

Lakshmi Ravi Chandra Talluri

401-416-9428 | tallurilakshmiravichandra@gmail.com | <http://www.linkedin.com/in/ravi-chandra>



EDUCATION

University of North Texas (G. Brint Ryan College of Business)

Master of Science in Business Analytics

GPA: 3.35

Course Work: Data-Warehousing, Applied Statistics, Enterprise applications for business intelligence, DBMS, Relational database, Data-visualization, Predictive Analytics, Data Mining, Information Technology

Denton, USA

Aug 2022 – May 2024

National Institute of Technology Agartala

Bachelor of Technology in Electronics and Communication Engineering

Course work: VLSI, Power Electronics, Analog Electronics, Digital Electronics, Semiconductors, Software-Development cycle, Data Structures and Algorithms, computer science

Agartala, India

Aug 2018-May 2022

EXPERIENCE

Alephys

Houston, Texas

Data Engineer

July 2024 - present

- Developed and optimized scalable real-time data pipelines in Databricks using Delta Live Tables and Autoloader, achieving a **20% improvement in processing throughput** to support e-commerce analytics.
- Implemented ELT processes to streamline the transformation and loading of datasets into BigQuery, enhancing query performance for business intelligence by **21%**.
- Architected performance improvements using partitioning, indexing, and caching, ensuring compliance with security protocols and improving query efficiency by **30%**.
- Designed and analyzed large structured and unstructured datasets, using Apache Flink and Apache Hadoop to process high-volume data in real time.
- Streamlined data integration workflows** using AWS services, including **EC2 for scalable compute power, S3 for secure and efficient data storage, Lambda for event-driven automation, Glue for ETL orchestration, and Athena for serverless querying**, ensuring seamless data processing and accessibility
- Streamlined workflows through Bitbucket for version control, Autosys for job scheduling, and Jenkins for CI/CD pipelines, reducing deployment times by 25%.
- Devised anomaly detection frameworks** leveraging **Python's SciPy and Statsmodels** libraries to proactively identify inconsistencies in grant allocations, reducing errors by **30%**.
- Installed and secured the Confluent Kafka platform on-premises, implementing SASL_SSL encryption for secure communication and creating partitioned Kafka topics to optimize data distribution.
- Automated ad hoc reporting requests** with SQL and Python, delivering accurate insights with a **10% improvement in accuracy** and significantly reducing response times
- Managed authentication using Kerberos by creating keytabs in LDAP nodes and integrating them into the on-prem Kafka cluster, ensuring robust security for data offsets and streaming operations.
- Collaborated with data scientists to integrate machine learning models into automated workflows, resolving moderately complex data issues through applied expertise.

Nuor

Hyderabad, India

Data Engineer Intern

April 2021 – June 2021

- Designed and developed notebooks in Databricks using Python and SQL for efficient data extraction, transformation, and loading (ETL).
- Led a team of 5 in the installation of Cloudera Manager, upgrading to the latest version to enhance system performance and scalability.
- Built customized dashboards** with advanced Tableau features like **calculated fields, LOD (Level of Detail) expressions, and parameter-driven views**, offering tailored insights for Business
- Designed and automated advanced Tableau dashboards** using Level of Detail (LOD) expressions, calculated fields, and dynamic parameters, enabling stakeholders to pull actionable insights effortlessly. This streamlined reporting process by **reducing manual effort by 30%** and ensuring updates could be generated in **under 5 minutes** with minimal intervention.

SKILL SET

- Programming skills:** Python (Pandas, PySpark, NumPy, Scikit, Matplotlib), Spark, Scala, SQL, NoSQL.
- Database Tools:** Oracle, MySQL, SQL Server, NoSQL (MongoDB, Cassandra)
- Data Mining:** Linear, logistic regression, Decision Tree, Random Forest, K-NN, Naïve Bayes, Clustering, neural networks
- Cloud Platforms & Data Warehouses:** Azure Databricks, snowflake, redshift
- Data-Visualization/Reporting skills:** QlikView, Tableau, Microsoft Power BI, SAS Data Miner, Excel, PowerPoint
- Big Data:** Databricks, Apache Spark, ETL, Data pipelines, Lakehouse, Data modeling and processing, Production, alerting, deployment, testing, security, Apache Kafka, Confluent Kafka, Kerberos.

CERTIFICATIONS

Databricks Certified Data Engineer Professional:

Sep 2024 – Sep 2026