

Neeraj Reddy Bandi

(585) 290-2136 | nb3938@g.rit.edu | [Github](#) | [LinkedIn](#)

Education

Rochester Institute of Technology, Rochester, NY

Master of Science in Information Technology Analytics, GPA 3.56/4.0

New York

Jan 2022 - May 2024

National Institute of Technology Jamshedpur

B.Tech. in Materials Engineering, GPA 7.28/10.0

India

August 2014 - May 2018

Technical Skills

Programming: Python, Java, SQL, Apache Spark, Hadoop, Hive, MongoDB, XML, R, Postgresql

Big Data: Airflow, Docker, Kafka, AWS, Azure, Snowflake, Kubernetes, Hadoop, Terraform, Pandas, Shell Scripts, Scala

Developer Tools: SAS, Tableau, PowerBI, qGIS, Plotly Dash, shiny, Streamlit

Certifications

- **Azure Data Engineer** associate certification.
- Databricks Certified Associate Developer for **Apache Spark**
- Certified **Scrum Master** with Proficiency in **Agile** Methodologies.

Professional Experience

Verticle Technologies - Data Engineer

Hyderabad, India (Aug 2020 - April 2022)

Tech Stack: Spark, Azure Data Factory, Azure DataBricks, Azure Devops, kafka, SQL, NoSQL, CI/CD

- Created automatic deployment infrastructure on the Azure cloud platform to increase code release frequency by **400%** by performing build, test, and deployment of artifacts in development, test, and production environments.
- Designed efficient ETL data pipelines to extract real time 20 million ecommerce events data using kafka and transformed using spark, hive tables reducing 30% latency and a 7% increase in fraud decision making for IGCC.
- Optimized Spark queries for processing 1TB daily data, enhancing business-critical workload efficiency and refund decision-making. Reduced query runtimes by 83%, from 30 minutes to under 5 minutes.
- Migrated from RDBMS to NoSQL denormalized schema, created partitions and indexes, reducing query latency by **54%**. Engaged proactively in transitioning to an updated data warehousing approach, gaining proficiency in Star Schema modeling to optimize data storage efficiency.

Volvo Trucks- Product Analyst

Bangalore, India (Jun 2018 - July 2019)

Tech Stack: SAS, PowerBI, Python, JIRA, SQL, ETL

- Built machine learning models using SAS to identify influential factors in purchasing a new vehicle across three customer segments. Achieved a 20% increase in targeting precision for segmented high-potential customers, enhancing targeted marketing strategies.
- Designed interactive dashboards and reports, backed by complex DAX formulae, to deliver actionable insights and support data driven decisions.
- Led successful Indian launch of VECTO, collaborating with a 19-person team to validate, customize, and add 16 new features.

Academic Projects

- **Enhanced LLM Application with Real Time Data Pipeline for recalled products:** Designed and built a real-time data pipeline using Kafka, Spark, and Airflow for streaming, processing, and orchestration. Leveraged Docker for containerization and PostgreSQL for data storage, ensuring scalability and reliability. Used GPT-4 LangChain framework which uses ReACT agent to generate reasonable queries from database by incorporating memory capabilities and deploying via Streamlit.
- **Database Optimization and Data Visualization:** Crafted a MySQL-driven emulation of Instagram's database, integrating statistical regression and clustering models for predictive analytics. Utilized Tableau for visual data representation, elevating actionable insights through mathematical analysis.
- **Dominick's fast food data warehousing and visualization:** Developed data warehouses using Snowflake by extracting supply chain, designing a STAR schema, performing ETL, and creating denormalized data marts and data lakes to indicate inventory flaws through visualization in PowerBI.