PREDICTING THE FUTURE SALES

Group 22 (Batch 1)

- Amaan Ranapurwala
- Kishan Mishra
- Jayesh Chavan

Introduction:

Predicting sales of a company needs time series data of that company and based on that data the model can predict the future sales of that company or product. So, in this research project we will analyze the time series sales data of a company and will predict the sales of the company for the coming quarter and for a specific product.

Purpose:

To predict the Maximum Sales in future by using different Libraries & Datasets.

Methodology:

The Libraries used in the project are Numpy, Pandas, Seaborn, Lightgmb, LGBM Regressor, Matplolib.pyplot, Plotly.express & Train_test_split. Also the datasets we have used are Items_categories.csv, Items.csv, Sales_train.csv, Shops.csv &Test.csv

In this research, The LGBM Regression model were trained and tested for our dataset. The raw data is undergoes for feature selection and feature extraction. After that we have applied LGBM regression models for the training dataset to train the model. This train model was then tested on test dataset and validation dataset for checking the accuracy of the model.

By using padas we have cleared the data. Also we have taken out Monthly sales & Yearly sales. We have detected most items sold per category, items sold most in a particular day of the month & busiest days of the shop. With using the NumPy library we have taken out in which month people spend more money, what of item category is famous among the customer also which are the famous shops. The Prediction of future sales is done by the help of LGBMRegressor.

Analysis:

By analyzing the data the Items are sold in December 2013 was highest during 2013-2015.The 2013 was the year were the shops was busiest in between 2013 to 2015.

In December the shops are busiest compared to the whole year. At the 1st day of month the shops are busiest compared to the whole month.

Prediction:

After Processing the data, By using LGBMRegressor we have predicted that in the mouth of November 2015 the sale well be Maximum.

Conclusion:

We have practically learn how to use libraries such as LGBM, Numpy, Pandas, Seaborn, Lightgmb, Matplolib.pyplot, Plotly.express & Train_test_split. We have successfully predicted the maximum probability of sales in the year.

Evaluation Measures:

Measures such as Classification error, Computational cost, Accuracy can be used for calculating the accuracy of days of highest sales using neural network.

Evolution & Scope:

By using different libraries we will get to know What item category will be most famous among the customers, Which will be the most famous shops,

Contact:

- Amaan Ranapurwala - amaanr.csed19@sbjit.edu.in
- Kishan Mishra - mishrakishan2017@gmail.com
- Jayesh Chavan - jbchavan9421@gmail.com

Reference:

http://lightgbm.resdthedocs.io/en/latest/pythonapi/lightgbm.LGBMRegressor.html

http://www.kaggle.com/c/competitive-data-science-predict-future-sales

http://jupyter.org/documentation