GANESH PASIGANTI

Denton, TX | ganeshpasiganti@my.unt.edu | (940)-843-3938 | LinkedIn | PGanesh384 (github.com)

EDUCATION

University of North Texas	**Texas, USA**
Master of Science, Advance Data Analytics, (Pursuing 2022 - 2024)	**May 2020**

Courses: Decision Analysis, Visualization for Analytics, Database Management Systems, Applied Predictive Analytics, Deep Learning, Principles Data Management/Mining, Big Data Needs Analysis

Jawaharlal Nehru Technological University	**Hyderabad, India**
Bachelor of Technology, Electrical and Electronics Engineering, GPA 3.5	**Apr 2017**

TECHNICAL SKILLS

- **Hadoop Ecosystem:** Hadoop, Spark, MapReduce, Hive, Sqoop, Airflow, Nifi
- **Languages:** Python, Java, R, SQL, UNIX, Scala
- **Databases:** MySQL, Oracle, PostgreSQL
- **IDE's:** PyCharm, Jupyter Notebook, R-Studio
- **Tools:** Tableau, Power BI, Git, Hortonworks Sandbox, Cloudera Big Data Framework, Docker, SSMS
- **Industry Knowledge:** AWS, Microsoft Azure, Agile, Machine Learning, Predictive Analytics
- **Certifications:** Microsoft Certified: Azure Data Fundamentals, AWS Certified Developer- Associate(In Progress)

PROFESSIONAL EXPERIENCE

Associate Data Engineer(IBM, Karnataka, India)	**Jul 2020 – Jan 2022**

Project: Data Migration Project from Click and SAP source system to Salesforce-Field Lighting Service

- Prototype solutions, prepare SQL test scripts, and conduct tests for data replication, loading, cleansing, and data review, validating data loaded in Field Lighting Service Salesforce application for accuracy
- Experienced in creating and using Stored Procedures, User Defined Functions, Sub - Queries and Joins for complex queries involving multiple tables
- Creating Ad hoc reports involving complex datasets and to query the database for Business Intelligence
- Deploying the transformed data in various environments i.e., DEV/SIT/UAT
- Developing proofs of concept to the client, and evaluate design options to deliver ingestion, metadata cataloging, making recommendations for future improvements
- Communicate with key stakeholders and system integration partners for target data mapping

Assistant Data Engineer(IBM, Hyderabad, India)	**Jan 2018 - Jul 2020**

Project: Creating an Ingestion Framework which can pull data from MySQL database into HDFS using Spark

- Developed an ingestion module to ingest data into HDFS to apply transformation logics
- Multiple tables provided in configuration to specify type of load, used Apache Hive to run Map Reduce jobs
- Managed and scheduled jobs on Hadoop Cluster using Apache Oozie
- Imported RDBMS data into HDFS using Sqoop to run Hive queries on data, worked with ORC, Avro, Parquet, JSON files, worked on various compression techniques to leverage the storage in HDFS
- Generated reports and dashboards from various sources to analyze and inform business decisions
- Built and delivered high quality data architecture and pipelines to support Business Analysts, Data Scientists, and Customer reporting needs, monitored analytical metrics and create automated alerts to trigger business response.

ACADEMIC PROJECTS

Retail Data Pipeline | *Python, Docker, Google Cloud Platform*

- Developed an end-to-end ETL(extract, transform, load) data pipeline for an e-commerce website
- Migrated the transactional data from MySQL database into PostgreSQL database for generating flash reports

Student Housing and Residence Life Database | *Relational Databases, MS-Visio*

- Designed a database system using entity-relational model, normalization, created database objects including tables, constraints and manipulating data
- Applied data query techniques such as row filtering, joins, single-row functions, aggregate functions, sub-queries, and views in Oracle 12c database environment