Candidate Project

Ecommerce Analytics – Data Scientist

By

Kunal Niranjan Desai

Task: Inventory Management

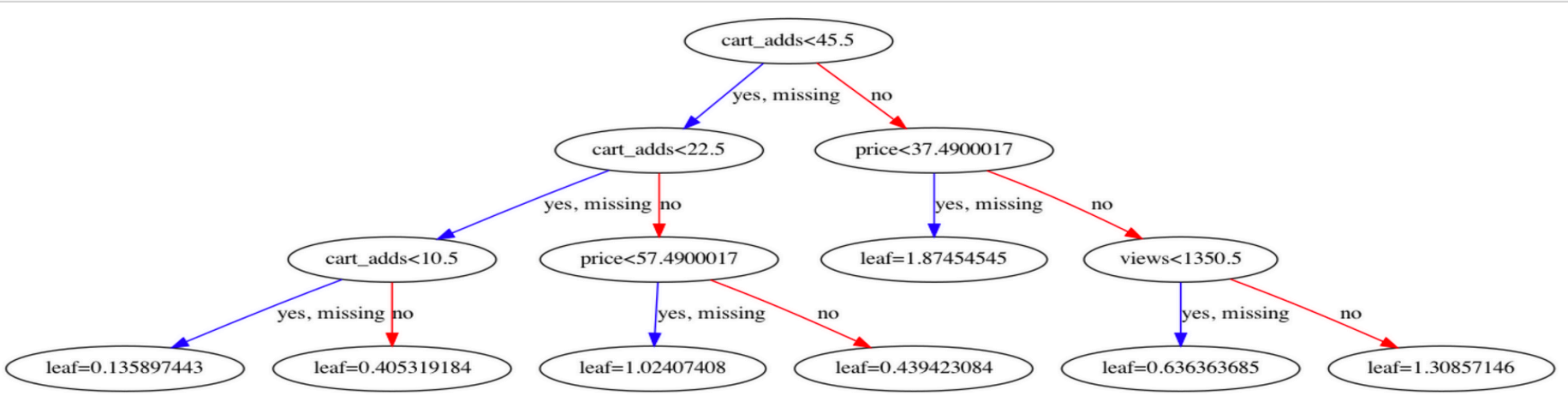
- Built a XGBoost model to predict the number of orders based on features: products, views, cart-adds, price, inventory and dayofyear (calculated from the week column).

- This model can be used by inventory analysts to forecast the number of orders and make sure the number of items of the particular product is available in the inventory for that day of the year.

- It is also useful to predict the spike in the number of orders due to a sale like Black Friday or Cyber Monday since the day of the year was taken into account during the model building.

Results

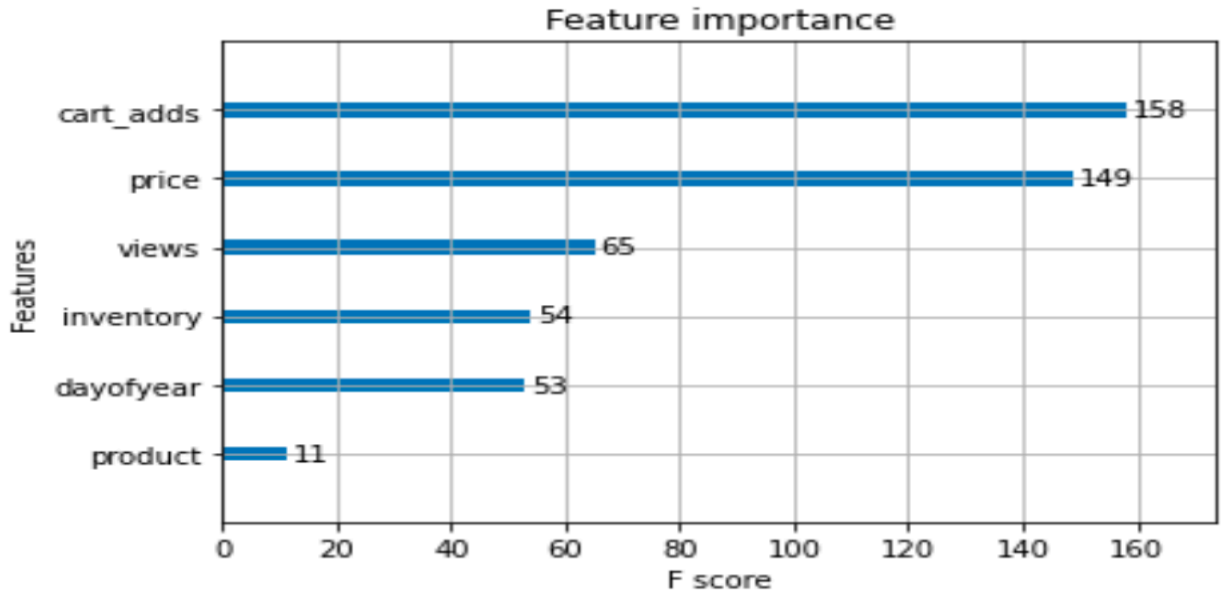
- The model's best hyperparameters were selected using GridSearchCV.
- The model was also evaluated using RMSLE (Root Mean Squared Logarithmic Error) and achieved an error value of just 0.169
- Here is one of the trees in the trained model:



Findings

Feature Importance:

- Based on the given dataset, it was found that the “cart_adds” feature played the most important role in predicting the number of orders.
- The “product” feature was the least important which suggests that all the 5 products influenced the number of orders in a similar way.
- The following is plot shows the detailed feature importance:



Thank you