Aditya Reddy Satti

Piscataway, NJ (Open to Relocation)
+1 551-344-7593 adityareddy1977@gmail.com aditya-reddy-satti Aditya-Reddy

Education

Stevens Institute of Technology **September 2022 – May 2024**
Master's degree in Data Science, CGPA:3.7 *Hoboken, NJ*

Vellore Institute of Technology **July 2018 – June 2022**
Bachelor's degree in Computer Science, CGPA:8.03 *Vellore, India*

Work Experience

InsightsGen **January 2024 – May 2024**
Data Scientist Intern *Jersey City, NJ*

- Modeled a serverless **AI-driven** Chat application for an educational platform, leveraging **Azure functions** with **Azure OpenAI**, resulting in a 30% reduction in server management.
- Engineered a conversation memory system using **LangChain** to integrate with **Large Language Models (LLMs)**, leveraging **Azure CosmosDB** for efficient storage and retrieval of user interactions and conversation summaries, resulting in a 15% improvement in system responsiveness and accuracy.
- Spearheaded front-end development using **JavaScript**, creating an intuitive user interface that resulted in a 10% uptick in user engagement and streamlined integration with back-end services for enhanced functionality.

Bluebonnet Data **January 2023 – August 2023**
Data Analyst Intern *Minneapolis, MN*

- Developed **Tableau** dashboards integrating trends and patterns from behavioral analysis, providing stakeholders with real-time insights and improving the accuracy of decision-making by 30%, streamlining campaign strategies.
- Engineered automated data pipelines using **Databricks** to **extract, transform and load (ETL)** USA census data with Minneapolis' 159 precincts voter data for voter behavior analysis, resulting in 70% reduction in data processing time.
- Utilized **Geopandas** for geospatial analysis, creating plots to visualize voter sentiment metrics in various precincts. Analyzed voter behavior patterns and preferences, optimizing campaign strategies.
- Employed **K-means clustering** for cluster analysis to identify five distinct behavior groups, effectively categorizing voter demographics, optimizing outreach methods, resulting in a 20% increase in voter engagement effectiveness.

Arachnomesh Technologies **Jan 2021 – May 2022**
Data Analyst *Hyderabad, India*

- Analyzed healthcare data from 15+ sources, such as electronic health records (EHR), patient survey, insurance claims data, to identify trends and patterns, reducing costs by 25% while enhancing patient care outcomes.
- Conducted in-depth **exploratory data analysis (EDA)** using **Python**, identifying critical inefficiencies in medical devices, resulting in a 25% boost in device utilization and optimizing healthcare delivery.
- Designed and maintained dynamic **Power BI** dashboards, visualizing key performance indicators (KPIs) and real-time healthcare metrics, resulting in a 20% improvement in decision-making efficiency and enabling faster, data-driven insights.
- Transformed legacy reports into modern, scalable **Power-BI** dashboards connected to **SQL-server**, improving data visualization and reduce manual reporting by 40%.
- Automated ETL workflows with **Snowflake** and **Alteryx**, ensuring consistent, accurate data processing and improving validation accuracy by 15%.

Academic Projects

Mental Health Analysis on Social Media Posts | Github **September 2023 – December 2023**

- Applied **NLP** techniques and deep learning models to analyze and cluster Reddit posts, predicting suicide risk at 80% accuracy using feature extraction methods such as **TF-IDF** and **Word2Vec**.

Time Series Modeling and Forecasting | Github **October 2023 – December 2023**

- Employed **Time Series Analysis (SARIMA, GARCH)** to forecast seasonal (monthly retail sales) and non-seasonal (JP Morgan stock price) data, achieving 92% accuracy in seasonal forecasts and a 90% improvement in non-seasonal predictions, with visualizations leading to a 15% reduction in decision-making time.

Twitter (X) Search Application | Github **March 2023 – May 2023**

- Developed a Twitter Search Application using **Python** and **Flask**, integrating **MySQL and MongoDB** for data storage and **Redis** for caching, which amounted to an enhanced retrieval efficiency.

ML Driven Diabetes Prediction | Github **October 2022 – December 2022**

- **Predicted** diabetes with **BRFSS** data by preprocessing and employing **Machine Learning** techniques such as **Logistic Regression** and **XGBoost**, achieving the highest F1-Macro Score of 0.47 with **Random Forest** and top accuracy of **75%** using XgBoost.

Technical Skills

Languages: Python, SQL, R, PostgreSQL

Frameworks/Libraries: Tensorflow, Keras, PyTorch, Pandas, Numpy, Scikit-Learn

Data Analysis and Visualization: Tableau, PowerBI, MS Excel, Matplotlib, Seaborn, Plotly, Alteryx

Machine learning: Linear Regression, Decision Tree, Random Forest, XGBosst, Neural Network, A/B testing, Data Science pipeline (cleaning, wrangling, visualization, modeling, interpretation).

Cloud/Development Stack: Amazon Web Services (AWS S3 bucket, EC2 Instance), Snowflake ,Docker, Git, Jupyter Notebook.