**Arun Kumar** **Singh**

|  |  |
| --- | --- |
| “Data Engineer | 4+ Years Experience | Talend Studio | Power BI | PostgreSQL| Pyspark| Databricks, Microsoft Power BI & Fabric Certified” | +917895855917  ak.singh1321@gmail.com |

**Summary**

Data Engineer with 4 years of experience in building scalable data pipelines and analytics solutions using Talend Studio,Databricks Power BI, PostgreSQL, Python, and PySpark. Certified Databricks Data Engineer Associate, Microsoft Power BI, and Microsoft Fabric professional, with proven expertise in data integration, performance optimization, and cloud solutions on Microsoft Azure.

**Experience**

**Data Engineer**, 06/2021 - Current

**Tata Consultancy Services** - Noida

Data Engineer with **4 years of experience** in data integration, analytics, and reporting, currently working on the **Power Conversion Project at General Electric (GE)**.

* Skilled in **Talend Studio, Databricks, Power BI, PostgreSQL, Python, and PySpark**, with expertise in building scalable ETL pipelines and optimizing large datasets for performance and accuracy.
* Hands-on experience with **Microsoft Azure** for cloud-based data storage, processing, and deployment of analytics solutions.
* Designed and optimized **complex SQL queries and PySpark jobs**, reducing execution time and improving system performance.
* Developed and maintained **interactive Power BI dashboards** to deliver insights on system performance, business trends, and process improvements.
* Collaborated with cross-functional teams to **improve data quality, automate workflows, and resolve performance bottlenecks** in critical projects.
* Certified **Databricks Data Engineer Associate, Microsoft Power BI, and Microsoft Fabric professional** with strong expertise in modern data engineering and BI platforms.

**Skills**

|  |  |
| --- | --- |
| * **Data Visualization:** Microsoft Power BI (Certified), Microsoft Fabric (Certified) * **Databases:** PostgreSQL, SQL (Advanced), Data Modeling * **ETL/Integration Tools:** Talend Studio, WinSCP, Workflow Automation | * **Programming & Big Data:** Python, PySpark * **Cloud Platforms:** Microsoft Azure, Databricks (Certified Data Engineer Associate) * **Tools & Utilities:** DBeaver, Git, Version Control * **Core Competencies:** Data Pipeline Design, Query Optimization, Performance Tuning, Data Quality & Governance |

**Education**

**Master of Computer Applications** - GLA University | 2022

**Bachelor of Computer Applications** - GLA University | 2020

**Certifications**

* **Databricks Data Engineer Associate** – Databricks
* **Microsoft Power BI (Certified)** – Microsoft
* **Microsoft Fabric (Certified)** – Microsoft

**Projects**

**GE Vernova Power Conversion**

Key responsibilities

* Actively contribute to the **Power Conversion Project** for General Electric (GE), focusing on debugging data issues, ensuring smooth data flow, and resolving technical challenges in Power BI reports.
* Designed and maintained **interactive Power BI dashboards** that deliver actionable insights to stakeholders, supporting data-driven decision-making.
* Built and optimized **data pipelines** using **Talend, PySpark, and Microsoft Azure**, enabling seamless data flow across multiple systems and reducing processing time.
* Optimized **complex SQL queries**, achieving **~60% reduction in execution time**, which improved report performance and lowered infrastructure costs.
* Developed and automated **ETL workflows** to integrate diverse data sources, ensuring consistency, accuracy, and scalability of reporting systems.
* Leveraged **PostgreSQL, Python, PySpark, and Azure services** to process large datasets efficiently and enhance analytics capability.
* Collaborated with cross-functional teams to strengthen **data governance, quality, and workflow automation**, ensuring reliable and timely project deliverables.