Naresh Pochaveni

nareshya@buffalo.edu | Buffalo, NY | 716-541-5411 | linkedin.com/nareshpochaveni | github.com/pochaveni

PROFESSIONAL EXPERIENCE

Rentastico OPC Pvt Ltd, India *(Data Engineer II)* **May 2020 - Jul 2021**

- Ensured the accuracy, completeness, and security of data. I created and maintained data quality checks to ensure data integrity, and I was responsible for managing access to sensitive data, ensuring that only authorized personnel had access to it.
- Owned and maintained a robust data pipeline to ensure daily accuracy and consistency within the database.
- Implemented optimizations that reduced data processing time, resulting in a more efficient pipeline.
- Played a key role in enhancing the Extract, Transform, Load (ETL) process, resulting in improved data quality and reliability.
- Continually improved ETL workflows to accommodate evolving business needs and larger datasets.
- Collaborated with data analysts to identify and resolve data inaccuracies for internal and external reporting purposes.
- adept at designing and developing SQL databases, monitoring and optimizing system performance, and ensuring data accuracy and security.
- Successfully developed and optimized Extract, Transform, Load (ETL) processes, leveraging AWS Glue for automated data extraction, transformation, and loading tasks, reducing manual intervention and enhancing efficiency.
- Ability to work both independently and as part of a team, fostering a collaborative work environment that promotes knowledge sharing and innovation
- Monitored database performance, identified potential problems, and made recommendations to optimize system performance.

Tata Consultancy Services, India *(Data Engineer)* **June 2018 - Apr 2020**

- Designed and developed SQL databases to ensure efficient data storage and retrieval, monitoring database performance, and optimizing system performance.
- Successfully designed, implemented, and maintained scalable data pipelines to process and ingest large volumes of data, reducing data processing time by 30%.
- designed, implemented, and maintained scalable data pipelines and solutions using AWS services such as S3, Glue, Lambda, and EMR.
- Developed robust ETL (Extract, Transform, Load) workflows using Apache Spark, Apache Airflow, and Python, ensuring data accuracy and consistency.
- Optimized data storage and retrieval systems, reducing storage costs by 25% through the implementation of data partitioning and compression techniques.
- Built efficient data pipelines that extracted information from diverse sources and seamlessly integrated it into a secure, cloud-based data analytics platform.
- Leveraged SQL for data querying and transformation, optimizing data retrieval processes.
- Implemented data quality monitoring and validation processes, identifying and rectifying data anomalies, leading to improved data integrity and decision-making.
- Led the integration of various data sources, including databases, APIs, and streaming data, resulting in a unified data repository for analytics and reporting purposes.
- Contributed for creating visually appealing dashboards and reports using Power BI, working closely with business analysts to understand their requirements and translating them into actionable insights through data visualization.
- Detail-oriented and thorough approach to work, ensuring accuracy and precision in data analysis and reporting.
- I provided technical support to users, identified and resolved technical issues, and created technical documentation to help users understand how to use the software.

PROJECTS

Customer Segmentation using RFM Analysis (Python*, NumPy, Matplotlib, Seaborne*) **Aug 2021 - Dec 2021**

- To solve the issue of most e-commerce enterprises missing out on potential growth due to a lack of understanding of their target client and their preferences, created a model that segments online customers into clusters based on how frequently they buy products when the last time they made a purchase and how much they spent in total.
- It helps in targeted advertising, reducing marketing costs and time.

Binary Classification Model for Conversion Prediction *(R, data table, ggplot2, caret)* **Aug 2021 - Dec 2021**

- Analysed data from Google Analytics and publicly available session data to forecast consumer intent on e-commerce websites to predict if a visitor will make a purchase.
- The discovery process allows e-commerce businesses to present relevant content to customers who have indicated a positive buying intent, which in turn allows them to attract more customers and increase revenue.

Ticketing Data Analysis of City Buses (R, IBM infosphere) **Dec 2017**

- By analysing the ticketing data of city buses, I was able to reduce the frequency of buses in less crowded areas and redirect them to areas that require a higher number of buses, resulting in optimal resource utilization for the benefit of the people and increased revenue.
- Project received appreciation by the project coordinator for the practical approach towards a real-world problem.

AWARDS & INVOLVEMENT

Vice President - Education, TCS Toastmasters Club, Nagpur, India **April 2019 – Dec 2019**

- Active in assigning meeting roles and in setting club agendas and goals and worked towards it.
- Worked as a primary Base camp Manager to monitor members to achieve their pathway goals.
- Managed club schedule planning and club meetings.

EDUCATION

State University of New York, Buffalo, USA **Feb 2023**
Master of Science in Data Science *3.53 GPA*
Relevant Courses: Data Intensive Computing, Data Modelling Query Language, Intro to Machine Learning

Gokaraju Rangaraju Institute of Engineering and Technology, India **May 2018**
Bachelor of Technology in Computer Science and Engineering

SKILLS

Programming Languages: Python, R

Big Data Technologies: Hadoop, Sqoop, Spark

Data Visualisation Tools: Tableau

Database and ETL Technologies: PostgreSQL, SQL

Web Technologies: HTML5, CSS, JavaScript, Angular

Other: AWS glue, S3, EMR, Athena, Lambda, Step Functions, JIRA, Agile Development, GitHub