Sanchana Mohankumar

📞 +1(206)209-9901| ✉ mohankumar.s@northeastern.edu

🔗 www.linkedin.com/in/sanchanamohankumar | ⌂ https://github.com/Sanchana1997

Experienced software engineer with 2 years in data platforms, backed by a **master's in Data Analytics Engineering**. Currently excelling as a **Data Scientist**, I am fueled by the quest for data-driven insights. Acknowledged for my adeptness in automating complex processes and delivering data-powered solutions that fuel success.

EDUCATION

Master of Science, Data Analytics Engineering   **GPA: 3.96**

Northeastern University, WA, USA   **May 2023**

- **Related Courses:** Data Visualization, Big Data, Machine Learning, NLP, Database Management, Algorithms

B. Tech, Electronics and Communication Engineering

Amrita University, TN, India   May 2019

WORK EXPERIENCE

Quadrant, Data Scientist   **August 2023 – Present**

- Designed and implemented an automated resume processing system using **LLM API**, effectively extracting key insights from resumes and generating structured JSON outputs.
- Developed a resume matching tool utilizing word embedding and similarity metrics, simplifying candidate selection process for recruiters by identifying most relevant resumes.

Northeastern University, Graduate Teaching Assistant   **January 2023 - April 2023**

- Demonstrated technical expertise in **python, statistical analysis, and machine learning** while conducting lab sessions, grading assignments, and assisting students with data mining course work.

Accenture, Software Engineer – Data Platform   **June 2019 - January 2021**

- Extracted data from MSSQL server with SAP data services **ETL tool** leveraging SQL queries to load data into SAP software after transforming data for validating cross functional **ERP system** such as inventory and production.
- Applied technical skills using SQL, Python in data analysis and reporting to provide solutions to client requests.
- Collaborated with **20-member cross-functional (sales, warehouse, material) team** using **Agile methodologies** to adeptly coordinate with **3rd party clients** and successfully launch **SAP ERP** system across 3 countries concurrently.
- Applied intermediate Excel skills in analyzing and exploring data for testing ERP system test scenario.

SKILLS

- **Language and Database:** SQL, MySQL MS SQL Server, PostgreSQL, Python, R, MongoDB
- **Tools and Software** : Tableau, Power BI, MS Excel, AWS S3, AWS EMR, Hive, Hadoop, Spark
- **Libraries** : Pandas, NumPy, Scikit-learn, Matplotlib, ggplot2, tidyverse, dplyr, keras, Shiny, plotly
- **Data Processing and Algorithm:** Text Analysis, Regression, Classification, Clustering

---

PROJECTS

Weather Forecast | *Big Data Parallel Processing using AWS* | SELF PROJECT

- Designed and executed an efficient ETL process using **AWS Glue** to transform and load temperature data for use in **ARIMA models** for weather forecasting and showcased output in **AWS QuickSight**.
- Leveraged **AWS S3** for secure and scalable storage of temperature data and model outputs, ensuring easy accessibility and efficient data management.

Meal Kit Delivery Database | *Database Design and Management using SQL*

- Build a MySQL meal kit delivery database system using 3rd Normal form modelling technique with views, constraints and computed columns to provide seamless end to end meal kit experience for users.
- Created **Tableau Dashboards** to execute user geographical concentration trend analysis further expand meal kit delivery geographically.

Global Issues | *Data Visualization using R*

- Extracted **10+ Global Issue datasets** containing over 1 million data to analyze trends and patterns on major global issues such as poverty, and food security through interactive graphs leveraging RShiny webpage.
- Performed **data wrangling and data cleaning** techniques on raw data set to retrieve informative data for data analysis.

PUBLICATION| A wearable device to detect blood volume changes| **Link**- https://ieeexplore.ieee.org/document/8728520

- Designed and developed a wearable device to detect changes in blood volume using Pulse Rate Variability (PRV) as a response to the increase in cardiovascular cases in the medical field.
- Developed a **MySQL database** to store the data gathered from human subjects and used an internet web host platform to securely access and manipulate the data.