NIVEDHA BALAKRISHNAN
• nivedha0702@gmail.com • LinkedIn: nivedhabkr • GitHub: NB • 669-293-4449 • San Jose, CA

EDUCATION

San Jose State University May 2023
Master of Science in Data Analytics (Awarded Academic Scholarship)

Linkoping University, Sweden May 2020
Master of Science (Master Thesis in Data Analytics and Artificial Intelligence)

Anna University, India May 2016
Bachelor of Engineering (Awarded Best Outgoing Student of the Year 2016)

WORK EXPERIENCE

San Jose State University | Graduate Research Assistant (Jan 2022 - Present)

- Performed data collection, data cleaning, feature extraction on the protein sequences from various sources.
- Applied **statistical analysis** and **visualization** techniques using **Python** to gain insights on the dataset.
- Implemented hypothesis test such as **Z test**, **p value** scores to identify significance of features.
- Build Two-Staged ML **Classification** and **Regression** models on **572** extracted features to identify proteins with antithrombotic activity and inhibition constant values respectively.
- Performed Recursive Feature Elimination (**RFE**) and Sequential Forward Selection (**SFS**) methods to identify most important features for both Classification and Regression models respectively.
- Tested the model with **10 million** peptides from protein databases and filtered **573** peptides.

Cognizant Technology Solutions | Data Analyst (Jun 2016 – Jul 2017)

- Developed high performance data pipeline using **MySQL** to organize, cleanse, and normalize the data to generate insights for reporting and AI applications.
- Performed data analysis using **SQL** queries and developed interactive dashboards using **Excel** and **Tableau**.
- Implemented **Predictive Analysis** using Healthcare Insurance data.

TECHNICAL SKILLS

Languages: Python, SQL, MongoDB, R, MATLAB.
Technologies: AWS, Data Mining, Machine Learning, Neural Networks and Deep Learning, NLP.
Software: Jupyter Notebook, Tableau, Power BI,
Databases: MySQL, MongoDB, Neo4j, DynamoDB.
Statistical Techniques: Hypothesis Testing, EDA, Inferential and Predictive analysis, Time Series Analysis and forecasting, Regression Analysis, Sentimental Analysis, and Market Analysis.

PROJECTS

NBA Basketball Data Analysis (GitHub)

- Fabricated the database using **MySQL** & **MongoDB**, compared their performances using **Apache JMeter**.
- Analysed the datasets and developed dashboards using **Tableau** and identified key insights.

Database Management System for Preowned Cars using AWS (GitHub)

- Developed data pipeline using **AWS S3 bucket**, **RDS**, **RedShift** to perform ETL operations.
- Established a connection to **Tableau** and built dashboards to identify key features affecting the car prices.

Investigation of Classification Complexity Algorithms & Identifying Conflicting Hand Movements (Thesis)

- Analysed and identified best features from EMG signals using **permutation feature importance** and **separability index** algorithms and improved the classification accuracy from **92% to 96%** using **MATLAB.**
- Investigated the presence of consistent conflicting neighbours between several hand movements in feature space using data visualization techniques and helped to set up therapeutic procedures based on the results.

Sentimental Analysis on Dating App Reviews using NLP (GitHub)

- Scraped the data from Google App Store using Python and performed data analysis and visualizations to understand the trends and patterns in Hinge data.
- Implemented **Vader Sentimental Analysis** to understand customer's emotions using reviews.

Time Series Forecasting on Weather data (GitHub)

- Implemented statistical analysis and visualizations to understand the weather data.
- Incorporated **Stationary test** using **ADF** and implemented **ARIMA** and **SARIMA** models for forecasting.

CO-CURRICULAR ACTIVITIES

President of the Machine Learning Club at SJSU

- Driving engagement and coordinating events, leading collaborations with other clubs and professors, working with members on projects involving Computer Vision and Natural Language Processing.