Pratik Chowdhury

(+91) 7003411604 | sandyhammer31@gmail.com | linkedin.com/in/pratikcr

SUMMARY

Experienced professional with 2 years of expertise in designing scalable pipelines for managing large-scale datasets, optimizing workflows using Azure Data Factory and Databricks. Skilled in ETL processes, seamless data migrations, and leveraging Azure Cognitive Search and Open AI models to excel in dynamic environments.

SKILLS

Languages: Python, SQL, PySpark

Tools/Technologies: Azure, Spark, Databricks, Alteryx, SAP BODS, Gen AI

CERTIFICATIONS

DP-203 (Microsoft Certified: Azure Data Engineer Associate)

Advanced SQL (HackerRank)

EXPERIENCE

Kagool   July 2022 - Present
Data Engineer   *Remote*

- **Alteryx To Azure Migration** — Jan 2024 - Present
    * **Evaluated** over 115 Alteryx workflows and **engineered** their migration to Azure using Azure Data Factory (ADF) and Databricks.
    * **Managed** data loads from 1,000 to 175 million records, ensuring accurate and timely processing for various schedules.
    * **Optimized** complex Bex queries into Databricks code, enhancing performance and scalability.
- **BW To B4 HANA Data Migration** — July 2022 - Jan 2024
    * **Developed** 125 jobs for a one-time historic migration from SAP BW to BW4, managing loads from 10,000 to 130 million records.
    * **Automated** the Reconciliation Process with a Python script, reducing manual efforts by 90%.
    * **Oversaw** two ETL cycles as a one-man team, conducting Data Sanity Checks, implementing Ad-Hoc Transformations, and resolving tickets promptly.
- **Semantic Search Proof Of Concept (POC)** — July 2023 - Sept 2023
    * **Investigated** Semantic Search Capabilities for a retail client using Azure Cognitive Search.
    * **Implemented** Semantic Ranking with Cosine Similarity Scores across a database of over 100,000 product details.
    * **Explored** Open AI models like GPT, Da Vinci, and Babbage to generate embeddings for User Queries and Product Data.
- **IoT Proof Of Concept (POC)** — Nov 2022 - Jan 2023
    * **Investigated** strategies for managing streaming data for reporting purposes in the manufacturing sector.
    * **Utilized** Azure Stream Analytics and Azure Data Explorer for real-time Power BI Dashboards integration.
    * **Implemented** Python code within Azure Logic Apps to transmit key performance indicators (KPIs) such as Heat, Humidity, and RPM for a motor.

EDUCATION

National Institute of Technology, Warangal
Master of Computer Applications - MCA, Computer Science — GPA: 7.57   *2019 - 2022*

Calcutta University, Kolkata
Bachelor's degree, Mathematics (Hons.) — Grade: First Class   *2016 - 2019*

Kendriya Vidyalaya
12th, Science — Percentage: 94%   *2016*