CMPE-257- Machine Learning

Submitted To: Dr. Ali Arsanjani
Teacher Assistant: Sravani Thota
Team Name: Team_Spartans

Team Members:

- Jack Kalavadia (016197519)
- Pramatha Nadig (016708588)
- Rishikesh Andhare (016726203)
- Rutvik Moradiya (016656367)

Prediction of Bike Rental Count

A.Business Purpose:

The primary objective of this machine-learning project is to address critical challenges faced by bike rental businesses: maximizing profitability and ensuring customer satisfaction. By leveraging machine learning and data science, the project aims to accurately forecast bike demand based on factors like weather conditions, seasonality, temperature, humidity, wind speed, and holidays. This enables businesses to optimize operations, enhance customer experiences through improved bike availability and tailored pricing strategies, and adapt to market conditions, especially during economic crises. The project fosters data-driven decision-making, offering valuable insights to support informed choices and sustainable growth while contributing to eco-friendly transportation solutions.

Goals:

The primary goal of this machine learning project is to aid bike rental businesses, particularly those grappling with revenue challenges during market crises, in optimizing profitability and ensuring efficient customer experiences. To accomplish this, the project aims to develop accurate machine learning models for bike demand forecasting, taking into account factors like weather conditions, season, temperature, humidity, wind speed, and holidays. This enables businesses to tailor their strategies to meet customer expectations and demand levels effectively. Additionally, the project provides insights into demand patterns in both current and potential new markets, guiding businesses in expansion decisions. Ultimately, it supports the long-term sustainability of bike rental companies by devising income-boosting strategies as economic and market conditions stabilize.

Data Narrative:

In this project, the primary dataset comes from Capital Bikeshare and includes a range of attributes such as weather conditions, season, and holidays. The data is believed to be sourced from bike rental companies and bike-sharing platforms. The dataset comprises various factors that could influence bike rentals, including weather conditions, season, temperature, humidity, wind speed, and holidays. Before modeling, quality checks and preprocessing steps like imputation and normalization will be carried out to ensure data integrity. To evaluate the performance of

machine learning models, the dataset will be divided into training and testing sets. Additionally, feature engineering techniques may be employed to introduce new attributes that could enhance model performance.

B.Experiments:

1-How can we accurately predict the demand for bike rentals based on various factors, such as weather, season, temperature, humidity, wind speed, and holidays?
2- What properties should be invested in to maximize bike rental profitability ?
3- What is the relationship between temperature levels and bike rental demand, and how can this information be used to optimize bike availability?

4-What are the yearly patterns of temperature and precipitation in the areas where bike rentals are offered, and how do these patterns influence demand throughout the year?

5) Key Performance Indicators (KPIs) for Bike Rental Prediction:

a) Accuracy KPI:
- Metric: Accuracy
- Formula: (True Positives + True Negatives) / (True Positives + True Negatives + False Positives + False Negatives)

b) Customer Satisfaction KPI:
- Metric: Bike Availability Rate
- Formula: (Number of Bikes Available / Total Bikes) * 100

C. Data set:

Capital Bike share

Bike Data

Historical Weather data sites

D. ML: Design an experiment (at least one) for each homework assignment

K-Means and Other Algorithms:-
K-means and other clustering techniques can be particularly helpful in identifying distinct patterns in historical weather data. By categorizing weather conditions into different clusters, these algorithms provide valuable insights into their potential correlation with various wildfire outcomes.
Clustering:
1. Segmentation by Demand Patterns: Group similar days based on demand.
2. Market Expansion: Identify similar demand profiles in new markets.
3. Customer Targeting: Segment customers based on preferences.
4. Service Customization: Offer different bike models or pricing strategies based on demand clusters.

Classification:
1. Demand Prediction: Predict high or low demand days.
2. Holiday Impact: Assess the effect of holidays on demand.
3. Weather Conditions: Categorize days based on weather conditions.
4. Seasonal Trends: Classify days into different seasons.

Regression:
1. Demand Forecasting: Predict the exact number of bike rentals.
2. Optimizing Inventory: Determine the number of bikes needed.
3. Pricing Strategy: Analyze how pricing affects demand.
4. Resource Allocation: Guide resource allocation in new markets.

Articles:

1. IJISET - International Journal of Innovative Science, Engineering & Technology- Bike share demand prediction.
2. J. Larsen. (April 25, 2013). Plan B Updates - 112: Bike-Sharing Programs Hit the Streets in Over 500 Cities Worldwide.
3. P. DeMaio, "Bike-sharing: History, Impacts, Models of Provision, and Future," Journal of Public Transportation, vol. 12, no. 4, p. 3, 2009.
4. Abdelhalim, A., & Traore, I. (2009). A new method for learning decision trees from rules. In Machine Learning and Applications, 2009. ICMLA 09. International Conference on (pp. 693–698).
5. Alippi, C., & Roveri, M. (2010). Virtual k-fold cross validation: An effective method for accuracy assessment. In The 2010 International Joint Conference on Neural Networks (IJCNN)
6. Joelsson, S. R., Benediktsson, J. A., & Sveinsson, J. R. (2005). Random forest classifiers for hyperspectral data. In Geoscience and Remote Sensing Symposium, 2005. IGARSS 05. Proceedings. 2005 IEEE International (Vol. 1, p. 4–pp). IEEE.