PARTHA SARATHI DEB

+91-7980460418 **|** partha.cse1@outlook.com

TECHNICAL SKILLS

Programming Languages: Scala, Python, R.
Deep Learning Frameworks: PyTorch.
Framework and Tools: Azure DevOps, Hadoop, Git.
Azure: Synapse, Azure Databricks, Azure SQL, Azure Data Factory (ADF), Data Lake (ADLS).

PROFESSIONAL EXPERIENCE

EPAM Systems, *Senior Software Engineer*   **November 2020 – Present**

- Analyzed, designed, and built data solution using Azure technologies to support visualization in Power BI and provide Data Science training data.
- Created ADF Pipelines to Extract, Transform and Load data from different sources like Azure SQL, Blob storage and SharePoint.
- Developed Spark applications in Scala for data extraction, transformation, and aggregation from multiple file formats for analyzing & transforming the data.
- Performed code refactoring of previously written Spark code to improve performance.
- Deployed and tested (CI/CD) the code using Visual Studio Team Services (VSTS).
- Working in Agile development environment in sprint cycles of two weeks by dividing and organizing tasks.
- Conducting code reviews for team members to ensure proper test coverage and consistent code standards.
- Implemented Row Level Security to restrict data access in SQL tables.

Indian Bank, *Assistant Manager – IT*   **July 2017 – November 2020**

- Managed a team of eight developers in design, development, and enhancement of big data solutions.
- Collaborated with government authorities like RBI, Income Tax Department, Finance Ministry and Bank management to understand the requirement and design the complete system.
- Implemented fraud detection methods by analyzing transaction patterns for transactions originating from Financial Inclusion endpoints.
- Implemented customer segmentation to offer personalized services and products to customers.
- Installed an on-premises Hadoop cluster comprising 16 nodes and 2 master in active standby configuration.
- Implemented a data pipeline for ingesting daily transaction data into the Hadoop cluster using Oozie and Spark.

- Performed transformation and data analysis on raw semi-structured data using Spark to aid in business decision-making and planning.
- Performed data import and export from RDBMS sources using Sqoop and Spark.
- Created Oozie workflow jobs for scheduling queries and actions.
- Performed Hadoop cluster management and monitoring using Apache Ambari.

Tata Consultancy Services, *IT Analyst*   **March 2011 – July 2017**
- Developed Hive queries to analyze the data stored in HDFS to recommend business Improvements.
- Created data processing scripts to automate reporting tasks using Python and Pandas.
- Used Hive to analyze the partitioned and bucketed data and compute various metrics for reporting.
- Performed migration of ETL process from Oracle to Hive.
- Designed Ozzie workflows to create a data pipeline to ingest daily data and email reports to higher management.

EDUCATION

University of Illinois, Urbana-Champaign   **Jan 2022 – May 2024**
Master of Computer Science, **GPA: 4.0/4.0**

West Bengal University of Technology   **July 2005 - August 2009**
Bachelor of Technology, Computer Science and Engineering, **DGPA: 8.14/10.00**

Kendriya Vidyalaya, Cooch Behar   **2004**
Higher Secondary (CBSE), **79.60%**

Kendriya Vidyalaya, Cooch Behar   **2002**
Secondary (CBSE), **86.80%**

CERTIFICATIONS

- Microsoft Certified Azure Data Engineer Associate.   **February 2021**
- Microsoft Certified Azure Developer Associate.   **May 2021**