Saddam Husain
Data scientist | Noida- 201309

+91-7054996930   saddamhusain7058@gmail.com   in linkedin.com/in/saddam-husain-73749b22b/   🌐 github.com/Saddam705

PROJECTS TECH STACK

Online Retail Customer Segmentation Languages
AlmaBetter Verified Project ↗
03/2023-04/2023

Tags : Clustering, K Means Clustering, RFM model, Elbow method, Silhouette score, hierarchical clustering, dendrogram, optimal cluster, DBSCAN clustering, TF-IDF, tokenization

o Built a **Clustering model** using **K means clustering**, **Agglomerative clustering** to identify major **customer segments** on a transactional data set that contains all the transactions occurring between 01/12/2010 and 09/12/2011 for a UK-based and registered non-store online retail.

o Applied **feature engineering** to obtain new features such as **Recency, Frequency, Monetary, busiest days, busiest time, month RFMGroup,** and **RFMScore** for getting more details about the customer.

o Obtained the optimal number of clusters using **Silhouette Analysis** and **Elbow Metho**d and similarly identified the optimal clusters at the threshold 85 **by visually inspecting the dendrogram from hierarchical clustering.**

Credit Card Default Prediction
AlmaBetter Verified Project ↗
02/2023-03/2023

Tags: Classification, KS Statistic, Gains table, SMOTE, Hyperparameter Tuning, SHAP Interpretability, Extreme, Gradient Boosting

o Developed a binary classification model using algorithms such as **Logistic Regression SVC** and **XGBoost** to predict whether a customer will default on credit card payments.

o Engineered a new class of attributes known as **decayed field variables** and developed out-of-pattern variables on historical credit and bureau data to identify risky customers and **reduced default rate** from 8.3% to 6.5%.

o Performed missing value imputation using **KNN-Imputer**, implemented **SMOTE boosting** to oversample the minority class observations and carried out **hyperparameter** tuning using **Bayesian optimization**.

o Obtained Model Reason Codes (MRCs) by leveraging the concept of **SHAP** plots to cater to customer grievances and analyzed the **Gains table** to decide rejection cutoff.

INTERNSHIPS

Internship from **Internshala**↗

01/11/2021 – 01/01/2022
That was an awesome internship where I have learned supervised and unsupervised machine learning models and implemented in many projects. I had learned ml algorithms like **linear regression, logistic regression, SVM**, **random forest, decision tree, k-Means clustering** and many more **algorithms**.

EDUCATION

B.Tech in Computer Science

SR Institute of management & tech Lucknow
2019 - 2023

Skills

TECH STACK Languages-

Python, SQL, C, Java, HTML, CSS

ML Frameworks-

Scikit-learn, spaCy, Keras, Pandas, Numpy, Matplotlib, Seaborn, PyTorch, NLP

Platforms-

Jupyter Notebook, Sublime, GitHub, Visual Studio, Tableau, PowerBI, Excel

Certifications

Data Science Certification at
FastLane Career - 2023

Data Science Certification at
AlmaBetter - 2023

Data Science Certification at
Coursera - 2022

ACHIEVEMENTS

SQL 5 star on **Hackerrank** ↗
Python 5 star on **Hackerrank** ↗