

# JWALIT SHAH

(732)-209-2041 | [shah.jwa@northeastern.edu](mailto:shah.jwa@northeastern.edu) | [Linkedin](#) | [Github](#)

## Education

### Northeastern University, Boston, MA

Jan 2022–Present

*Master of Science in Industrial Engineering(Concentration: Data Science), 4.00 GPA*

Coursework: Machine Learning, Data Mining, Database Design, Data Visualization, Deep Learning

### IITRAM, India

Aug 2016–June 2020

*Bachelor of Technology in Mechanical Engineering, 3.6 GPA*

## Technical Skills

**Languages:** Python(Pandas, Scikit-learn, Numpy, Matplotlib, Seaborn, flask, Tensorflow), SQL, R

**Machine Learning Algorithms:** Regression(Linear, Logistic, Ridge, Lasso), Classification(Decision Tree, Support Vector Machine, Naive Bayes, K-Nearest Neighbour) Clustering, Time Series Forecasting, Principal Component Analysis

**Databases:** MySQL, Microsoft SQL Server, SQLite, MongoDB, PostgreSQL

**Data Analytical Tools:** Rstudio, Tableau, Google Data studio, MS Excel, Jupyter Notebook, GitHub, Flourish, Power BI

**Certifications:** AWS and GCP Fundamentals - Coursera, Advanced Excel, Programming with Python

## Projects

### Taxi Demand Prediction | *Python, Machine Learning, Time Series Forecasting, Jupyter Notebook* Sept 2022–Nov 2022

- Performed Exploratory Data Analysis on 30 million records of trips by yellow cabs in New York during January 2015.
- Divided New York into equal clusters by K-means Clustering and bifurcated time interval into equal bins of 10 minutes.
- Used Linear Regression, Random Forest and XGBoost, to forecast pickups for a given location and 10 minute interval.
- Evaluated model performance using MAPE and MSE and obtained the best one with MAPE of 12.94%.

### Predictive Maintenance Analysis | *Python, Machine Learning, Google Colab*

Sept 2022–Nov 2022

- Analyzed data of 10,000 records and 14 features regarding the factors leading to machine failure.
- Developed classification models to predict if a given machine fails or not considering the 5 main features.
- Handled data imbalance using SMOTE, leading to oversampling of minority class to 50%.
- Determined best performing model with recall of 0.89 by comparing performance using precision, recall and f-1 score.

### Application for Blood Bank Management System | *MySQL, MongoDB, Python*

Sept 2022–Nov 2022

- Designed an application for Blood Bank Management System to simplify and automate process of searching for blood.
- Developed database on MySQL, populated it to simulate real world and facilitated CRUD Operations on 9 tables.
- Incorporated table level check constraints, triggers and stored procedures by operating advanced level SQL queries.
- Accessed the database using Python and visualized the data to gain information regarding blood donors and recipients.

### Flight Delay Prediction | *Python, Machine Learning, Data Mining, Google Colab, Excel*

June 2022–July 2022

- Cleaned and preprocessed data of 0.58 million records and 22 fields, regarding the flights delayed during January 2019.
- Visualized the data by creating interactive plots to gain insights about the relation among 9 attributes.
- Designed and trained 5 machine learning models on the split data to categorize flights that were late.
- Evaluated models and established the best one with recall value of 0.85 based on the Classification report generated.

## Experience

### Northeastern University, Boston, MA

Dec 2022–Present

*Graduate Teaching Assistant - Data Mining in Engineering*

- Facilitated Professor's work by conducting weekly Python sessions and grading Homeworks, Quizzes, and Exams.
- Worked with team of 4 TA's and 8 Graders to manage class of 200 students and to provide assistance with doubt solving.
- Mentored around 25 students in their Academic Projects that involved the use of Machine Learning concepts.

### Paras Engineering Co, India

Aug 2020–Dec 2021

*Technical Analyst*

- Administered supply management and inventory control to track sales of 4 types of electrodes utilized for welding.
- Analyzed potential market opportunities in city by performing ad hoc analysis in SQL, leading to 15% increase in profit.
- Developed and monitored KPIs for customer acquisition activities in Tableau and increased revenue by 18%.
- Streamlined the scheduling of maintenance jobs to reduce overall equipment operating time by 25% using available data.
- Implemented material handling methodologies to decrease transportation and logistics cost by 20%.