SHARMADHA PARTHIBAN

+1(857) 701 7502 | parthiban.s@northeastern.edu | LinkedIn | Github

EDUCATION

Northeastern University, Boston, USA **May 2023**

Master of Science in Information Systems – GPA 3.8

Relevant Courses: Database Design and Management, Computational Visualization, Data Science Engineering Methods and Tools, Data Warehousing and Business Intelligence, Algorithmic Digital Marketing

Rajalakshmi College of Engineering, Anna University, India **April 2020**

Bachelor of Engineering in Computer Science and Engineering – GPA 8.3

SKILLS

Database Technologies : SQL Server, MongoDB, Snowflake, SQL Server Integration Services (SSIS), SQL Server Analysis Services (SSAS)
Data Visualization Tools : Tableau, PowerBI
Data Science Frameworks : Pandas, Numpy, Sklearn, BeautifulSoup, Matplotlib, Seaborn, Keras**,** TensorFlow
Machine Learning : Regression, Classification, Clustering, Neural Networks, Natural Language Processing, Time Series Forecast
Programming Language : Python, R
Other : Git, HTML, CSS, MS Excel, MS Office

PROFESSIONAL EXPERIENCE

Tata Consultancy Services, Banking Financial Service Team – Client: PayPal, India (System Engineer) **Oct 2020 - Aug 2021**

- Created ETL job infrastructure using Informatica and tuned the upsurge performance of ETL jobs by 20% using optimized query
- Analyzed and extracted large datasets with advanced SQL queries, managed production database, resized database space and related transactional logs and documented the confluence
- Generated stored procedures, triggers, packages, PL/SQL tables and tuned existing procedure to reduce overall job time by 10%
- Constructed one to one engagement with key business stakeholders and platform team to understand the requirements and expected functionality for data migration to Google Cloud Platform (GCP)
- Worked independently and demonstrated leadership skill with a team of 8 in planning, scheduling and allocating resources

PROJECTS

[ML] Classification of Ocular Diseases (Python, EDA, SHAP) [GitHub] **Sep 2021**

- Constructed a machine learning model to classify ocular disease of an individual from structured ophthalmic large database of patients consisting a collection of normal and abnormal images of retina
- Trained the model with **Deep Neural Network** using TensorFlow and Keras. Implemeted VGG-16 and VGG -19 models
- Performed exploratory data analysis, augmentation of images, one-hot encoding and image generator techniques to enhance results. Executed hyperparameter tuning for the models and obtained classification accuracy of **89%**

[ML] Prediction of Grocery Store Sales (Python, EDA, Deep Neural Networks) [GitHub] **Oct 2021**

- Determined a machine learning model to predict the sales of the grocery store using **Python** on large dataset. Performed exploratory data analysis (EDA), feature engineering, data wrangling for better insight of the data
- Achieved feature importance for the dataset using **Random Forest Regression** model and **Linear regression model**. Obtained insight of the data through visualization using PyPlot and Seaborn
- Implemented interpretability using **SHAP** and obtained model with highest accuracy of **77%**

[ML] Forecasting sales of Champagne (Python, ARIMA, SARIMA) [GitHub] **Nov 2021**

- Determined a machine learning model to forecast the monthly sales of champagne using Python on large dataset.
- Pre-Processed the dataset and converted non stationary data into stationary data. Implemented Dickey Fuller test
- Plotted the co-relation and auto co-relation charts and visualized the Time Series Data
- Constructed and compared ARIMA and Seasonal ARIMA models. Forecasted future sales of champagne using the model

[ETL and Data Analysis] Customer Review Analysis (NLP, AWS, Tableau) [GitHub] **Nov 2021**

- Scrapped reviews from the Trustpilot site using Beautiful Soup and performed **NLP** pre-processing using stemming, lemmatization, bag of words in the scraped dataset. Performed sentimental analysis to gain insights of the data.
- Obtained batch results via **MTurk** and used **Amazon S3** bucket for data storage. Deployed **Deep Neural Network** and **Random Forest** algorithms for review prediction on the dataset in Jupyter
- Visualized results using **Tableau** dashboards for the analyzed results

[Database and Data Visualization] Restaurant and Food Delivery Database System (SQL, MYSQL, Tableau) **Dec 2021**

- Designed a logical data model using **Lucid Charts** and implemented physical model database architecture on restaurant management system with customers, restaurants and delivery partners considering affecting factors using **SQL Server**
- Reduced data retrieval time by developing optimized complex SQL queries, triggers and stored procedures and developed **Tableau** dashboards to extract business insights