**Bike Sharing Demand Prediction**

This is a project that predicts the demand for bike sharing based on various attributes such as weather conditions, time of day, and user demographics. The dataset used in this project is publicly available.

**Project Goal**

The main goal of this project is to build a machine learning model that can accurately predict the demand for bike sharing based on various input features.

**Dataset**

The dataset used in this project contains hourly bike rental data spanning two years. It includes information such as date, time, temperature, humidity, wind speed, and the number of bikes rented in that hour. There are a total of 17 attributes in the dataset.

**Exploratory Data Analysis**

The dataset was first analysed to gain insights into the distribution of data, correlations between attributes, and missing values. The analysis showed that the data was mostly normally distributed and there were some correlations between attributes. Additionally, there were no missing values in the dataset.

**Data Pre-processing**

The data was pre-processed by removing unnecessary attributes and converting categorical attributes to numerical attributes using one-hot encoding. The data was also normalized to ensure that all attributes were on the same scale.

**Model Building**

Several machine learning algorithms were used to build the prediction model, including linear regression, ridge regression, lasso regression, decision tree, random forest, gradient boosting, and cat boost. Each model was evaluated using various performance metrics such as mean squared error (MSE) and root mean squared error (RMSE).

**Conclusion**

The results of the model evaluation showed that the cat boost algorithm performed the best with the lowest MSE and RMSE. The final model was trained on the entire dataset and can be used to predict the demand for bike sharing given various input features.