Brief approach to solve the problem.

1.Given train dataset of shape **(18247,3),** with **3** columns of all numerical type.

2.Basic data overview in terms of shape, missing values, min, and max values

3.Followed by Feature engineering to extract features from the given data.

4.Exploratory data analysis (EDA) to get more depth insights about the data and give some business-related perspectives from data.

5.Under EDA -univariate and bivariate analysis

6.Data splitting into train and validation

7.Standardization of train data.

8.Finding right Machine learning (ML) model, where we fit train data to the given ML algorithms

And evaluate them using a RMSE metric.

9.Choose ML algorithm and hyper tuning it and check for RMSE score.

10.Test ML algorithm with validation data and check for RMSE score.

11.Predcit ML algorithm on test data

Data-pre-processing / Feature Engineering ideas really worked? How did you discover them?

1.Given train data with 3 columns, where one is converted into datetime format and other 2 columns are in numeric datatype. Using pandas' data time function, generated 3 new features date, week, month and year.

2.There were no missing values in data, and since features are valued in dates, weeks and years there were no outliers.

3.Days and months Featured engineered column will help us to give the demand across the time frame. This demand variation was capture in EDA.

4.Since data was in time series format, to get better patterns from the date column. Extracted the 3 new features.

5.From correlation there are 3 independent features which are more correlated to demand feature (dependant feature). Following is the order: Hour>week>month>year.

6.Dropped date column from train data as we have already extracted features from it. And kept remaining columns later feed them into models.

What does your final model look like? How did you reach it?

1.From 7 ML + 1 Deep neural network model based on RMSE score was able to find 3 model which are having good RMSE score.

      a) Deep neural network: **38.70442025947229**

      b) XGBoost regressor with RMSE score: **37.587497001282905**

      c) Gradient boosting regressor with RMSE score: **37.58390963641146**

2.Final model after hyper parameter tunning, which performed good with low RMSE score was: **Gradient boost regressor** with RMSE score of: **37.12983198436382** (on validation data) **with** given hyper parameter values are n_estimators=700, max_Depth=3

#Summary of data
A. Observations

1. More demand for rental cars in months of January, February, October, and November.
Least in July.
2. More demand for car is starting at 10 am to night 8pm. And minimum at 12 am
3. From monday to tuesday demand is almost same but on friday demand is more compared to rest of other days followed by thursday.
4. More demand has shot up in 2019 and 2020.

B. Way forwards

1. Introducte new offer like 5-10% off per 1km in the month of july to increase revenue.
2. Need to allocate more cars in the months of January, February,
october to match with the supply. And also increase the Car collecting points areas.
3. Add some attractive offers on thurdays since its demand ranked at 2nd ,it can have potential to at tract new customers and add more revenue to company.