Shubham Shrotri

+1 682-408-0345 | shubham.shrotri@gmail.com | Linkedin | Github | Tableau Profile | Machine Learning Web App

Objective - Innovative and passionate professional, seeking opportunity in field of Data Analysis, Data Science, Data Engineering, Machine learning to demonstrate and enrich technical as well as interpersonal skills.

EDUCATION

Master of Science, Business Analytics - The University of Texas at Dallas, Richardson, TX GPA-3.62 | **May 2019**
Bachelor of Engineering, **Electrical** - University of Pune, Pune, India GPA-3.80 | **May 2016**

TECHNICAL SKILLS

Machine Learning: Regression, Decision Tree, Dimensionality Reduction, k-NN, SVM, Naive Bayes, Deep Learning
Languages: Python (numpy, pandas, scikit-learn, Streamlit, PyCaret, TensorFlow, Keras), PySpark, R, SQL, C
Cloud Technologies: Amazon Web Services
Tools: Power BI, Tableau, Anaconda, Salesforce lightning, R Studio, SAS, Microsoft Excel, Informatica, Cognos
Certifications: AWS Solutions Architect Associate

BUSINESS EXPERIENCE

TaskUs, San Antonio, Texas, US | Business Intelligence Analyst [Python, Power BI,NLP,SQL] February 2019-July 2020

- Designed and developed robust and scalable reporting automation in Power BI to perform various ad-hoc analysis including customer satisfaction analysis, quality of service analysis, contractual KPI driver analysis, and financial performance analysis on diverse data sources
- Presented actionable insights in high level executive review through interactive reports and dashboards built in Power BI to facilitate Operations and Client Services team in making data driven business decisions
- Achieved 20% increase in customer satisfaction rate by performing sentiment analysis on customer feedback text data using Natural Language Processing and Machine Learning in Python and Power BI
- Discovered underlying topics through analyzing semantic structures in text data by applying Latent Dirichlet Allocation algorithm on 15K customer feedback comments using Python libraries NLTK, SpaCy, re
- Automated 30 Power BI reports by building data pipelines designed to extract data from Zendesk API, Talkdesk API, to transform and load data in AWS Redshift using AWS Lambda, SQL queries and Python scripts

Capgemini, Pune, India | Data Analyst [Informatica, SQL, Tableau, python] July 2016 -June 2017

- Executed informatica mappings for ETL operations and masked sensitive information to avoid misuse of data
- Built reports in Tableau to analyze data and drew business insights to improve user experience
- Preprocessed data and performed feature extraction, feature scaling and removed class imbalance in data
- Developed machine learning models in python on historical data to predict fraudulent transactions and performed hyperparameters tuning to evaluate performance of various models on metrics Precision, Recall

KEY ACADEMIC PROJECTS

Real or fake news prediction [Python, NLP, TFIDF, naive bayes, logistic regression, Artificial Neural Network]

- Built and Trained models based on TFIDF matrix to classify 5000 articles using supervised machine learning
- Assessed models on their F1 score and achieved 0.93 F1 score using MLP artificial neural network

Prediction of car mileage [Python, Linear, polynomial, Ridge, LASSO regression, Streamlit, PyCaret] (Link: Web App)

- Performed exploratory data analysis, data cleaning, feature extraction, one hot encoding in Python
- Built polynomial regression model with regularized parameters to predict mileage with r2 score 88.6%
- Built web app through Streamlit and deployed Github repo on Heroku to perform online or batch prediction

Customer churn analysis [Python, Logistic Regression, k-NN, XGB]

- Preprocessed data, developed classification models, compared performances, tuned parameters of XGB classifier to increase F1 score of predicting customer would churn or not from 77.85% to 79.77%

Expansion Analysis Insights [Excel, Tableau]

- Applied data source level filters, performed data blending to develop workbook from multiple data sources
- Assessed different Business environments to suggest region with high turnover, high sales per capita for the expansion of business
- Analyzed competitors' net profit margin in 4 regions using box plot and calculated range, median of net profit
- Created a forecast of sales per capita through 2023 considering seasonal patterns every 12 months