Rohan Gonjari

Boston, MA | (774) 301-6307 | rohan.gonjari@outlook.com | LinkedIn | GitHub | Research Papers

PROFESSIONAL SUMMARY

Versatile Data Scientist & Machine Learning Engineer with a Master's in Data Science & proven track record of 3 years in professional Data Analytics roles. Specialized in imbalanced classification of Multi-View data & hyper-tuning machine learning models as a MLE. As a Data Analyst, leveraged regression models & tableau dashboards to identify target regions to increase sales. Executed projects encompassing end-to-end implementations of ML pipelines, designing database schemas, conducting data analysis & building interactive visualizations. Contributed to research publications on Multimodal Auditory BCI Classification using Graph CNNs.

TECHNICAL SKILLS

- **Technologies:** Python, MATLAB, R, SQL, MySQL, SAS, Java, JavaScript, HTML, CSS, Tableau, PowerBI, CUDA, Docker, Google Analytics, MS Excel, Linux
- **Libraries**: PyTorch, TensorFlow, Matplotlib, Pandas, Numpy, PySpark, XGBoost, Seaborn, Selenium
- **AWS**: SageMaker, S3, Snowflake

PROFESSIONAL EXPERIENCE

Machine Learning Engineer/Research Assistant | *UMass Dartmouth* | *North Dartmouth, MA* *Aug 2022 – Aug 2023*

- Performed imbalanced classification of **Multi-View** data through feature selection, feature elimination, & data sampling.
- Generated & manipulated graph data for **Graph CNNs** & popular ML models to implement supervised **machine learnin**g.
- Modelled neural networks, manipulating layers & hyperparameters to improve performance for **GCNs**.
- Performed dimensionality reduction **(t-SNE)** to help visualize graph nodes and edges using **Seaborn**.

Data Analyst | *Destek Infosolutions* | *India* *Aug 2020 – July 2022*

- Collaborated with **120+** clients to implement **GA4** via GTM to meet project requirements with a **95% success** rate.
- Initiated a **data sourcing** project, establishing a robust data pipeline for sourcing new data & cleansing old data, then further process data for feature elimination & selection using **NumPy** & **Panda's** libraries.
- Used **regression** models for targeted customer segmentation, resulting in a substantial **18%** sales boost.
- Developed different **Tableau** dashboards to have more visibility of companies' sales portfolio for different products

EDUCATION

University of Massachusetts Dartmouth | *North Dartmouth, MA* *Sept 2021 –Aug 2023*

- Masters of Science in Data Science
- Coursework: High-Performance Scientific Computing, Advanced Data Mining, Advanced Machine Learning, Data Visualization, Data Design & Modeling, Business Analytics, Graph Neural Networks

National Institute of Technology Karnataka (NITK), Surathkal | *India* *July 2016 – June 2020*

- Bachelor of Technology in Electronics & Communications Engineering
- Coursework: Numerical Analysis, Discrete Mathematical Structures, Data Structures & Algorithms, Statistical Analysis.

PROJECTS

Multimodality-Enhanced Graph Generation & Multimodality-Driven GCN | *Master's Thesis* *Aug 2022 – Aug 2023*

- Proposed methods for graph generation & neural networks to work with **Multi-modal** data across various industries.
- The novel GNN implemented in **PyTorch** is able to combine information from multiple modalities to improve BCI-Systems.
- Novel methods showcased an improvement in performance by **16.25% & 21.65%** in two distinct studies.

Hospital Management System (HMS) *Sept 2022 – Dec 2022*

- Built a data architecture to implement a Health Management System. Then generated tables & attributes using **MySQL**.
- Applied **ETL** with **Selenium** to extract medical conditions & prescription data for patients from NHS surveys.
- Transformed prescription string data using **NumPy** & **Pandas** in to be loaded in our HMS database.

Evaluating Medical Condition of Patients *Feb 2022 – April 2022*

- Diagnosed patient health based on predicted health scores using **EDA** and modelling.
- Built **Regression** model with Cross-Validation & Recursive Feature Elimination with significant & engineered features.

Visualizing Olympics Performance | *Link* *Jan 2022 – April 2022*

- Sourced Olympics data between 1960 – 2016 to build dynamic & interactive visualizations using **D3** in **JavaScript**.