RAJ SARODE

Boston, MA | 617-653-0918 | sarode.r@northeastern.edu | LinkedIn | GitHub | Portfolio

EDUCATION

Master of Science in Information Systems (GPA: 3.82) | Northeastern University *Sept 2022 - May 2024*
Coursework: Designing Data Architecture for Business Intelligence, Data Management and Database Design *Boston, MA*

Bachelor of Engineering in Computer Engineering | University of Mumbai *Aug 2018 - May 2022*
Coursework: Data Warehouse and Mining, Big Data Analytics, Database Management Systems *Mumbai, India*

TECHNICAL SKILLS

Programming Languages: Python, R, SQL, JavaScript
ETL and Business Intelligence Tools: Power BI, Tableau, Alteryx, Talend, ER Studio, Navicat
Database Systems: Microsoft SQL Server, PostgreSQL, NoSQL, MongoDB, Spark SQL
Cloud Platforms: Amazon Web Services (AWS), Microsoft Azure, Snowflake
Publications: Image Super-Resolution using Enhanced Super Resolution Generative Adversarial Network, ICACC, 2022
Other Skills: Streamlit, API Integrations, Rest APIs, Docker, Microsoft Excel, Statistical Analysis, Data Modeling, Git, GitHub

WORK EXPERIENCE

Northeastern University | Graduate Teaching Assistant | Boston, MA *Jan 2024 - Present*
- Instructed 200+ students in **"Prompt Engineering for Generative AI"** and **"Data Science Engineering Methods and Tools"** courses, emphasizing actionable methods to improve prompt quality and project performance
- Assisted the instructor in translating academic concepts into hands-on exercises, significantly boosting students proficiency in leveraging generative AI for real-world applications

Wizcom Tech | Data Engineer Intern | Mumbai, India *Feb 2022 - Jul 2022*
- Utilized **SQL**, **NumPy**, **Pandas**, and **Amazon S3** for streamlined data cleaning, achieving a 20% error reduction and enhanced forecasting accuracy
- Leveraged **Snowflake** utilities and **SnowSQL** techniques for efficient data processing, simplifying navigation and reporting of complex financial data
- Streamlined **ETL** processes by developing **Alteryx** workflows and leveraging **AWS Glue**, reducing manual operations by up to 35%
- Crafted financial data-driven visualizations in **Tableau**, fine-tuning ad spend based on financial metrics and achieving a significant boost in marketing campaign ROI

Wizcom Tech | Data Operations Trainee | Mumbai, India *May 2021 - Aug 2021*
- Optimized insurance sector client's data analysis by leveraging **Spark SQL** in **Azure Databricks**, improving data processing efficiency by 30% through advanced scripting and model adjustments
- Deployed real-time transaction ingestion and historical data pipelines in **Azure Data Factory**, enhancing fraud detection by integrating external databases
- Developed **T-SQL** database objects in **Azure Data Warehouse** to support **ETL** processes, optimizing data flow in OLTP and relational environments for enhanced performance
- Generated comparative **Power BI** charts, illustrating variations in annual premiums and claims across diverse insurance products

PROJECTS

Web Wise - RAG Chatbot (Python, Streamlit, Chroma DB, LangChain, Flask, Beautiful Soup) |
- Orchestrated the integration of a Retrieval-Augmented Generation (RAG) chatbot, leveraging **LangChain** for website interaction, ensuring the system's capability to parse and utilize web content effectively
- Employed **Python** to develop a data processing pipeline, integrating web scraping, text preprocessing, vectorization, and **Chroma DB** for embedding storage, significantly enhancing the chatbot's performance and response relevance
- Designed an accessible chatbot interface with **Streamlit**, improving the overall user experience by making interactions engaging

Collision Scope: NYC Traffic Dynamics (Azure Data Studio, Talend, Navicat, Power BI, Tableau) |
- Analyzed the robust NYC-MVC dataset with **Azure Data Studio**, executing **SQL** queries to extract valuable insights through data validation and cleaning techniques
- Enhanced NYC traffic data reliability and integration by 35% through **Talend Open Studio** ETL processes, optimizing the outcomes
- Architected a cutting-edge dashboard through **Tableau** and Power BI, seamlessly integrating data from Azure Data Studio to display collision trends in high-risk areas

IOWA Liquor Sales Analysis (ER Studio, Alteryx, Talend, Tableau, Power BI) |
- Conducted trend analysis on 25M Iowa liquor sales records with **ER Studio** for dimensional data modeling, identifying key patterns
- Improved data quality and consistency in Iowa Liquor Sales Analysis using **Alteryx** for ETL staging, modeling, and integration
- Created impactful visualizations with **Power BI** and Tableau, resulting in efficient drill down for specific data points to gain a deeper understanding of trends