PRASHANT SHUKLA

Boston, MA 02120 | (857)-217-6281 | shukla.pra@northeastern.edu | LinkedIn

EDUCATION

Northeastern University, Boston — April 2021
Master of Science in Analytics with concentration in Statistics — (3.9/4)
Amity University, India — June 2016
Bachelor of Engineering in Electronics and Communication — (7.1/10)

WORK EXPERIENCE

OCTO TELEMATICS NORTH AMERICA - Data Scientist, Boston, USA **February 2021 - July 2021**

- Collaborated with various cross-functional teams and stakeholders to gather business requirements following AGILE methodology
- Applied and validated filtration method (Savitzky–Golay) for smoothing of data which reduced fuel measurement error by 50%
- Optimized window length and degree of polynomial for smoothened Telematics data using Grid Search approach
- Translated business requirements into KPI's to capture the effect of COVID-19 on the mobility in Italy
- Amplified multiple interactive dashboards to analyze 1 million senors data and provided insights to understand the COVID impact
- Integrated and tested two layered Conventional Neural Network (CNN) activation function (ReLu) which reduced error by 25%

INTENSE IP SERVICES – Data Scientist, Gurugram, India **February 2019 - December 2019**

- Extracted patent data from multiple sources, performed data mining techniques to prepare data for further analysis using Python
- Implemented clustering algorithm (K-Means Clustering) on cleaned data to classify patents based on various sectors
- Forecasted monetized value of patents by identifying KPI's and KRI's to avoid litigation in future
- Migrated extracted data to MySQL Server and wrote complex SQL queries, store procedures, joins as per business requirements
- Retrieved data using SQL queries and created interactive dashboards as well as detailed documentation using Tableau

INNOVATIVE SYSTEMS – Data Analyst, Delhi, India **January 2017 - January 2019**

- Proposed end- to - end solution to meet business requirements of sales and marketing teams by identifying KPI's
- Spearheaded project focused understanding customer behavior by applying A/B testing which resulted in increase of 10% sales
- Mentored a team of 3 members to improve technical depth in defining a vision and built a roadmap for sales & marketing services
- Initiated detailed business intelligence reports to capture metrics requested by stakeholders to enable in better decision making
- Developed detailed documentation and procedures for report generation thus enabling smoother transition to the analytics team

TECHNICAL SKILLS

- Languages: Python (Scikit, NumPy, Pandas, Matplotlib, Seaborn, NLTK, Tensorflow, Keras, beautifulsoup), R, HTML, CSS
- Tools: Alteryx, Talend, R Studio, Jupyter Notebook, Microsoft Excel, Microsoft Office (Word, PowerPoint), Databricks
- Database: Postgres SQL, MySQL, Hadoop, AWS S3,
- Big Data: HDFS, PySpark, Spark, Hive, Impala, HBase, Pytorch
- BI & Visualization: Tableau, Power BI, Microsoft Visio, Star UML, Alteryx, IBM SPSS
- Machine Learning: Linear Regression, Lasso and Ridge Regression, Logistics Regression, KNN, Naive Bayes, K-Means Clustering, Decision Tree, Random Forest, Time Series Analysis, A/B Testing, Neural Network (CNN, ANN, DNN)
- Statistics: Descriptive Statistics, Chi-Square Test, Hypothesis Testing, T-Test, ANOVA, Pearson Test, Estimators
- Cloud: AWS (S3, Redshift, Quicksight, RDS, Lambda, Cloudwatch, Cloudformation), VPC, Subnet, Azure, GCP

ACADEMIC PROJECTS

HOUSE PRICE PREDICTION USING MODELING

- Executed EDA by cleaning dataset and plotting correlation matrix to reduce computational complexity using feature selection
- Performed ETL Technique and applied Linear Regression model and built a prediction model with accuracy of 80%
- Applied KNN classification, Random Forest, Decision Tree, and compared algorithms to check model fit properly

HEART DISEASE PREDICTION

- Designed and developed an android application which worked on collecting, extracting real time data from wearable devices & storing it using MySQL as well as monitored patients based on severity and sent alerts in case of emergency
- Predicted severity of heart disease using Decision Tree (83%) and classified patients in four categories based on severity

S&P 500 STOCK PRICE PREDICTION

- Scrapped data using pandas library from multiple data sources and visualized data using matplotlib library
- Applied ARIMA, Facebook's Prophet, staked LSTM, and AutoML to compare and obtain RMSE value of 3.23
- Built predictive model by defining functions based on moving average, exponential smoothing, double exponential, random forest

IMAGE CLASSIFICATION WITH A CONVENTIONAL NEURAL NETWORK

- Classified images with CNN and visualized tuning results of activation function, cost function, network initialization and number of hidden layers with TensorBoard
- Integrated PyTorch network with Leaky – ReLu, and Adam Optimization by applying 4 hidden layers to get best results