



## PROFESSIONAL SUMMARY

- Data Engineer with 6 years of experience, specialising in designing and optimizing ETL data pipelines on cloud platforms like **AWS, Azure and GCP**.
- Lead successful implementation and adoption of Agile/scrum ceremonies(Jira software) with product managers, architects and other engineers, resulting in a 5% increase in delivery efficiency and 90% improvement in team communication and collaboration.
- Developed code to encapsulate complex business logic, leveraging emerging trends to drive data strategy and management in the banking sector.
- Skilled in aligning data governance with enterprise data requirements to support customer acquisition and operational efficiency.
- Scripting Languages Python, Pyspark, Scala, Advanced Sql Querying, PL/SQL, T-SQL, JAVA, C# , Postgres, Object -Oriented languages, Relational Databases, Dimensional Modeling, Development Operations(Devops).
- Soft Skills Leadership, Innovation, interpersonal Skills, Organizational skills, Consistency.
- Proficient in implementing and optimizing various file format **CSV, Parquet, Avro, and ORC** for efficient data storage, access, and analysis, enhancing data processing performance and scalability in distributed environments.
- Proficient in Database Management, importing/exporting data between Relational databases, including **RDS, Sql Server, Teradata, Snowflake**. Skilled in NoSQL databases like **HBase,Cassandra, MongoDB, DynamoDB**.
- Experienced in building **Redshift** data warehousing solutions and employing **BigQuery** for storing and querying large datasets using partitioning.
- Experienced in real time event data streaming leveraging **Kafka, Spark Streaming, Kinesis**.
- Experienced working with Databricks/Cloud Storage like **Hadoop distributed file systems/Dfs, Data Lake Gen2, S3 , Delta Lake, BigLake**.
- Proficient in Serverless architecture tools like AWS **Lambda, API Gateway, GraphQL** Fargate, Azure Functions, Cloud Functions and EC2, ECS, Virtual Machines, SQL Database, **Kubernetes(AKS)/GKE**.
- Expert in cloud integration, leveraging **Data Factory, Glue, Nifi, Step Functions, Logic Apps, Airflow** Dataflow, MuleSoft and Talend to drive operation efficiency.
- Leveraged MPP architecture principles to optimize data processing, Query processing and system scalability.
- Experienced in applying advanced data modeling frameworks such as ER, Dimensional Modelling, Object-Role modeling to structure, define, and analyze complex data systems enhancing data quality and business insights.
- Developed edge to edge unit testing for production environment, ensuring stability and operational efficiency.
- Strong Writing skills, Experience in writing clear and concise documentation, reports and code comments.
- Willingness to continuous learning and stay up to date with latest industry trends and technologies.

## EDUCATION

- |  |   |          |
|--|---|----------|
| • Master's In Data Science & Analytics | The University Of Oklahoma, Norman U.S. | 12/2023. |
| • Bachelor's In Computer Science       | IIT Dhanbad, India.                     | 06/2018. |

## WORK EXPERIENCES

UNIVERSITY OF OKLAHOMA Norman, Oklahoma | United States

02/2023 to 12/2023

Research Assistant.

- Led a 50% improvement in data, system, and application integrations through an expert-guided legacy system migration to modern platforms..
- Fostered professional partnerships within and outside university, enhancing collaborative efforts and project outcomes.

Research Assistant.

- Multitasked to earn a MS in Data Science & Analytics while working at the university's Project, taking a unique challenge such as developing and enhancing a Stock market analysis application by big data (ETL/ELT) pipeline for user sentiment analysis on the U.S stock market.
- Implemented data ingestion and utilised it for batch layer processing and live streaming leveraging tool.

**Company: System Info Solutions. - Hyderabad | India**

**06/2020 to 03/2022**

Role: Senior Data Engineer.

- Led a cross functional team, fostering collaboration among team members to address technical issues and drive project management initiatives in a dynamic environment.
- Implemented process improvements and managed data operations, utilizing management tools to optimize data strategy and ensure adherence to policies and technical issues and drove project management initiatives in a dynamic environment.
- Synthesized complex datasets and business requirements to develop a robust data infrastructure , enhancing data management and governance within the banking sector.
- Developed DevOps initiative to streamline development operations, applying project management best practices to achieve consistent process improvements and meet organization goals.

**Company: Indian Eagle Pvt Ltd – Hyderabad, India.**

**05/2016 to 07/2020**

Role: Data Engineer.

- Crafted scalable solutions for handling complex data sets, improving data representations and supporting enterprise data initiatives.
- Deployed applications in google cloud, optimizing data centres and infrastructure to enhance data strategy and governance, while ensuring high uptime and operational efficiency.
- Drove customer acquisition strategies through insightful data analysis and visualizations, leveraging performance metrics to monitor and report on success.

## GRADUATE & ACADEMIC WORK

- **GRADUATE RESEARCH PROJECT**, Oklahoma Holleveraging Prediction (Machine Learning) – Project to identify patterns and techniques for micro-level analysis on Price. Boosted prediction accuracy from 45% to 85%. Performed advanced data cleaning techniques outlier detection. Used ML design patterns to build machine learning models in Azure cloud.
- **GRADUATE RESEARCH PROJECT**, ADHD Classification significance Using Neural Networks – Derived meaningful attributes from ADHD re-fMRI data and identified classification significance leveraging niLearn neuroimaging package. Detailed analysis of neural networks and its classification with ADHD rs-fMRI data.
- **GRADUATE RESEARCH PROJECT**, Covid-19 Exploratory Data Analysis – Visualised Covid-19 data and identified geographic areas with increasing cases, Abstracted pattern, and distributions.

## CERTIFICATIONS & PROFESSIONAL DEVELOPMENT

- Microsoft Certified : Azure Data Scientist Associate – Issued Nov 23 [Credential](#)
- AWS Certified Data Analytics – Specialty – Issued Nov 23. [Credential](#).
- Databricks Certified Associate Developer for Apache Spark. [Credential](#).