

# Saransh Srivastava

Data Scientist | Faridabad

✉ saranshoffice@gmail.com

☎ +918178118585

🌐 linkedin.com/in/saranshsrivastava13

🐙 github.com/saransh2396

## WORK EXPERIENCE

### Constructional Site Engineer

Mansi Ganga Builders & Engineers Pvt. Ltd.

06/2019 - 09/2020

Greater Noida, Noida

Tasks/Achievements

- Planning, Reporting, Billing and Coordination with Subcontractor, Client & PMC/TPI for Material & Progress for CGD project of IGL Project – Natural Gas CS Steel Pipeline

## PROJECTS

### Cardiovascular-Risk-Prediction

Almabetter Verified Project 

11/2021 - 12/2021

Tags - (Classification, Label Encoding, SMOTE, Hyper parameter Tuning).

- The early prognosis of cardiovascular diseases can aid in making decisions on lifestyle changes in high risk patients and in turn reduce the complications. Predict the overall risk.
- Implemented **Smote Boosting** and implemented Classification Models like **Logistic Regression**, **SVM**. Finalized SVM and carried out Hyper parameter tuning using **Grid Search CV**.
- Deployed the model in **Heroku** with **Flask web framework**.

### Seoul-Bike-Sharing-Demand-Prediction

Almabetter Verified Project 

11/2021 - 11/2021

Tags- Regression, Multicollinearity, Skewness, Label Encoding, Random Forest Regression.

- The prediction of bike count required at each hour for the stable supply of rental bikes.
- Performed log transformation on the variables, Removed Multicollinearity, Label Encoded the Categorical Values, **Implemented Linear, Polynomial and Random Forest Regression** among which Random Forest Regression was giving best results so chose it
- Carried out **Hyper parameter Tuning using Grid Search CV** on the finalized model to reduce over fitting.
- Deploying the model using Django Framework.

### Netflix Content Clustering

Almabetter Verified Project 

12/2021 - 12/2021

Tags: K-Means Clustering, Elbow Method, Silhouette score

- This dataset consists of tv shows and movies available on Netflix as of 2019. The dataset is collected from Flixable which is a third-party Netflix search engine. **Clustering similar content** by matching text-based features.
- Manipulated the dataset using **Pandas & analyzed the data using Matplotlib & Seaborn**. After analyzing gain some insights about the type of contents (movies, TV shows) Netflix has.
- Performed **NLP methods** on the description column of each content and then converted it into a **Tfidf Matrix**
- Implemented **K Means Algorithm** in which estimation of value of k using **elbow method**.

## TECH STACK

### Language

Python, SQL

### ML Frameworks

Scikit-learn, Pandas, Numpy, Matplotlib, Seaborn, NLTK, spaCy, Tensorflow, Keras, Flask

### Databases

MySQL

### Platforms

Jupyter Notebook, Google Colab, MS Office, Excel, GitHub

## EDUCATION

### B.Tech (Mechanical Engineering)

Manav Rachna College of Engineering  
affiliated by (MDU, Rohtak)

08/2014 - 07/2018

Faridabad

## CERTIFICATES

Master's Certificate in Artificial Intelligence  
Engineer from Simplilearn, Bangalore  
(10/2020 - 06/2021)

Machine Learning, Deep Learning, Computer Vision, NLP

Python from DUCAT India- Noida  
(03/2021 - 05/2021)

## ACHIEVEMENTS

Star Student at AlmaBetter (01/2022)

Achieved star student badge from AlmaBetter institute

## INTERESTS

Football

Music

Gaming

Travelling