Data Storm 3.0

Team Name:24CODERS

Kaggle Username: ds2282

Display Name: ds2282

Details of Team members:

Member1:

- Name: W.A.Ravinath
- Phone Number: 0769734553

Member2:

- Name:A.A.P.M.Ranasinghe
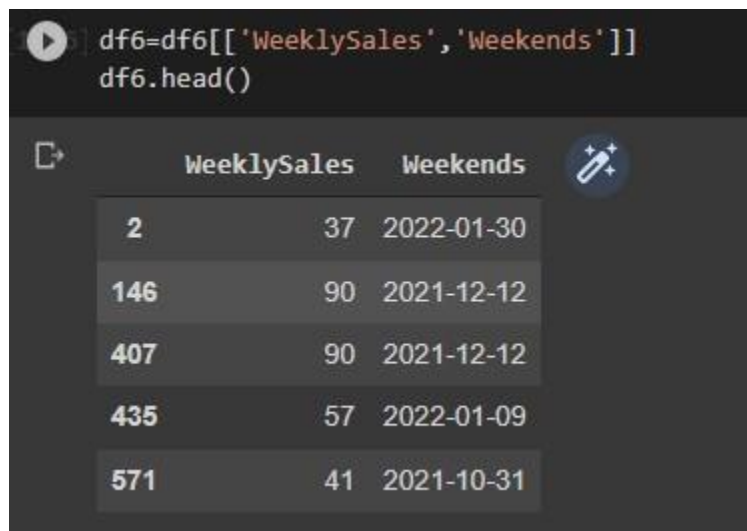- Phone Number:0768351756

Member3:

- Name:Manula Bojith Gunawickrama
- Phone Number:0776902299

GitHub repository link:https://github.com/Ravinath98/24CODERS-Data-Storm

Lowest Total MAPE: 64.16049(in Kaggle)

Steps for the approach

- Coding part was done on the Google colab notebook
- 1st done the datapreprocessing on training dataset
- We looked at some insights of the dataset using various commands
- Because the problem was sales forecasting,we identified it's wanted consider this problem as a time series problem
- So, we made the train dataset suitable for time series problem
- With the help of given DateID, by doing some preprocessing tasks, we got the weeklysales sum from the training dataset(for a specific itemcode with a specific category)
- We got the every week Sunday's date as the date for weeklysales('Weekends' column). So, the dataframe columns suitable for time series problem as below(example for specific item with specific category)

Figure 01: Example dataframe which made suitable for time series forecasting

- We used the Prophet Model for this time series forecasting problem
- We forecasting for period of 4(because the problem hope for 4 weeks forecasting for a specific item with specific category)
- So, by iterating through the train dataset and made a specific dataframe for each item with it's specifc itemcode and category code and then train with the model and predict for the future 4 weeks to that item sales(weekly sales). Then assign those predicted value for the test set by condidering the item code ,category and the week.
- Below given a specific itemcode with specific category code prediction illustration
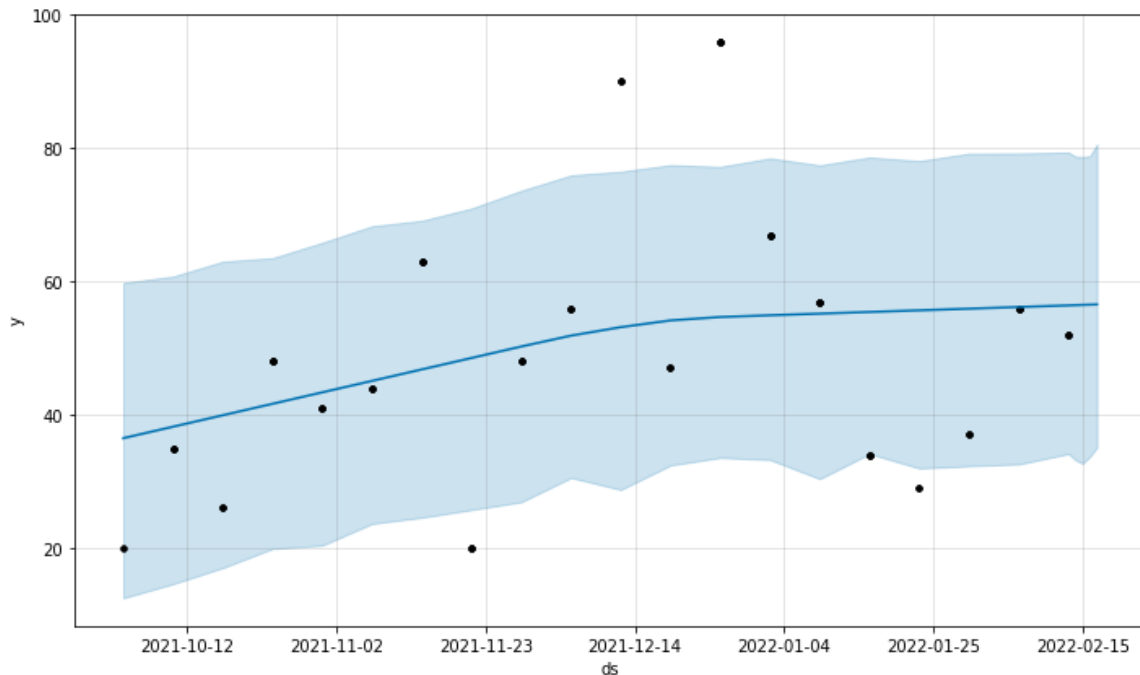
Figure 02: Prediction graph for specific itemcode with specific category(here blue line is the prediction line and black dots are real values and shaded blue area is the upper and lower y hat values)

- After training the model(with train set), We used the validation dataset for get the best hyperparameters for the model. So,we tried with various hyperparameters and checked the model with mean absolute percentage error score and then changed the main model parameters and do the submission with it's output
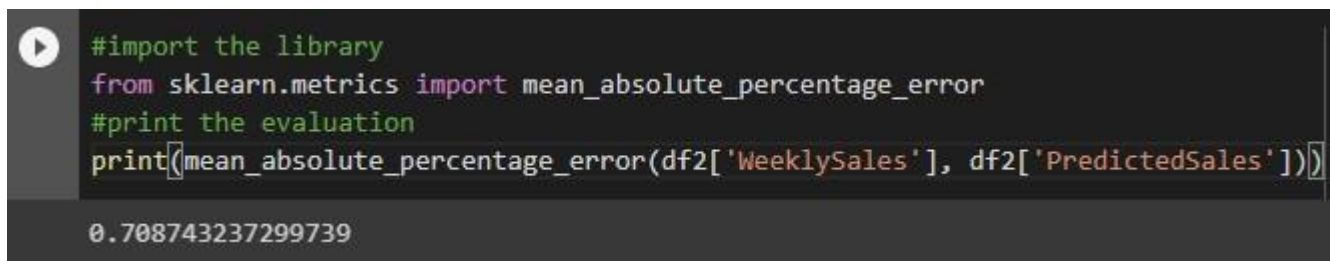
Feature Engineering Ideas

- We created the input dataset(train dataset) suitable for feed to the Prophet model by various preprocessing tasks and get the correct dataframe shape(filtering the specific itemcode with specific categories and made dataframes for each item)
- And with the help of validation dataset we built a good relationship between input features and target variable(weekly sales) by tuning the hyperparameters. By that we increased the performance of the Prophet model

Final Model and How It was reached

- First of all, we choosed the Prophet model for this sales forecasting problem because, it can automatically find seasonal trends with a set of data and offers easy to understand parameters
- We checked various hyperparameters with the validation dataset(for Prophet model) and checked the MAPE score and get the hyperparameters with least MAPE score and put that hyperparameters for the original model and get the predictions for this problem

Evaluation metrics

Evaluation done using the Mean Absolute Percentage Error(evaluation done on the validation dataset which was provided). And there was a scikit-learn library to get that.

```python
#import the library
from sklearn.metrics import mean_absolute_percentage_error
#print the evaluation
print(mean_absolute_percentage_error(df2['WeeklySales'], df2['PredictedSales']))

0.708743237299739
```

Figure 03: Evaluation code segment(MAPE)

Business Insights

- This kind of timeseries forecasting helps business to understand timely patterns in the data and analyse trends in business metrics
- They can monitor the demand of a particular product,brand and category with this kinds of time series forecasting
- Business owners can get the trend using charts like(Figure 02) , if their items selling in a good or bad way(upward and downward trending)
- By that they can arrange their stock inventories
- And it helps for better management of warehouse
- They can mark down and mark up with changing seasons and coming festivals(seasonal performance of their products)