

Shubham Rai

Highly motivated and results-driven professional with a years of experience as a Software Developer. Possessing strong analytical & problem-solving skills, I have demonstrated proficiency in Python, SQL, and Machine Learning. My extensive background includes academic roles & hands-on development experience. Eager to leverage my expertise in a challenging data science role, where I can continue making meaningful contributions to impactful projects.

shubhamrai1318@gmail.com

+91-9918621318

linkedin.com/in/shubhamrai1318

github.com/shubhamrai1318



ACADEMIC PROJECTS

Netflix Movies and TV Shows Clustering

AlmaBetter Verified Project

08/2023 - 09/2023

Tags: K-Means, Silhouette score, Hierarchical Clustering, Dendrogram, Agglomerative Clustering

- Engineered an **Unsupervised ML** model to perform **clustering** on a comparable dataset and applied **data transformations**, including **one-hot coding** and **dummy variable conversion**, while ensuring **standardized** data for optimal model fitting.
- Conducted **hypothesis testing** to explore **correlations** between content types & viewer engagement on the Netflix platform. Attained a **silhouette coefficient** of 3 through Silhouette Analysis & the **Elbow Method**, indicating cluster separation.
- Visualized** clustering outcomes through a comprehensive **Hierarchical Clustering Dendrogram**. The primary **goal** was to refine **content recommendations**, ensuring a more tailored and **engaging user experience** on the Netflix platform.

Credit Card Default Prediction

AlmaBetter Verified Project

07/2023 - 08/2023

Tags: Supervised ML, Classification, Sampling Method, Feature Engineering, Decision Tree, XGBoost

- Developed a **Classification Model** for **predicting** credit card payment defaults.
- Implemented comprehensive **feature engineering** steps, including Feature Selection, **Standardization**, **Treating Multicollinearity**, and **One-Hot Encoding**.
- Employed **GridSearchCV** for **hyperparameter tuning**, optimizing the model's predictive capabilities & comparative analysis revealed that **XGBoost** outperformed the **Decision Tree Classifier**, achieving a remarkable **Recall of 88%**.
- Demonstrated expertise** in Hyperparameter Tuning to **fine-tune model** parameters and enhance overall **predictive accuracy**.

TED Talk Views Prediction

AlmaBetter Verified Project

05/2023 - 06/2023

Tags: EDA, Supervised Learning, Regression Analysis, Handling Outliers, Ensemble Technique.

- Conducted **data preprocessing**, including meticulous cleaning and feature engineering and employed a variety of **regression models**, including **Linear Reg**, **Regularization**, **Random Forest**, to capture the complexity of viewership trends.
- Assessed **model performance** using a suite of metrics, including **Mean Squared Error**, **Root Mean Squared Error**, **R Square**, and **Adjusted R Square**.
- Utilized advanced **Cross-validation** techniques such as **Grid Search CV** and **Random Search CV** to fine-tune model hyperparameters, optimizing performance.
- Achieved remarkable results with **Random Forest**, attaining a minimum **MSE** and an impressive **86% R Square** after careful tuning.



TECH STACK

Expertise in Languages & Tools (x/5)

Python - 4.5 || SQL - 4.5 || Tableau - 4.0 || Excel - 4.0 || PowerBI - 4.2 || GitHub - 4.7 || Looker Studio - 3.5

Platforms

Jupyter Notebook, Google Colab, MS Office, GitHub, Kaggle, SQL Server, PostgreSQL, SQLite, Hacker Rank

ML Frameworks

Pandas, NumPy, Matplotlib, Sci-kit Learn, Spacy, Keras, Plotly, Seaborn, Streamlit



WORK EXPERIENCE

Dot Net Developer

Yoma Technologies Pvt. Ltd.

02/2022 - 02/2023

Gurugram, Haryana

Skills: ASP.Net, SQL Server, MVC, Reposo, Team Work

- Developed and maintained software applications** using the **.NET framework**.
- Collaborated with cross-functional teams** to gather requirements and design scalable solutions.
- Conducted **code reviews** and implemented best practices to ensure **code quality**.
- Played a key role in the **Research & Development** team, specializing in **Python ML projects**.



TRAINING EXPERIENCE

Data Science Trainee

AlmaBetter

02/2023 - Present

Bengaluru, Karnataka

Skills: Excel, Supervised ML, Unsupervised ML, SQL, Tableau, Power BI, Good Communication skill, Team Work

- Explored and learnt technical skills such as **Python**, **SQL**, **Exploratory Data Analysis & Visualizations**. Acquired knowledge by building **dashboards** using **Tableau** and **PowerBI**.
- Secured the **4th rank** in a cohort of **300 students**, placing in the **Top 5%** and earned the opportunity to work as a **Teaching Assistant** (Data Science).



EDUCATION

Master of Computer Application

AKTU, Lucknow

10/2020 - 07/2022

CGPA || 8.7

Bachelor of Computer Application

MGKVP, Varanasi

08/2017 - 10/2020

Percentage || 66%



ACHIEVEMENTS

Gold Badge in Python and SQL

HackerRank

Teaching Assistant (Data Science)
(10/2023 - Present)

AlmaBetter