

# Architecture Design Bike Share Prediction

Number – 1.0

Date : 10-06-2023

## Document Version Control

Date	Version	Description	Author
10-06-2023	1.0	Abstract Introduction Architecture	SAHIL JOSAN

## Contents

Abstract.....	3
1. Introduction.....	4
2. Architecture.....	4
3. Architecture Design.....	5
3.1 Data Collection.....	5
3.2 Data Description.....	5
3.3 Data Pre-Processing.....	5
3.4 Modelling Process.....	5
3.5 UI Integration.....	6
3.6 Data from user.....	6
3.7 Data Validation.....	6
3.8 Rendering the Results.....	6
3.9 Deployment.....	6

## Abstract

Bike sharing systems are new generation of traditional bike rentals where whole process from membership, rental and return back has become automatic. Through these systems, user is able to easily rent a bike from a particular position and return back at another position.

Apart from interesting real world applications of bike sharing systems, the characteristics of data being generated by these systems make them attractive for the research. Opposed to other transport services such as bus or subway, the duration of travel, departure and arrival position is explicitly recorded in these systems. This feature turns bike sharing system into a virtual sensor network that can be used for sensing mobility in the city.

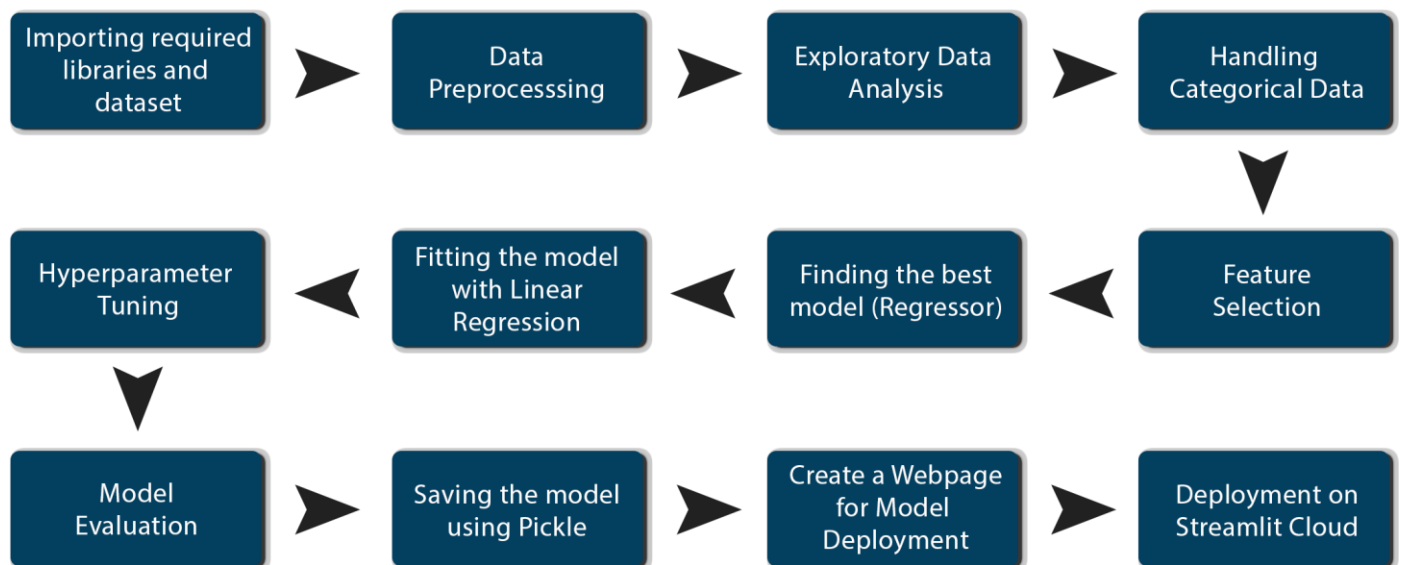
I am considering variables as season, yr,mnth, holiday, weekday, workingday, weathersit, Temp, atemp, hum, windspeed, casual, registered to predict the count of total rental bikes That will be required.

# 1. Introduction

## 1.1 Why this Architecture Design Document?

The main objective of the Architecture design documentation is to provide the internal logic understanding of Bike Share Prediction. The Architecture design documentation is designed in such a way that the programmer can directly code after reading each module description in the documentation.

# 2. Architecture



## 3. Architecture Design

### 3.1 Data Collection

The data for this project is collected from UC Irvine Machine Learning Repository. The link for the data is <https://archive.ics.uci.edu/dataset/275/bike+sharing+dataset>

### 3.2 Data Description

Bike sharing dataset is publicly available on UC Irvine Machine Learning Repository. This dataset contains the hourly and daily count of rental bikes between years 2011 and 2012 in Capital bikeshare system with the corresponding weather and seasonal information.

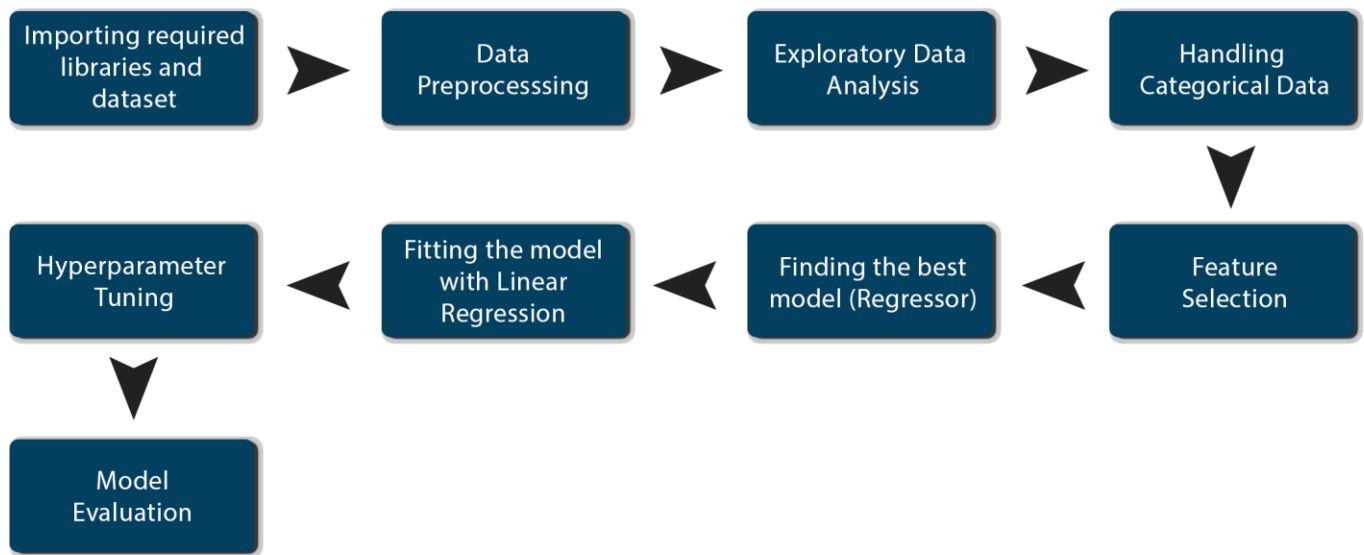
### 3.3 Data Pre-processing

- Checked the datatype of features in dataset using `df.info()`
- Checked for Null values, because the null values can affect the accuracy of the model.
- Perform Label Encoding for the features that have categorical data.
- Checked the distribution of the features to interpret its importance.

Now, the data is prepared to train a machine learning model.

### 3.4 Modelling Process

After pre-processing the data, we visualize our data to gain insights and split into two parts, train and test data. After Splitting the data, we use different machine learning algorithms like – Linear Regression, Random Forest Regressor, Decision Tree Regressor to predict the Count of Bike share



### 3.5 UI Integration

Streamlit files are being created and are being integrated with the created machine learning model. All the required files are then integrated to the app.py file and tested locally.

### 3.6 Data from user

The data from the user is retrieved from the created streamlit web page.

### 3.7 Data Validation

The data provided by the user is then being processed by app.py file and validated. The Validated data is then sent to the prepared model for the prediction.

### 3.8 Rendering the results

The data sent for the prediction is then rendered to the web page

### 3.8 Deployment

The tested model is then deployed to streamlit cloud. So, user can access the project from any internet device.