

RUPAM SAU

+91-9064869734

rupamsau721401@gmail.com

rupamsau0010

rupamsau0010

About

Data Engineer (Databricks Certified Data Engineer Professional | Azure Certified Data Engineer Associate)

A proficient data engineer with strong computer science fundamentals and more than two years of experience in the Azure framework, having a proven track record of building scalable systems to process large-scale data for multinational enterprises.

Technical Skills

Languages:	Python, SQL, Java, Scala, JavaScript, C, C++.
Tools/Frameworks:	Pyspark, Databricks, ADF, ADLS Gen2, Delta Tables, Azure Synapse, Power BI, ETL, Hadoop, Map-Reduce, Hive, Pig, HBase, Kafka, NumPy, Pandas
Certifications:	Databricks - Databricks Certified Data Engineer (Professional & Associate) Azure Cloud - Azure Certified Data Engineer (DP-203), DP-900, AZ-900, AI-900

Experience

Cognizant Technology Solutions	Oct 2022 – Present
<i>Data Engineer - Programmer Analyst</i>	Kolkata, WB, India
<ul style="list-style-type: none">Contributing to data engineering initiatives for one of the largest property insurance providers in North America.Designed and implemented a scalable Operational Data Store (ODS) on Azure from scratch.Optimized ETL pipelines to improve data ingestion speed, handling 7+ TBs of data efficiently.Collaborated with Data Scientists & BI Engineers to ensure high-quality data for analytics and reporting.Developed cost-effective Spark-based ETL solutions, reducing execution time by 40%.	
Cognizant Technology Solutions	Feb 2022 – Aug 2022
<i>Big Data Engineer Intern</i>	Kolkata, WB, India
<ul style="list-style-type: none">Gained hands-on experience with open-source Big Data technologies like Hadoop, Hive, and Kafka.Developed a POC for flattening large machine-generated JSON/XML files, reducing transformation time by 150%.Collaborated with senior developers to learn data engineering best practices and performance optimization.	
Mentorsity	Mar 2021 – May 2021
<i>Backend Developer Intern</i>	Kolkata, WB, India
<ul style="list-style-type: none">Designed APIs using Node.js, leveraging microservice architecture to ensure modularity, scalability, 100% availability.Utilized MongoDB as the database to design efficient data models and handle CRUD operations.Collaborated closely with the frontend engineering team to ensure seamless integration and functionality of applications.	

Projects

Simplified Task Dashboard: Featured in CALIBER 2023 – A Collaboration Hackathon by Microsoft Azure and Cognizant / Pyspark, Databricks, Delta Tables, ADF, ADLS Gen2, PowerBI, Node.js, Express.js, React.js
<ul style="list-style-type: none">Developed a unified platform for managers and associates to prioritize tasks and increase productivity by 50%.Implemented visualizations for task assignments and completion records, providing insights into associate performance
Covid-19 Data Processing / Databricks, Pyspark, ADF, ADLS Gen2 & Azure SQL
<ul style="list-style-type: none">Processed and analysed a dataset of Covid-19 pandemic on top of Azure stack.Designed ETL pipelines to transform raw data into structured formats, enabling reporting and data visualization.
Formula1 Racing Data Analysis / Databricks, Pyspark, ADF, ADLS Gen2 & Unity Catalog Delta Tables
<ul style="list-style-type: none">Designed a complete ETL flow of Formula1 Racing data set on top of Azure to get insightful understandings through visualizations.

Education

Future Institute of Engineering and Management, Kolkata (MAKAUT University)	2018 - 2022
<i>Bachelor of Technology in Computer Science & Engineering (CGPA: 9.10)</i>	
Contai Model Institution	2016 - 2018
<i>Higher Secondary in Science (90.0%)</i>	

Achievements

- Leetcode: solved 470+ questions / [rupamsau_0010](#)
- Published research paper on IEEE - Smart Attendance Tracking and Performance Monitoring System | [Link](#)
- College Level Champion – Smart India Hackathon 2019