

BHARATH K

+1 (716) 939-6952 | bharathk9927@gmail.com | [LinkedIn](#) | New York , US

EDUCATION

University at Buffalo, New York, US

December 2024

Masters in professional studies in Data science

Coursework: Database Management Systems, R, Python, Probability & Data Analysis, Statistics, Data Mining, Machine Learning.

Chennai Institute of Technology, Anna University, Chennai, India

April 2021

Bachelor Of Engineering

PROFESSIONAL EXPERIENCE

Softtek

Bangalore , India

Data Engineer

November 2021 – August 2023

- Designed, developed, and maintained over 15 data pipelines using Spark, **Databricks**, **Informatica**, and Teradata, efficiently processing 200 million records daily from S3 to Teradata.
- Spearheaded the migration of Teradata SQL pipelines to Databricks Spark, achieving an annual **cost reduction of \$350k** and enhancing overall pipeline performance.
- Optimized existing pipelines by refining cluster usage and applying advanced Spark techniques, leading to a 50% improvement in CPU utilization and reduced compute expenses.
- Orchestrated 100+ Spark jobs with **Apache** Airflow and **Databricks workflows**, managing terabytes of data daily with high efficiency.
- Implemented **Jenkins** pipelines for **CI/CD** processes, automating code deployment and streamlining workflow orchestration. Automated data testing processes and contributed Python scripts for diverse data operations

Data Engineer Intern , Softtek India

December 2020 – October 2021

- Assisted in the creation and maintenance of **Spark** data pipelines, including the development of **pipelines** to transfer data from S3 buckets (JSON files) to **Teradata tables**.
- Contributed to the **migration** of **Teradata SQL pipelines** to Databricks Spark by performing data validation, troubleshooting, and optimizing initial pipeline configurations.
- Supported performance tuning of data pipelines and automated daily data testing processes, including building and validating testing queries and notifications.
- Developed Python scripts for automating routine tasks such as data copying, **DDL & SQL** generation, and dynamic data transformations, enhancing overall workflow efficiency.

Projects :

Stock Price Prediction Using Machine Learning

- Developed a **machine learning** model to forecast future stock prices based on historical data and financial indicators. Employed data collection through financial APIs, performed extensive data preprocessing and feature engineering, and experimented with multiple models including **ARIMA**, **Random Forests**, **Gradient Boosting Machines**, and **LSTM networks**.
- Developed interactive dashboards in **Tableau** to visualize predictive analytics, enabling intuitive interpretation and strategic decision-making.

Customer Experience and Labor Cost Optimization : PowerBI, Python, SQL

- Defined and tracked KPIs such as "Average Customer Wait Time," "Sales per Hour," and "Labor Cost as a Percentage of Revenue" to measure project success. Regularly monitored these KPIs to ensure that customer experience and cost control objectives were met .

Skills

Programming: Python ,SQL, R Programming

- **Libraries:** NumPy, Pandas, Matplotlib, Scikit-learn, PyTorch

Database Management : Teradata , MySQL.

- **Tools and Frameworks:** ETL , Informatica, Microsoft Power BI, Data visualization, Tableau, Machine Learning , Microsoft Excel, JIRA, Jenkins , Appworx , Autosys , Putty , WinSCP , Toad for SAP , GIT, ETL Pipeline , Data Mining , Data Bricks ,Data cleaning, Datasets , Data Analytics , Service now .