

BUSINESS

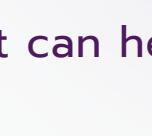
FORTH WEEK TASK

Explore and clean the data

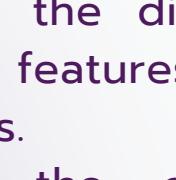


- Explore the dataset to understand the structure and content of the data.
- Check for missing values and handle them appropriately (e.g., impute or drop).
- Clean the data by converting data types, handling outliers, and removing duplicates.

Feature engineering



- Extract relevant features from the dataset that can help in predicting the fare amount.
- Engineer new features, such as:
 - Time-based features (e.g., hour, day, month, year)
 - Distance-based features (e.g., distance between pickup and dropoff locations)
 - Weather-based features (e.g., categorical encoding of weather conditions)
 - Traffic-based features (e.g., categorical encoding of traffic conditions)
 - Car condition-based features (e.g., categorical encoding of car conditions)

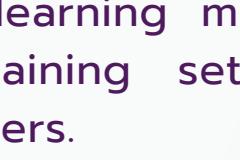


Visualize and analyze the data

- Visualize the distribution of fare amounts and other relevant features using histograms, scatter plots, and box plots.
- Analyze the correlations between features using heatmaps and correlation matrices.



Train a machine learning model



- Train a machine learning model (KNN) on the training set to classify the customers.

Presentation



We would create some super canva presentation to apply and sum and show our findings

WEB app (IF WE GOT TIME)



We would create some small single one page application to show the for which degree would a customer be

BY THE TEAM