Jayanth Dasamantharao

📍 New Brunswick, NJ - 08901 📞 +1 (848)-313-5617 ✉ dasamantharao.jayanth@rutgers.edu

🔗 linkedin.com/in/JayanthDasamantharao ⦿ github.com/JayanthDasamantharao

EDUCATION

Rutgers University, *Masters in Data Science*   *Sept 2022 - May 2024 | New Brunswick, NJ*

Relevant Coursework: Statistical Learning, Statistical Modeling and Computing, Neural Networks, Machine Learning (math-oriented), Data Structures and Algorithms, Data Wrangling (R), NLP, Data Mining. Projects Overview: Portfolio.

Andhra University   *2017 - 2021 | Visakhapatnam, India*

Bachelor of Technology in Electrical and Electronics Engineering

Research Paper: Self Driving Cars Using Image Processing and Deep Learning.

TECHNICAL SKILLS

Programming: R, C++, Scala, SAS, **Tools:** TensorFlow, Pytorch, Numpy, Pandas, Scikit-learn, Matplotlib, Seaborn.

Data Engineering: Data Warehousing, Postgres, MongoDb, Oracle Db, SQL, MySQL, SQLite, PySpark, Hadoop.

Deep Learning and Machine Learning(ML): Regression, Clustering, Random Forests, NLP - Nltk, Keras, Hugging Face Transformers, spaCy, LLM's, NER, Computer Vision, Recommender Systems, Language Models, CNN, RNN, Feedforward, Transformers, LSTM, BERT, GPT.

Cloud Services and Data Analysis: AWS Athena, Amazon Lambda, Redshift, Amazon S3, RDS, Data Wrangling, EDA, PowerBI, Tableau, RedCap, SAS - statistical tests using Proc ANOVA, Proc GLM, Pearson correlation and others. SAS macros, Proc SQL for data retrieval.

PROFESSIONAL EXPERIENCE

Harvest Software Solutions, LLC, *Data Science Intern*   *June 2023 - Sept 2023 | Jacksonville, Florida*

Real-time Ad Campaign Optimization:

- Leveraged data-driven insights and state-of-the-art machine learning techniques to enhance the effectiveness of ad campaigns and drive increased revenue. Acquired and extracted data from Aniview and CMCBeacons using APIs.
- Handled datasets with over 5 million rows, involving EDA to uncover patterns, relationships within the data, data cleaning, outlier handling, data format correction. Calculated CPM and added it as a critical metric to the dataset.
- Employed data visualization, summary statistics, and hypothesis testing to gain insights into ad campaign performance.
- Utilized XGBoost for predictive modeling and achieved a low Mean Squared Error (MSE). Implemented a 5-fold cross-validation approach to evaluate model performance and avoid overfitting. Leveraged H2O's AutoML to automate model building processes, selecting the best-performing model for CPM prediction.
- Employed AutoML tools to automate labor-intensive tasks, including feature engineering and hyperparameter tuning, selecting the best-performing model for CPM prediction.

Accenture Solutions Pvt. Ltd., *Application Development Associate*   *June 2021 - Aug 2022 | Hyderabad, India*

Database Analyst - Regeneron Inc:

- **Data Analysis:** Leveraged Athena and Redshift for clients data analysis to detect two-thirds of integrity exceptions, enhancing process efficiencies.
- Designed Python pandas data pipelines to automate the script, reducing support calls by 50% and ensuring efficient storage and maintenance of customer campaigns.
- Led a cross-functional team of 3 in Agile development, driving enhancement in data visualization.
- **Customer Data Warehousing:** Performed business analysis, leading to enhanced data warehousing, improved data accuracy, and a substantial reduction in data processing duration.
- Developed SQL queries via SQL Workbench to cater the business needs, thereby boosting data quality and precision.
- Implemented technological solutions, including DDL updates and the deployment of data into production servers.

RESEARCH PROJECTS

Bayesian Classification on Diabetes dataset, *GitHub Repository*   *Rutgers University, 2023*

Implemented Bayesian logistic regression models to predict patients' diabetes status, utilizing the PyMC3 library to apply informative prior distributions. Evaluated model performance using recall and f1-score metrics, comparing outcomes across logistic regression with uniform priors and normal distribution priors.

Twitter Search Application, *GitHub Repository*   *Rutgers University, 2023*

Created a user query-responsive tweet retrieval system by integrating dual databases: Postgres for user data management and MongoDB for tweet storage. Optimized data storage and retrieval efficiency while implementing advanced caching strategies for superior system performance.

Auto Insurance Fraud Detection, *GitHub Repository*   *Rutgers University, 2023*

Implemented ML models, including Random Forests, KNN, Logistic Regression, SVM, Decision Tree, and XGBoost. Due to the importance of minimizing false negatives, KNN was chosen as the best model with a recall rate of 97%.