**Bytexl’s guided projects**

**Students’ User Guide**

**INSTRUCTIONS FOR STUDENTS:**

**Project based learning course overview:**

This project-based learning course will guide you through the process of building a machine learning model to predict bike-sharing demand using real-world data. You'll gain practical experience in data analysis, feature engineering, model selection, and evaluation.

**About the project:**

The Bike Sharing Demand Prediction project aims to develop a robust and accurate regression model for forecasting bike rental demand in urban areas. This project addresses the operational challenges faced by bike-sharing companies by ensuring optimal bike availability while minimizing resource wastage and maximizing customer satisfaction.

By accurately predicting bike rental needs, these companies can streamline operations, reduce costs, and improve user experience. The project leverages historical bike usage patterns and weather data to identify key factors influencing rental demand. By understanding these relationships, the model learns to make informed predictions about future demand. This information enables bike-sharing companies to efficiently allocate bikes, optimize maintenance schedules, and ensure a smooth rental experience for customers.

**Prerequisites:**

* Basic understanding of Python programming and data analysis.
* Familiarity with fundamental machine learning concepts.

**What will you learn?**

* Understand the problem of bike sharing demand prediction and its importance.
* Explore and analyze real-world bike sharing data using Python libraries.
* Perform data preprocessing and feature engineering to prepare data for modeling.
* Select and train appropriate machine learning models for regression tasks.
* Evaluate model performance using relevant metrics and interpret results.
* Gain insights into the factors influencing bike rental demand.
* Communicate your findings and present your project effectively.

**Skills you will practice:**

* Python programming language
* Jupyter Notebook or Google Colab environment
* Libraries like Pandas, NumPy, Seaborn, and Scikit-learn
* (Optional) Cloud-based machine learning platforms for deployment

**How to execute? Your learning platform:**

**1. Setting Up the Environment:**

* **Google Colab:** Create a new notebook in Google Colab. This provides a cloud-based environment with pre-installed libraries.
* **Libraries:** Import necessary libraries (Pandas, NumPy, Seaborn, Scikit-learn) using import statements at the beginning of your notebook.
* **Data:** Upload your bike sharing dataset to your Colab environment. You can use the files section in Colab's sidebar or mount your Google Drive to access the data.

**2. Data Exploration and Preprocessing:**

* **Exploration:** Use Pandas functions to load the data, examine its structure (df.head(), df.info()), and summarize key statistics (df.describe()).
* **Visualization:** Use Seaborn or Matplotlib to create visualizations (e.g., scatter plots, histograms) to understand data distributions and relationships between features.
* **Preprocessing:** Handle missing values (e.g., imputation or removal), convert categorical variables using one-hot encoding, and potentially scale numerical features.

**3. Feature Engineering and Selection:**

* **Engineering:** Create new features based on existing ones (e.g., combining date and time information, deriving features from weather data).
* **Selection:** Choose relevant features for modeling using techniques like correlation analysis, feature importance from tree-based models, or domain knowledge.

**4. Model Selection and Training:**

* **Selection:** Choose a regression model from scikit-learn based on the nature of the data and project goals (e.g., LinearRegression, RandomForestRegressor).
* **Training:** Split the data into training and testing sets, then use the training data to train your chosen model using its fit() method.

**5. Model Evaluation and Tuning:**

* **Evaluation:** Use appropriate metrics like RMSE, R-squared, and MAE to evaluate the model's performance on the test data.
* **Tuning:** Adjust model hyperparameters using techniques like grid search or randomized search to find the best settings for your data.

**6. Deployment and Reporting:**

* **Deployment:** If desired, consider deploying the model using a cloud platform like Google Cloud AI Platform for real-time predictions.
* **Reporting:** Create a comprehensive report summarizing your project, including data exploration, feature engineering, model selection, evaluation, and insights gained.

**Learn step-by-step:**

**In this guided project, you will find your educator giving you a walk-through to complete your project in 20 hours.**

**Structure for educators:**

**Welcome to Bike Sharing Demand Prediction. This is a guided project which will take about 20 hours to complete.**

**Here are the course objectives and structure:**

**Course Objectives:**

**In this project we will focus on the following objectives:**

* Understand the problem of bike sharing demand prediction and its importance.
* Explore and analyze real-world bike sharing data using Python libraries.
* Perform data preprocessing and feature engineering to prepare data for modeling.
* Select and train appropriate machine learning models for regression tasks.
* Evaluate model performance using relevant metrics and interpret results.
* Gain insights into the factors influencing bike rental demand.
* Communicate your findings and present your project effectively.

**By the end of this project you will be able to predict You will deploy the project on the Nimbus Platform using Python tools / platforms**

**Course Structure:**

**This course is divided into 3 parts:**

**Course overview: This is the introductory reading material.**

**Project structure:**

**The hands on project on Bike Sharing Demand Prediction  is divided into following tasks:**

**Task 1: Data Acquisition and Initial Exploration**

**Task 2: Data Cleaning and Preprocessing**

**Task 3: Exploratory Data Analysis and Feature Selection**

**Task 4: Model Selection and Training**

**Task 5: Model Evaluation and Tuning**

**Task 6: Deployment and Communication**

**Meet your educator:**

Hi I am Ayushman Dubey and I will be your instructor for your course. I have about 2 years of experience in Artificial Intelligence, Machine Learning. I have worked in Amazon, ByteXL. I also work with different business organizations. I have a Bachelor’s degree in Aeronautical Engineering. When I am not teaching, I enjoy Singing, Travelling, Playing Flute. I also love to visit places.

**About the Nimbus Platform and the tools required for this project:**

Nimbus is a cloud-based coding environment developed by byteXL. It serves as a self-hosted solution that allows students to access essential coding tools without the need for local software installation.

**Quiz Questions :**

**Question 1:**

Which of the following is NOT a typical feature used in bike sharing demand prediction?

(a) Temperature  
(b) Humidity  
(c) User's Age  
(d) Holiday

**Answer:** (c) User's Age

**Question 2:**

What type of machine learning problem is bike sharing demand prediction?

(a) Classification  
(b) Regression  
(c) Clustering  
(d) Dimensionality Reduction

**Answer:** (b) Regression

**Question 3:**

Which Python library is commonly used for data manipulation and analysis in this project?

(a) NumPy  
(b) Pandas  
(c) Seaborn  
(d) Scikit-learn

**Answer:** (b) Pandas

**Question 4:**

What is the purpose of feature engineering in this project?

(a) To collect raw data from different sources  
(b) To select the best machine learning model  
(c) To transform and create new features to improve model accuracy  
(d) To deploy the model in a real-world environment

**Answer:** (c) To transform and create new features to improve model accuracy

**Question 5:**

Which metric is commonly used to evaluate the performance of a regression model?

(a) Accuracy  
(b) Precision  
(c) Recall  
(d) Root Mean Squared Error (RMSE)

**Answer:** (d) Root Mean Squared Error (RMSE)

**Earn a Certificate: After you have completed the Bike Sharing Demand Prediction hands-on project, you should**

**upload your code for it to be assessed**

**complete the Quiz to assess your knowledge.**

**You will earn a certificate if you score 80 % or more.**