

AMIT KUMAR

DATA ENGINEER

+91 8178796716 | mramit4848@gmail. | [LinkedIn](#) | [GitHub](#)

SUMMARY

Dynamic and results-driven Data Engineer with expertise in **ETL processes**, **big data analytics**, and **cloud-based data solutions**. Skilled in leveraging tools like **Databricks**, **Azure Data Factory**, and **PySpark** for processing and analyzing large-scale datasets. Experienced in **predictive maintenance**, **real-time tracking systems**, and **market analysis**, contributing to enhanced decision-making and operational efficiency.

WORK EXPERIENCE

EMBEDDED DESIGN SERVICES INDIA PVT LTD.
DATA ENGINEER

Oct 2022 – Present

Vehicle Tracking System (VTS)

- Developed a real-time **Vehicle Tracking System** integrating GPS, sensor data, and vehicle performance metrics.
- Leveraged **CosmosDB** for real-time data storage with high availability and low-latency retrieval.
- Utilized **PySpark** for processing and analyzing tracking data, improving response times.
- Designed interactive dashboards using **Power BI** to enhance fleet management and operational monitoring.

Remote **Diagnosis and Predictive Maintenance (RDPM) for Indian Railways**

- Implemented **predictive maintenance models** for **Indian Railways**, reducing downtime and improving train operations by forecasting potential equipment failures.
- Integrated IoT sensor data from railway systems with maintenance logs for model training and validation.
- Automated data ingestion using **ETL pipelines** and optimized storage with **CosmosDB**, enabling efficient real-time analytics.
- Developed a real-time alerting system to notify railway operators of potential issues, enabling proactive maintenance and reducing service disruptions.

E-commerce On-premise Solution

- Developed an on-premise e-commerce solution utilizing Azure services such as **Azure SQL Database** and **Azure Data Factory** for robust data management.
- Implemented Azure Functions for **serverless computing**, ensuring efficient business logic execution and scalability.
- Designed data security measures using **Azure Security Center** and **Azure Active Directory** to adhere to industry standards.
- Utilized **Azure DevOps for CI/CD pipelines**, enabling seamless deployment and updates for optimized performance and reliability.

CLOUDLOGZ (Internship)

Feb-2024 – Aug 2024

AZURE DATA ENGINEER INTERN

- Conducted **data analysis** using **Python** and **SQL** to derive actionable insights on market trends, aiding financial institutions in strategic decision-making.
- Designed and implemented **Azure Data Factory pipelines** for seamless data ingestion and transformation.
- Managed large-scale data integration with **CosmosDB** and **MySQL**, ensuring efficient storage and retrieval.
- Applied **machine learning algorithms** to predict future market behaviors, enabling financial forecasting.
- Automated data workflows and developed dashboards for financial trend analysis using **Power BI**, enhancing reporting accuracy and stakeholder engagement.

KEY SKILLS

- **Languages:** Python (Pandas, PySpark), SQL, Scala
- **Databases/Data Warehouses:** CosmosDB, MySQL, Hive, Snowflake, Microsoft SQL Server
- **Tools:** Azure Data Factory, Databricks, Power BI, GitHub, Azure DevOps
- **Analytics:** Big Data, Hadoop, ETL, Data Analytics, Predictive Maintenance
- **Cloud Platforms:** Azure (Data Factory, SQL Databases, Synapse Analytics, Databricks)

PERSONAL PROJECTS

Health Care COVID-19 Analysis

- Developed a Health Care COVID-19 Analysis solution leveraging Azure services like **Azure Databricks** and **Azure Machine Learning** for data processing and predictive modeling.
- Implemented **Azure SQL Database** for secure storage and efficient querying of COVID-19 patient data
- Utilized Azure Machine Learning to build **predictive models** for forecasting COVID-19 trends and analyzing patient outcomes.
- Designed data visualization dashboards using **Power BI** for healthcare professionals to monitor and analyze real-time COVID-19 statistics.

Movies Rating Analysis

- **Optimized data pipelines** and storage structures, resulting in faster response times and improved data processing efficiency.
- **Introduced new features** and enhanced functionalities, creating a more user-friendly interface.
- **Streamlined data** visualization and reporting, enabling better decision-making and accurate analysis.
- **Tech involved:** Azure Data Factory, Databricks, Hadoop, Apache Spark, PySpark, SQL.

EDUCATION

Kurukshetra University, Kurukshetra
Master of Business Administration (Information Technology)

Kurukshetra, Haryana
2024 - 2026

J.C. Bose University of Science and Technology, YMCA
Bachelor of Technology (Computer Science Engineering)

Faridabad, Haryana
2019 - 2023