GOPI SELVARAJ

DATA SCIENTIST

Enthusiastic and motivated person with high level of Data Analysis knowledge and an expertise in Machine Learning concepts. Focused individual possessing personal interest in the field of Data Science and Data Engineering. Looking forward to work for an organization where I can contribute directly to the data-to-day growth of the company as well as to build my individual career.

gopiprofessionalmail@gmail.com
+91-7904157464
Chennai, India
https://www.linkedin.com/in/gopi-selvaraj-4b634296/

TECHNICAL SKILLS

PROGRAMMING LANGUAGES:

Python (Libraries: Numpy, Pandas, Matplotlib, Seaborn, Scikit-learn, Statsmodel)

DATABASE MANAGEMENT:

SQL (MYSQL: Joins, Subqueries, Windows function, Views, Normalization)

EXPLORATORY DATA ANALYSIS:

Data sourcing, Data Cleaning,
Data Preprocessing,
Data transformation, Data scaling

DATA VISUALIZATION:

Python Visualization (Univariate, Bivariate & Multivariate analysis)

STATISTICAL ANALYSIS:

Hypothesis formation/testing, Z-test, T-test, Chi-square, ANNOVA, POST HOC Tuckey test

MACHINE LEARNING:

Supervised Learning Regression:
Linear Regression (Best fit line)
Regularization
(Lasso, Ridge, Elastic net)

Supervised Learning Classification:
Logistic Regression,
Naive Bayes
(Gaussian, Bernoulli, Multinomial),
Decision Tree,
Random Forest, Bagging, Stacking
Boosting
(Ada Boost, Gradient Boost, XG Boost)

Clustering algorithms:
K-Means, Hierarchical, DBSCAN

PERSONAL PROJECTS:

- **Second hand car sales Prediction**
  - Built a Supervised Learning Regression model to analyze and predict the cost of a second-hand car sales website.
  - Executed Exploratory Data Analysis (EDA) on the dataset prior to the regression model building and evaluated the model performance with appropriate evaluating measures such as RMSE and r2 scores.
  - Tools Used: Jupyter Notebook, NumPy, Pandas, Matplotlib, Seaborn, Scikit Learn- Regressors.
- **Student performance Prediction**
  - Built a Supervised Learning Regression model to analyze and predict the academic performance of a student based on the past data.
  - Executed Exploratory Data Analysis (EDA) on the dataset prior to the regression model building and evaluated the model performance with appropriate evaluating measures such as RMSE and r2 scores.
  - Tools Used: Jupyter Notebook, NumPy, Pandas, Matplotlib, Seaborn, Scikit Learn- Regressors.
- **Credit card fraud detection**
  - Built a binomial machine learning model to classify transactions or activities into two categories: "fraudulent" and "non- fraudulent," based on historical data.
  - Performed the required EDA pre-processing steps prior to model building and performed Hyper Tuning to derive out the perfect classified outcome during model building process.
  - Evaluated the model performance with appropriate evaluating measure i.e., Recall Score.
  - Tools Used: Jupyter Notebook, NumPy, Pandas, Matplotlib, Seaborn, Scikit Learn – Classifiers.

WORK EXPERIENCE

Senior Systems Programmer– IBM & Kyndryl Limited – Chennai (4.10)

EDUCATION

- POST GRADUATE PROGRAM IN DATA SCIENCE AND ENGINEERING — Year of completion: 2024
  Great Lakes Institute of Management, Chennai
  Certification Link - https://olympus1.mygreatlearning.com/certificate/KGGELRGA
- IBM Professional Data science certificate (Coursera) — Year of completion: 2021
  Certification Link - https://www.credly.com/users/gopi-selvaraj.475df637/badges
- Master of Computer application (MCA) — Year of completion: 2018
  Anna university - Chennai
- Bachelor of Computer Science (BSc. Computer Science) — Year of completion: 2015
  Chennai National College (Affiliated to Madras University) - Chennai