

# Md Sarfaraz Iqbal

Data Analyst / Data Scientist

 sarfaraziqbal7860@gmail.com

 8005957046

 linkedin.com/in/mdsiqbal

 github.com/sarfaraziqbal

## WORK EXPERIENCE

### ESMITO SOLUTIONS

#### Battery Health Analysis and Prediction

03/2023 - Present

Tags: Machine Learning, Statistical Modelling, Feature Engineering, Data Analysis, battery Health Prediction, SOH, Cycles, Optimization, Research, Mathematical Modelling.

- Analyzed different factors that impacted **battery health**. Explored factors like charging cycles, temperature, cell Voltage, and usage patterns to identify the factors that most contributed to battery degradation.
- Presented findings and **recommendations** based on the analysis of battery health factors, enabling stakeholders to make informed decisions regarding battery **maintenance, replacement** and optimization.
- Developed battery **health prediction** model using techniques like **Coulomb Counting, Cyclic** and **Calendar degradation** method to predict and extend its **lifespan**.

### TIGER ANALYTICS

#### My Depot Marketing Campaign

06/2022 - 12/2022

Tags: Machine Learning, Classification, Feature Engineering, Supervised Learning, Clustering, Uplift Modelling, Market Mix Modelling, Decision Tree, Recall, Precision, ROC-AUC

- Conducted **analysis** analysis to identify factors that influence customer behavior, such as purchase history, demographics, and interests
- Developed and implemented **uplift models** to improve the effectiveness of personalized marketing campaigns.
- Reduced marketing costs by **5%** by targeting customers with the highest likelihood of conversion.

## PERSONAL PROJECTS

### Credit Card Default Prediction

01/2022 - 03/2022

Tags: Machine Learning, Classification, Imbalanced data, Feature Engineering, Supervised Learning, Hyperparameter Tuning, Decision Tree, Recall, Precision, ROC-AUC

- Developed a **binary classification model** to predict whether a customer will default on credit card payments.
- Implemented techniques to **oversample** the minority class observations, encoded categorical variables using various encoding techniques and carried out hyperparameter tuning using **GridSearchCV**.
- Experimented with algorithms such as Logistic Regression RandomForest, XGBoost and SVC and selected XGBoost based on **Recall rate** of **86%** and **AUC score** of **91.4%**.

### Online Retail Customer Segmentation

09/2021 - 12/2021

Tags: *k means, Recency, Frequency, Monetary, Silhouette and Elbow score*

- Built a clustering model using **K-means** to identify major customer segments on transactions dataset for the UK based non-store online retail.
- Applied feature Engineering to obtain new features such as **Recency Frequency Monetary** and RFM score for getting more details about the customer.
- Obtained optimal number of clusters using **Silhouette analysis, Elbow method** and visual inspection of Dendrogram resulting from Hierarchical Clustering.
- Generated optimal customer clusters which could be further used in various use-cases like **customer targeting, marketing campaigns, churn prediction, etc.**

## TECH STACK

### Languages

Python, Sql

### Frameworks and Libraries

Numpy, Pandas, Scikit-learn, Matplotlib, Seaborn, TensorFlow, Keras, Flask, Streamlit, OpenCV

### Tools and Platforms

Power BI, Jupyter Notebook, Git, Tableau, Excel

## RELEVANT COURSEWORK

The Data Science Course 2021: Complete Data Science Bootcamp (Udemy) (12/2020) [↗](#)

SQL MySQL for Data Analytics and Business Intelligence(Udemy) (08/2021) [↗](#)

## PUBLICATIONS

*Publication Type*

**Regularization and its importance in Machine Learning** [↗](#)

01/2022

## ACHIEVEMENTS

Machine Learning and MySql Badge

LinkedIn

Python and Sql Badge

Hackerrank

## EDUCATION

— **B.Tech - Electronics And Communication**  
MIT, Muzaffarpur [↗](#)

2017 - 2021

7.57

— **XII- Higher Secondary**  
Rose Public School [↗](#)

2015 - 2016

70.8%

## INTERESTS

Books

Cricket

Exercises