E-Commerce Fraud Detection System

SAMBIT S. MONDAL[1], SOUMYADEEP MAJI[2], SUBHOSHRI GHOSH[3], VIKASH KR. SONI[4], SHRUTI KARMAKAR[5], SNEHASHIS BALL[6], SOURAV KANTI MAITY[7]

[1] *Supervisor, Asansol Engineering College*

[2, 3, 4, 5, 6, 7] *Member, Asansol Engineering College*

Abstract— The supply of Used products returned to the companies is increasing. The situation is very alarming. It is the time for the world to be educated on how to utilize goods honestly and sincerely. The improvement in decreasing these types of fraud to the companies will lead to a better economic growth.Working and implementing this type of ML model with its technologies can result the big ecommerce companies like Amazon, Flipkart, Myntra, etc to get free from the fraud activities made by their customers.

I. INTRODUCTION

Our Project consists of 4 phases: First include the data collection. Second include ML model to fit in our dataset. Third include the integration of Frontend / UI and lastly includes backend and database to provide a user-friendly getup.

To build a model for predicting the final value of product we have applied machine learning techniques.

We created a heruko (Platform as a Service) platform-based survey page which was left for a Month and people came to give valuable reports which enumerated our datasets.

After proper KDD (Knowledge Discovery of Datasets), we found that Linear model is the best plausible approach.

Now, Integration with Angular (UI), Spring-Boot (UX) and MySQL(database) to give the proper output.

The technology stack of our Car Price Prediction System is:

- Material UI and Bootstrap
- FrontEnd (Client Side):
  - Angular
  - Typescript & Javascript
  - HTML & CSS
- Backend (Server Side):
  - Spring Boot, Hibernate (ORM)
- Database
  - MySQL CLI
- Machine Learning

II. EASE OF USE

A. Front end

While running the page, the first page which we are getting on screen is our homepage which consists of 4 components Navbar, Hetero section, Cards, Footer.

B. Back end

Create a object Of Model/Entity of Spring Boot Application. Then create a Repository for the Model which will directly linked to database. Next comes the creation of Service Interface and its implementation class which will be responsible for the creation of methods for Repositories. Handle all the exceptions using exception handler . Data is sent via Post Method(Rest API) and fetched through Get Method. Creation of @RestController which controls all the method with its respective Rest API's. Configure the Application. properties file so that spring-boot should connects with Hibernate JPA (ORM tool) then hibernate will add those objects to MySQL. Establishment of the connection between spring-boot and ML model. After the final price prediction , all values and the predicted value are stored in the database. The predicted value is fetched from database to client side.

C. Data Collection

We have created a fake dataset which contains various factors depending on the fraud that is conducted everyday on E-Commerce websites.

D. Data Pre processing

Raw data is created using faker library . Then valuable data is extracted from the raw dataset.

After it is processed through different steps like Null value extraction , categorization , missing value detection , feature extraction and deletion to get efficient dataset For our Ml model.

- What is Linear regression?

A Linear regression analysis is used to predict the value of a variable based on the value of another variable. The variable you want to predict is called the dependent variable. The variable you are using to predict the other variable's value is called the independent variable.

E. Model Selection

At first dataset is divided into test and train set. The we select some Ml algorithms for our dataset. After train set is trained by ML Algorithm. After training Model is tested by test data set. Then we detect the accuracy for different techniques. The model which obtain better accuracy, is integrated with server side for predicting the price.

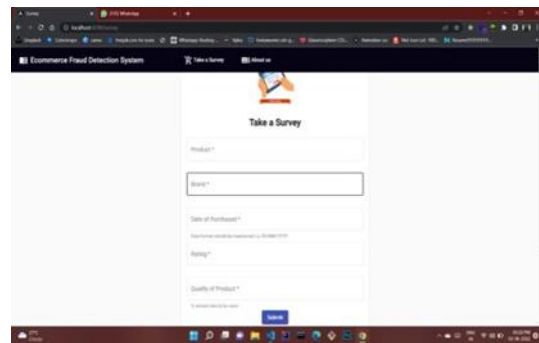Training algorithm used - Linear Regression

F. Testing Our Model
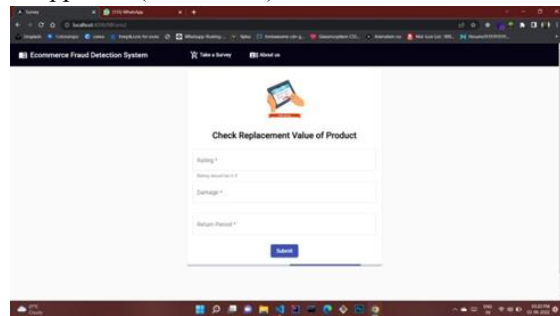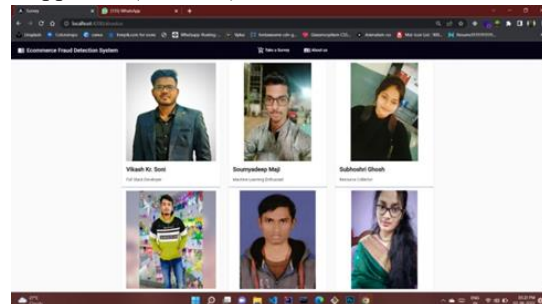
- Url Mapping using Postman Software

1. app.route("/")

2. app.route("/survey/")

3. app.route("/fillForm")

4. app.route("/about")

- ML model testing
- Whole dataset is divided into training and testing dataset in the ratio of 80:20.

- Algorithm is then applied to the testing data and since we are using linear regression the mean squared error is calculated.

G. Comparative analysis

In our project we have used Postman app to check our application.

The URL is pasted in the box and selecting the method we click on send box.

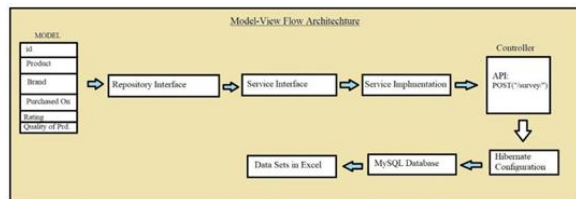According to the server https status is shown on the screen .

Otherside we find the accuracy of our model applying some in-built function in sklearn.

III. DEFINING THE DATA STORAGE

We have used MYSQL CLI as our database.

IV. FLOWCHART

V. RESULTS

Home Page

Form Page

Result

VI. CONCLUSION

The e-commerce transaction fraud dataset is a database that has a class imbalance.

Recent research about fraud detection in e-commerce transactions still determine feature extraction, purpose of this paper is to find the best model to detect fraud in ecommerce transactions.

Used and Value extracted products are returned to the companies without proper payment for the trail.

Fraud detection research on e-commerce is only limited to the determination of features or attributes which will be used to determine the nature of fraud or non-fraud transactions in e-commerce

This project undertakes a systematic ML model with proper UI and a powerful backend to determine the final value of product to create a balance between provider and user

REFERENCES

[1]   https://www.kaggle.com/datasets/zanjibar/100-million-data.csv

[2]   https://www.pdfdrive.com/machine-learning-step-by-step-guide-to-implement-machine-learning-algorithms-with-python-e158324853.html

[3]   https://f1000research.com/browse/articles?utm_source=CPB&utm_medium=cms&utm_campaign=JQB18928

[4]   https://scikit-learn.org/stable/datasets/toy_dataset.html

[5]   https://careers.uw.edu/blog/2021/10/05/21-places-to-find-free-datasets-for-data-science-projects-shared-article-from-dataquest/

[6]   https://data.world/solutions/data-mesh/