulatory Reporting Framework**, replacing **COGNOS**, for Volcker 2.0 Regulatory Compliance, saving **$4M a year** in licensing costs using **Python, HTML, CSS, Java, Shell**.
- Automated **End-To-End testing application** for Ingestion Layer, saving **approx 80 man-hours a month** and **reducing manual testing effort** for the organisation's Data Ingestion Process using **Spark, Scala, Java, SQL, Hive**.

PROJECTS

§ **ProcAlloc (Python)** - Designed a Python framework for project allocation, prioritizing new-joiners' skills and project requirements, with the flexibility of resume scanning to improve candidate selection and communication using Text Mining.

§ **Weather Monitoring System (Python)** – Designed and built end-to-end weather monitoring system using Python that displays real-time temperature, humidity etc. for any global city through OpenWeather API calls and data visualization.

§ **Personal Portfolio Website (HTML/CSS/JavaScript)** – Created a responsive personal website to showcase my projects and skills. Deployed on GitHub pages.

AWARDS/CERTIFICATES

- India Excellency Award, Deutsche Bank
- 9th Global Position in Dabble (Deutsche Bank Hackathon)
- Completed Associate Level GCP Training with Distinction (Deutsche Bank)
- Joint 3rd Position in IdeaBox (Deutsche Bank Event)