Jayanth Dasamantharao

New Brunswick, NJ ✉ dasamantharao.jayanth@rutgers.edu ☎ +1(848)3135617 in jayanth ⬡ jayanth ➤ portfolio

Education

Masters in Data Science, *Rutgers University*   2022 – 2023 | New Brunswick, USA
Coursework : Data Mining, Statistical Modeling & Computing, Statistical Learning.

Bachelors in Electrical and Electronics Engineering,   2017 – 2021 | Visakhapatnam, India
Anil Neerukonda Institute of technology and Sciences
Coursework : Data Structures, Computer Architecture, Matlab/Simulink, C++.

Skills

Programming & Python Libraries:
R, C++, Scala, Python : Matplotlib, Numpy, Pandas, Scikit-learn, Seaborn, NLTK, Tensor Flow.

Database Management:
Postgres, MongoDb, Oracle Db, SQL, MySQL, SQLite, RDBMS, Data Warehousing.

Cloud Services:
AWS Athena, Amazon Lambda, Redshift, Amazon S3, RDS Microsoft Azure, GCP Big Query, Cloud SQL.

Data Science & Analytcis:
Tableau, Data Wrangling, Machine Learning, Regression and Time Series Analysis, Statistical Inference.

Professional Experience

Harvest Software Solutions, *Intern - ML Engineer*   06/2023 – 09/2023 | Florida, United States
- Contributed to the development and deployment of machine learning models utilizing Python, R, and other programming languages.
- Analyzed data statistically to uncover trends and patterns. Collaborated with diverse teams to identify and prioritize data science projects.
- Preprocessed and transformed extensive datasets for analysis. Crafted data visualizations and dashboards to convey insights to stakeholders.
- Generated and managed documentation for data science projects. Assisted the data science team in various ad-hoc projects as required.

Accenture, *Database Analyst, CDW admin with Regeneron.* ✎   2021 – 2022 | Hyderabad, India
Data Analysis
- Identified two-third of the integrity exceptions using **Athena, Redshift** by utilizing data analysis to monitor process efficiencies.
- Reduced up-to 50% of support calls by designing and implementing data pipelines using **python** pandas for storage and maintenance of customer campaigns.
- Lead a team of 3 in updating the dashboards using **Tableau** for the User Interface, resulting in a 50% improvement in data visualization and a two-third increase in user engagement.
- Developed and integrated machine learning models using **TensorFlow and PyTorch** libraries, resulting in the improvement in user retention.

Customer Data Warehousing
- Conducted business analysis for a data warehouse solution, resulting in the improvement of data accuracy and reduction in data processing time. Lead data modeling and testing of program modules to ensure compliance with design specifications, resulting in a decline of decrease in defects and a major increase in program efficiency.
- Executed **SQL queries using SQL Workbench** to fulfill business needs and deployments, leading to an enhancement in data quality.
- Created detailed reports and closed incidents promptly, resulting the downtime and increase in efficiency.
- Delivered technology solutions comprising a team of 3, to support user needs by performing DDL updates for the required deployments.

Tools Used: Python, Pandas, AWS Services - Athena, Redshift, S3, Lambda, RDS SQL Workbench, Tableau, Tensor Flow, PyTorch

Projects

Auto Insurance Fraud Detection, *Rutgers University* ✎   2023
- Developed machine learning model using Random Forest, KNN, Logistic Regression, SVM, Decision Tree and XG Boost to accurately detect auto insurance fraud and classify fraudulent claims without prior knowledge of the specific type of fraud, helping insurance companies avoid significant financial losses.
- Multiple models were tested on a balanced dataset to improve recall levels for identifying instances of auto insurance fraud. The Random Forest and KNN algorithms achieved an f1-score of 91%, with KNN having a recall rate of 97%.

Bayesian Classification on Diabetes dataset, *Rutgers University* ✎   2023
- This project uses Bayesian logistic regression to predict whether a patient is positive or negative for diabetes, based on input variables such as insulin levels and age.
- PyMC3 is used to apply prior distributions, and various prior distributions are experimented with to compare the results to standard logistic regression techniques. Calculated the recall, f1 score by employing a logistic regression model, logistic regression model with uniform distribution, normal distribution as its prior.

Twitter Search Application, *Rutgers University* ✎   2023
- This project centers around creating a search engine for retrieving tweets based on user input. It employs both relational and non-relational datastores - Postgres for user data and MongoDB for tweets - to ensure efficient storage and access. The search engine employs caching techniques for improved performance and aims to provide optimal tweet retrieval and a seamless user experience.