JAIDEEP REDDY ALURU

| Dallas, TX | +1(469)-349-2844 | [jaideepreddyaluru@gmail.com](mailto:jaideepreddyaluru@gmail.com) | https://[www.linkedin.com/in/aluru-jaideep-reddy-](http://www.linkedin.com/in/aluru-jaideep-reddy-) 138564151/

# CAREER OBJECTIVE

A motivated Data Science graduate with strong technical competence and attention to detail seeks an entry-level position in data engineering. Skilled in Python, SQL, and AWS, with a reputation for end-to-end ownership and problem- solving.

# EDUCATION

## UNIVERSITY OF NORTH TEXAS JAN 2022 – MAY 2023

Master of Science in Data Science (GPA**:** 3.7)

**Relevant Courses**: Harvesting and Data Retrieval, Big Data Engineering, Data Analytics, Data mining, Data Modelling, Data-Visualization and Knowledge Discovery.

## JAWAHARLAL NEHRU TECHNOLOGICAL UNIVERSITY HYDERABAD,

**INDIA AUGUST 2017 – JUNE 2021**

Bachelor of Technology in Electronics and Communication

# TECHNICAL SKILLS

AWS, GCP, Jupiter, Linux, MATLAB, Microsoft Office Suite, RapidMiner, SQL, Tableau, Eclipse, PyCharm, Anaconda, IntelliJ.

**Programming Languages:** C, C++, C#, Python, Scala.

**Databases:** MySQL, HBase, No SQL.

**Bigdata Ecosystem:** HDFS, Hive, Spark, Sqoop, EMR, Kafka, Nifi.

**Areas of Expertise**: A/B testing, Data Extraction, Data Governance, Data Preprocessing, Data visualization, Data Wrangling, Exploratory data analysis.

**Project Management:** Agile project management, Leadership, Scrum ceremonies, and Team Management.

# WORK EXPERIENCE

### Jr. Data Engineer , COGNIZANT, INDIA

JAN 2021 – DEC 2021

For development and testing, I set up EC2 instances and EMR clusters.

* Executed step execution in EMR clusters, data load was done from Spark RDDs and data frame into Hive..
* Loaded and transformed large amounts of semi-structured data in formats such as XML, JSON, AVRO, and Parquet. Sqoop was used for imports. Managed Hive Tables and created partition-based child tables.
* Worked with streaming data from Kafka and processed it in Spark.
* Contributed to the delivery of the resulting data to Cassandra.
* Used Scala to process web API data and convert it to data frames for further processing and stored in JSON data for easy access. IT was stored in HDFS/S3 based on business need.

# ACADEMIC PROJECTS

### Project Name | Credit Card Fraud Detection System

* Hive-managed HBase tables were created, Imported, and exported data from AWS-RDS to Hive.
* Sqoop was used to import jobs, Performed Spark-HBase integration for table data.

### History of Modern Olympics | Tableau:

* Created dual-axis advanced charts, workbooks, and dashboards for better insightful understandings and conducted presentations, established technical, and performance progress reports, and storytelling on the data- driven insights to the business partners, program managers, and a large group of stakeholders.

### IoT Based Home Automation | PYTHON C | GCP | JavaScript:

* Used Python to Initialize GPU in the Raspberry pi board. GCP is used as connection between application and Hardware. HTML, CSS, Java Script is used for development of mobile application.

### Smart Parking System | Embedded C:

* + Embedded C is used to initialize ports in Node MCU. GCP is used to send the data from hardware to the cloud.