

A Research paper on Trend Shopping Web application

Siddharth Sarkar, Diwakar Yagyasen, Ojaswi Gupta, Sakshi Yadav, Swati Singh, and Suchi Tiwari

Siddharth Sarkar is with CSE, BBDNITM, Lucknow, Uttar Pradesh, India.

Abstract—Abstract—The main objective of this web application is to help sellers in predicting the estimated selling price of the product using the database and of sold products. Using this the sellers can manage stocks and profits and can decide whether they wish to sell this product and at what rate they should sell it. Also in this application users can log in as shopkeepers and regular users. The regular user can buy products and sell second hand products while the shopkeeper can sell products. The regular customer can use prediction algorithm while selling the second hand product.

Languages used :-

Python for data analysis, data cleaning , data transformation, machine learning, data visualization and scripting.

HTML for designing web pages. Framework used:- Django. Database used:- dbsqlite3(Django's default data base) Libraries used:-Numpy, Pandas, Matplotlib, Seaborn and Scikit Learn.

I. INTRODUCTION

Prediction plays an important part in our lives. Without prediction, we wouldn't be able to forecast the future events and could not plan for them accordingly. Prediction in general terms means that we can estimate the future events or future values, We use prediction in our daily lives for example we predict the budget for future month by analyzing previous month budgets. In the same way sellers and producers also need this prediction to predict their success of their product. If a seller would know at what price a particular product would be sold at then they could actually estimate what their profit would be and thus accordingly decide at what price they should sell the product and whether they should sell the product or not thus saving them a huge money which would be wasted on buying the product. Also, this will prevent overpricing and under pricing the product thereby indirectly helping the customer as well.

This project aims at providing the correct prediction of prices for the sellers and helping them. The project contains a Django web application which contains two zones the shopkeeper and regular customer. User has a choice to log in as any option. As a shopkeeper the user has an option for selling products just like any e-commerce application, but they have a price prediction algorithm which will help them in managing prices, profits and stocks.

I. TECHNICAL DETAILS

The project is programmed in python language as it is a popular language and can be used for multiple uses. In This Project python is used for Scripting, data analysis, data cleaning and machine learning. Also, python has a very good web framework Django which can be used for creating web applications with ease and it has its own database dbsqlite3 so there is no issue for database connectivity.

The dataset used for learning my model is amazon's dataset from data.world site.

The dataset contains 17 columns and 21617 missing values which is a huge value so this needs to be cleaned.

Not all columns will have an equal importance on the target attribute price thus unnecessary columns are dropped.

The columns in prediction used are :-price, average_review_rating, amazon_category_and_sub_category and manufacturer.

The relationship between these attributes and target attributes are shown in charts given.

By these charts we can select attributes for predicting target attribute price.

For the dataframe Pandas library is used.

For analysis and visualization Matplotlib and Seaborn library is used.

For implementing this algorithm Scikit Learn library is used and the algorithm used is Random Forest Regressor which has better accuracy than linear regression and can work with categorical values.

II. CONCLUSION

The conclusion of this project is to help the sellers in managing the prices, profits and stocks so that the problems of overpricing and underpricing do not occur.

III. ACKNOWLEDGMENT

We would first like to thank our instructor and director for letting us choose this project. We would like to thank our project coordinator prof. Diwakar Yagyasen for guiding us in this project. We would like to thank our teachers for their time and help and would also thank our parents for supporting us.

APPENDIX A APPENDIX

Python:- An interpreted programming language which can be used for scripting as well.

Django:- A web framework in python for designing web applications.

Pandas:- A library of python used for dataframe operations.

Machine Learning:- A process of training a machine to do a certain task without explicitly programming it.

Linear Regression:-A process of fitting the training examples using a linear function.

Random Forest Regressor:-A machine learning algorithm in which multiple decision trees are used for prediction.

Scikit-Learn:-A library in python used for machine learning algorithms.

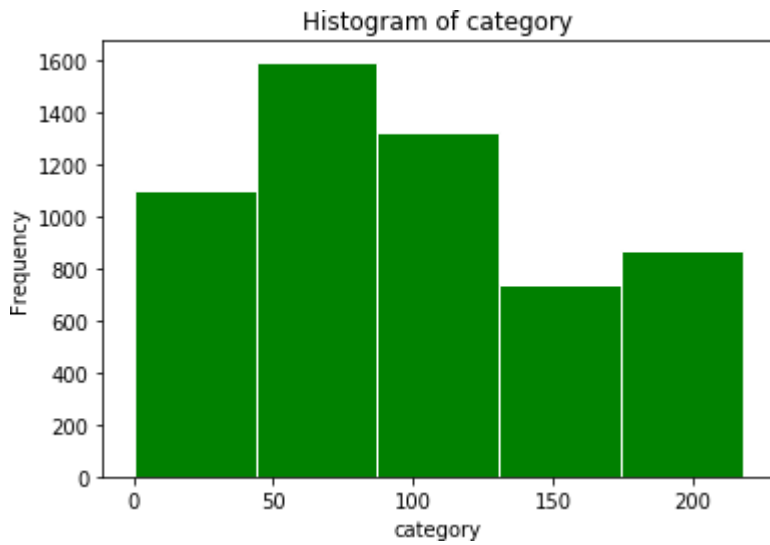


Fig. 1. 1. Histogram of category vs frequency

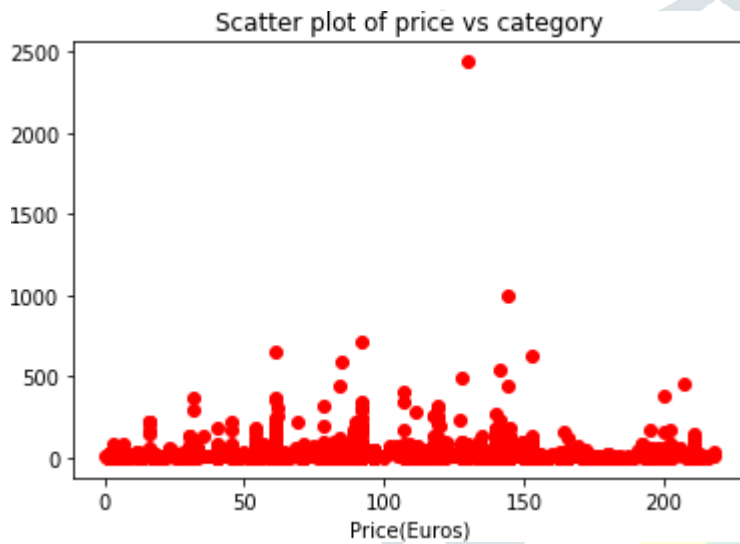


Fig. 2. 2. Scatter plot of price vs category

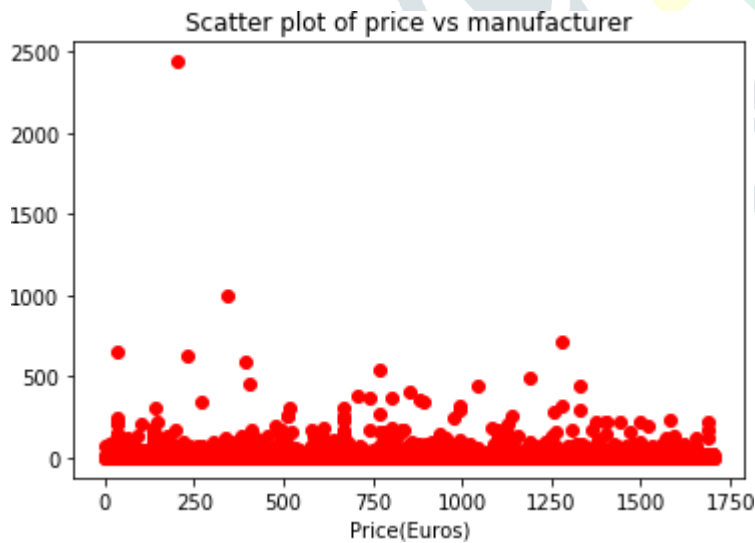


Fig. 3. 3. Scatter plot of price vs manufacturer