

PARDHA SARADHI PAMARTHI
+1(940) 977-4206, 1200 Dallas Dr Apt 713, Denton, Tx - 76205
pardhapamarthy@gmail.com ♦ www.linkedin.com/in/pardha-saradhi-pamarthi-60a5a6138

EDUCATION

Master of Science - Data Engineering, University of North Texas **CGPA - 4.0** **January 2023 - May 2024**
Relevant Coursework: D&P databases, Bigdata, Machine Learning, Data Visualization, NLP.
Bachelor of Technology, VIT University **CGPA-3.5** **June 2016 - May 2020**

TECHNICAL SKILLS

Hadoop Eco-system: HDFS, MapReduce, Spark, Hive, Sqoop, Kafka, Zookeeper.
Programming and Query languages: Python, SQL, HiveQL, Pyspark, Scala, PyCharm.
Tools and Methodologies: Databricks, ELT, ETL, Git Version, Tableau, Power BI.
Cloud Platform: AWS EC2, S3, AWS EMR, GCP, Big Query, Data flow.
Relational and NoSQL Databases: Oracle, MySQL, PostgreSQL, HBase, MongoDB, Snowflake.

EXPERIENCE

Data Engineer

June 2021 - December 2022

Cognizant Technology Solutions-- Client BHP Billiton, Hyderabad, India

- Served as the primary point of contact and coached operations team for over 70 applications during the warranty period, ensuring seamless operations and issue resolution, supporting end-users, ensuring effective dashboards.
- Developed a Python script that automated the migration of existing Oracle table definitions to Snowflake, eliminating hours of manual work and increasing efficiency by 60%.
- Built data processing pipelines using Stream Sets, seamlessly integrating API, SFTP flat files, and database sources, enabling real-time data ingestion after communicating with the end users.
- Spearheaded the creation and maintenance of a Git repository for Snowflake schema and table generation, incorporating appropriate confidentiality levels and enabling automated execution through CI/CD pipelines, leveraging Amazon S3 for intermediate staging.

Bigdata Engineer

September 2019 - May 2021

Cognizant Technology Solutions-- Client BHP Billiton, Chennai, India

- Deployed Apache Sqoop to efficiently migrate 1.5TB of data from Hadoop to Oracle databases, achieving a 45 percent increase in data processing speed and improving analytics accuracy for large datasets.
- Managed comprehensive Spark SQL DDL processes, involving the creation, modification, and removal of tables, views, and indexes; optimized data structure and improved query performance by 22%.
- Implemented robust data ingestion and transformation workflows, adeptly utilizing Spark Data Frames and RDDs to process and manipulate data from heterogeneous sources, including text, JSON, XML, AVRO, ORC, and PARQUET files.
- Engineered robust Spark Resilient Distributed Datasets (RDDs) transformations and actions, including filtering, mapping, reducing, grouping, and aggregating data.
- Completed comprehensive training in big data technologies, including Hadoop, Hive, Spark, and Kafka, enhancing skills in large-scale data processing and analytics.

PROJECTS

Machine Learning Approach to Predicting and Managing Flight Delays

- Modeled a logistic regression model to predict flight delays with 82% accuracy, enabling decision making for airlines.
- Delivered actionable insights to improve operational efficiency and enhance passenger satisfaction.

Pollution Impact Analysis: Cloud-Based Data Infrastructure Deployment

- Leveraged GCP tools (Big Query, Dataproc, Hadoop, Hive, Spark) to build a data pipeline for analyzing pollution's impact on health.
- Performed comprehensive data cleaning and wrangling using Open Refine and SQL queries to derive actionable insights from disparate data sources.
- Demonstrated Spark's 70% faster performance over Hive for large-scale data processing, optimizing analysis pipelines.

ACHIVEMENTS:

Paper published on "Pet and Welfare services"- <https://www.ijariit.com/manuscript/pet-and-animal-welfare-services/>