A

Mini Project

On

**ANALYZING CUSTOMERS PERSONALITY FOR
BUSINESS IMPROVEMENT USING MACHINE LEARNING**

(Submitted in partial fulfillment of the requirements for the award of Degree)

BACHELOR OF TECHNOLOGY

In

COMPUTER SCIENCE AND ENGINEERING

By

NEELAM PRIYANSHA (197R1A0539)

PASUMARTHI PRASHANTHI(197R1A0542)

ZEBA UNNISSA(197R1A0560)

Under the Guidance of

Dr. BAGAM LAXMAIAH

(Associate Professor)

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
CMR TECHNICAL CAMPUS

UGC AUTONOMOUS

(Accredited by NAAC, NBA, Permanently Affiliated to JNTUH, Approved by AICTE,

New Delhi) Recognized Under Section 2(f) & 12(B) of the UGCAct.1956, Kandlakoya

(V), Medchal Road, Hyderabad-501401.

2019-2023

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

CERTIFICATE

This is to certify that the project entitled **"ANALYZING CUSTOMERS PERSONALITY FOR BUSINESS IMPROVEMENT USING MACHINE LEARNING"** being submitted by **NEELAM PRIYANSHA(197R1A0539), PASUMARTHI PRASHANTHI(197R1A0542),and ZEBA UNNISSA(197R1A0560)** in partial fulfillment of the requirements for the award of the degree of B.Tech in Computer Science and Engineering to the Jawaharlal Nehru Technological University Hyderabad, is a record of bonafide work carried out by them under our guidance and supervision during the year.

The results embodied in this thesis have not been submitted to any other University or Institute for the award of any degree or diploma.

Dr. Bagam Laxmaiah   **Dr. A. Raji Reddy**
(Associate Professor)   (Associate Professor)
INTERNAL GUIDE

Dr. K. Srujan Raju   **EXTERNAL EXAMINER**
   HOD

 Submitted for viva voice Examination held on _____

ACKNOWLEDGEMENT

Apart from the efforts of us, the success of any project   depends largely on the encouragement and guidelines of many others. We take this opportunity to express our gratitude to the people who have been instrumental in the successful completion of this project.

We take this opportunity to express my profound gratitude and deep regard to my guide **Dr. Bagam Laxmaiah,** Associate Professor for his exemplary guidance, monitoring and constant encouragement throughout the project work. The blessing, help and guidance given by him shall carry us a long way in the journey of life on which we are about to embark.

We also take this opportunity to express a deep sense of gratitude to the Project Review Committee (PRC) **Dr. Punyaban Patel, Ms. Shilpa, Dr.T. Subha Mastan Rao & J. Narasimharao** for their cordial support, valuable information and guidance, which helped us in completing this task through various stages.

We are also thankful to **Dr. K. Srujan Raju,** Head, Department of Computer Science and Engineering for providing encouragement and support for completing this project successfully.

We are obliged to **Dr. A. Raji Reddy,** Director for being cooperative throughout the course of this project. We also express our sincere gratitude to Sri. **Ch. Gopal Reddy,** Chairman for providing excellent infrastructure and a nice atmosphere throughout the course of this project.

The guidance and support received from all the members of **CMR Technical Campus** who contributed to the completion of the project. We are grateful for their constant support and help.

Finally, we would like to take this opportunity to thank our family for their constant encouragement, without which this assignment would not be completed. We sincerely acknowledge and thank all those who gave support directly and indirectly in the completion of this project.

NEELAM PRIYANSHA	**(197R1A0539)**
PASUMARTHI PRASHANTHI	**(197R1A0542)**
ZEBA UNNISSA	**(197R1A0560)**

ABSTRACT

Doing business now-a-days is not just sitting in a store and selling items, but with the advancements in technology most of the businesses are making a move towards e-commerce via which customers can buy the products at any time and from anyplace and their products get delivered at their door steps. Day by day customers and their needs are changing, and the type of products used according to them are also changing.

Customers are an essential factor for a business. The focus of modern grocery superstore business has been shifted to the customer-centric organization. Our project about Customers Personality Analysis is a detailed analysis of a company's potential customers. It helps a company in understanding its consumers, so that the company can change goods to meet the demands, habits, and concerns of various sorts of customers.

LIST OF FIGURES/TABLES

FIGURE NO.   FIGURE NAME   PAGE NUMBER

LIST OF SCREENSHOTS

TABLE OF CONTENTS

TABLE OF CONTENTS

1. INTRODUCTION

1. INTRODUCTION

1.1 PROJECT SCOPE

This project titled "Analyzing Customers Personality for Business Improvement Using Machine Learning." Studying customer behavior is an age-old problem. Though there are many studies to understand customer purchase patterns, they are not sufficient. Every business wants to keep their customers on the long-term basis, because bringing a new customer to the commercial enterprise is fee-in depth.

1.2 PROJECT PURPOSE

Our project about Customer Personality Analysis is a detailed analysis of a company's potential customers. It helps a company in understanding its consumers, so that the company can change goods to meet the demands, habits, and concerns of various sorts of customers.

1.3 PROJECT FEATURES

Daily clients and their needs also are converting, and so the sort of merchandise consistent with them to be modified.There are algorithms like SVN and XGBoost in system studying which goes upon the datasets amassed and primarily based upon the algorithms that make predictions about the customers and likes and dislikes and display them products according to it. We have used the feature engineering technique to achieve more accurate results. Our study shows that machine learning and feature engineering can be a legitimate tool for potential customer prediction. The abilities of ML were extensively examined and achieved in field of Language Translation[3], [4], Self Driving[5], Text Recognition[6], Image Recognition[7], [8] and Voice Recognition[9], [10].

2. SYSTEM ANALYSIS

2. SYSTEM ANALYSIS

2.1 INTRODUCTION

Analysis of System is the important section in the system development procedure. The system is studied to the minute information and analyzed. The system analyst plays an vital function of an interrogator, and dwells deep into the operating of the existing gadget. In analysis, an in depth take a look at of these operations achieved by using the system and their relationships within and outdoors the device is executed. A key question considered right here is, "what should be carried out to solve the problem?" The system is regarded as a whole and the input to the machine is identified. Once analysis is finished the analyst has a firm expertise of what is to be carried out.

2.2 PROBLEM DEFINITION

The focal point of contemporary grocery superstore employer has been shifted to the customer-centric enterprise employer. Clients are the most essential factor for an organization. A few clients can help the commercial company to generate extra seasoned fit compared to the others.A loyalty-prone consumer intends to stay with the supplier who can provide the first-class merchandise. Alternatively, a deal-inclined customer will always look for a higher offer from a competitor. Clients may be classified as profitable and unprofitable[1]. Bringing a brand new consumer to the commercial enterprise is fee-in depth, because it involves severe advertising and marketing techniques without having prior information about the customer. According to the investigation of Reichheld and Teal[2], a business can rise its profit up till 95% by rising 5% client retention.

2.3 EXISTING SYSTEM

The dataset is used for personality trait identification. Multi-label Naive Bayes classifier has been used to predict results[7]. In this work various features are extracted on the basis of gender, age, etc. Machine learning algorithms such as Naive Bayes, kNN, Random Forest have been over the dataset[8]. Big Five Inventory [8] is used for this work. Results of this work have been calculated by taking both supervised and unsupervised methods into consideration.

2.3.1 DISADVANTAGES OF EXISTING SYSTEM

- Expected accuracy level of 70% is not attained
- Method of gray prediction is not widely accepted
- Dataset used in this study is publicly not available

2.4 PROPOSED SYSTEM

To raise the revenue boundary and stay ahead of the competitors it's far essential to recognize customers purchase conduct. Distinctive business industries proposed special guidelines to discover the potentiality of a person based on statistical evaluation. In this project, we would rather recommend a machine learning method to pick out potential customers for a retail superstore. The paper proposed an engineered method to categorize potential clients, based on formerly recorded purchase behavior. Using this classification as floor fact, we then apply machine learning algorithms to find a pattern to predict potential customers with an accuracy of 99.4%.

2.4.1 ADVANTAGES OF PROPOSED SYSTEM

- Multilevel predictive model proposed in this study is unique approach which was not raised before.
- Increased sales
- More accurate sales forecasting

2.5 HARDWARE & SOFTWARE REQUIREMENTS

2.5.1 HARDWARE REQUIREMENTS

- System   : Intel Core i3
- Hard Disk   : 1 TB
- Input Devices   : Keyboard, Mouse
- Ram   : 4 GB

2.5.2 SOFTWARE REQUIREMENTS

- Operating system : Windows XP/7/10
- Coding Language : Python
- Tool   : Visual Studio Code

3. ARCHITECTURE

3. ARCHITECTURE

3.1 ARCHITECTURE

A system architecture is the conceptual model that defines the structure, behavior, and more views of a system.
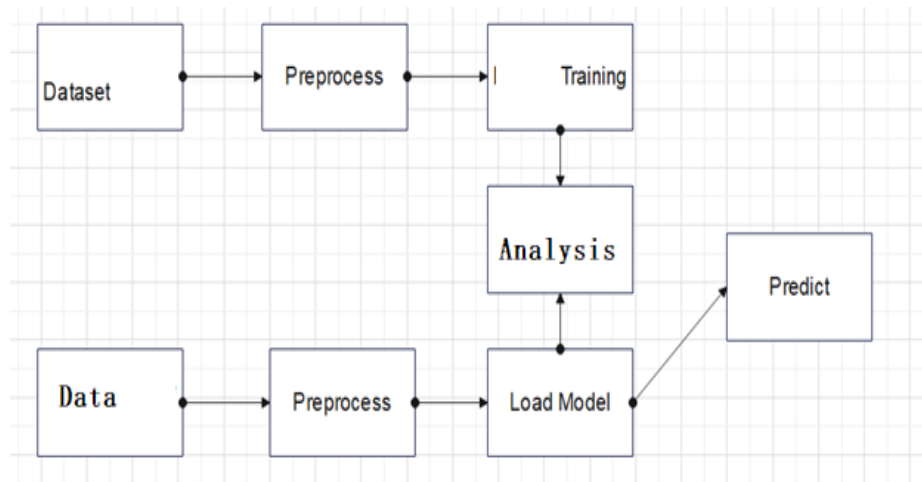

Figure 3.1: Project Architecture of Analyzing Customers
Personality for Business Improvement

3.2 DESCRIPTION

The three-tier software structure (a three-layer structure) emerged in 1990's to overcome the restrictions of the two-tier architecture. The third tier (middle tier server) is between the consumer interface (purchaser) and the data management (server) additives. This middle tier offers technique management where business logic and rules are executed and can accommodate hundreds of customers (compared to simplest one hundred users with the 2 tier structure) through presenting features consisting of queuing, software execution, and database staging. The 3 tier structure is used while a powerful distributed client/server layout is needed that provides increased overall performance, flexibility, maintainability, reusability, and scalability, even as hiding the complexity of disbursed processing from the user.

3.3 USE CASE DIAGRAM

A use case diagram in the Unified Modeling Language (UML) is a type of behavioral diagram to present a graphical overview of the functionality provided by a system in terms of actors, their goals which are represented as use cases, and any dependencies between those use cases. It is defined by and created from a Use-case analysis. The main purpose of a use case diagram is to show what system functions are performed for which actor.
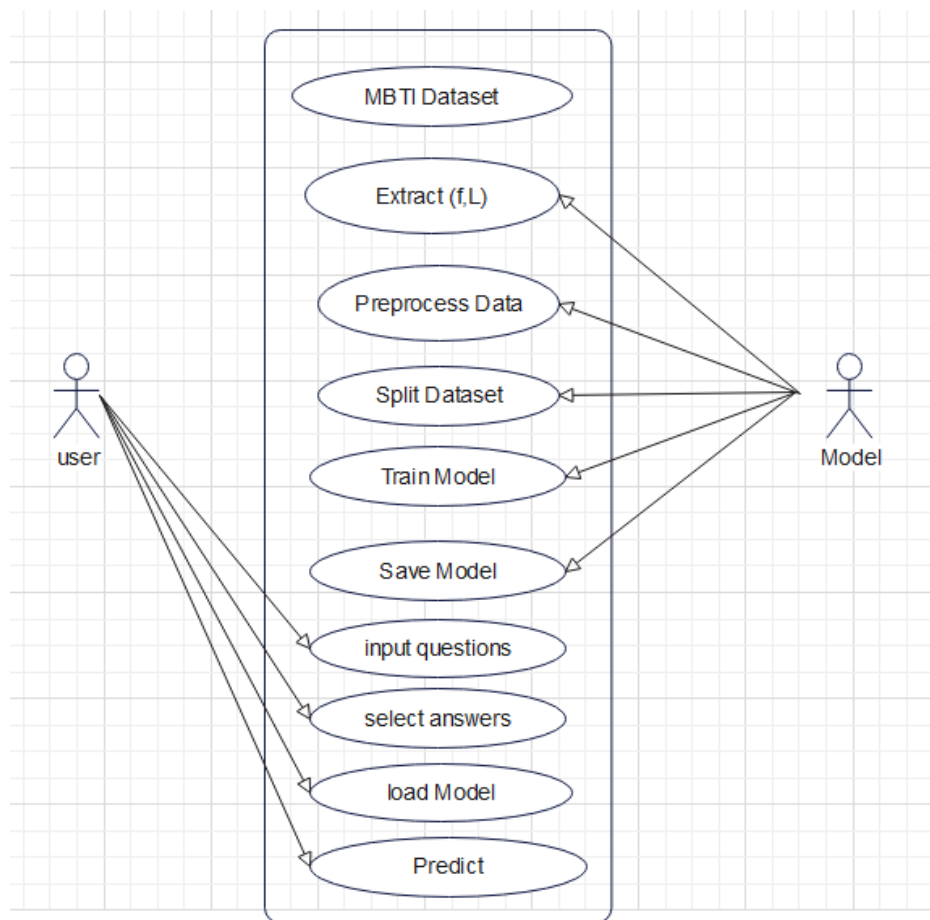

Figure 3.2: Use Case Diagram for Analyzing Customers
Personality for Business Improvement

3.4 CLASS DIAGRAM

A class diagram in the Unified Modeling Language (UML) describes the shape of a system by showing the system's classes, their attributes, strategies, and the relationships among the class. It explains which class carries data.
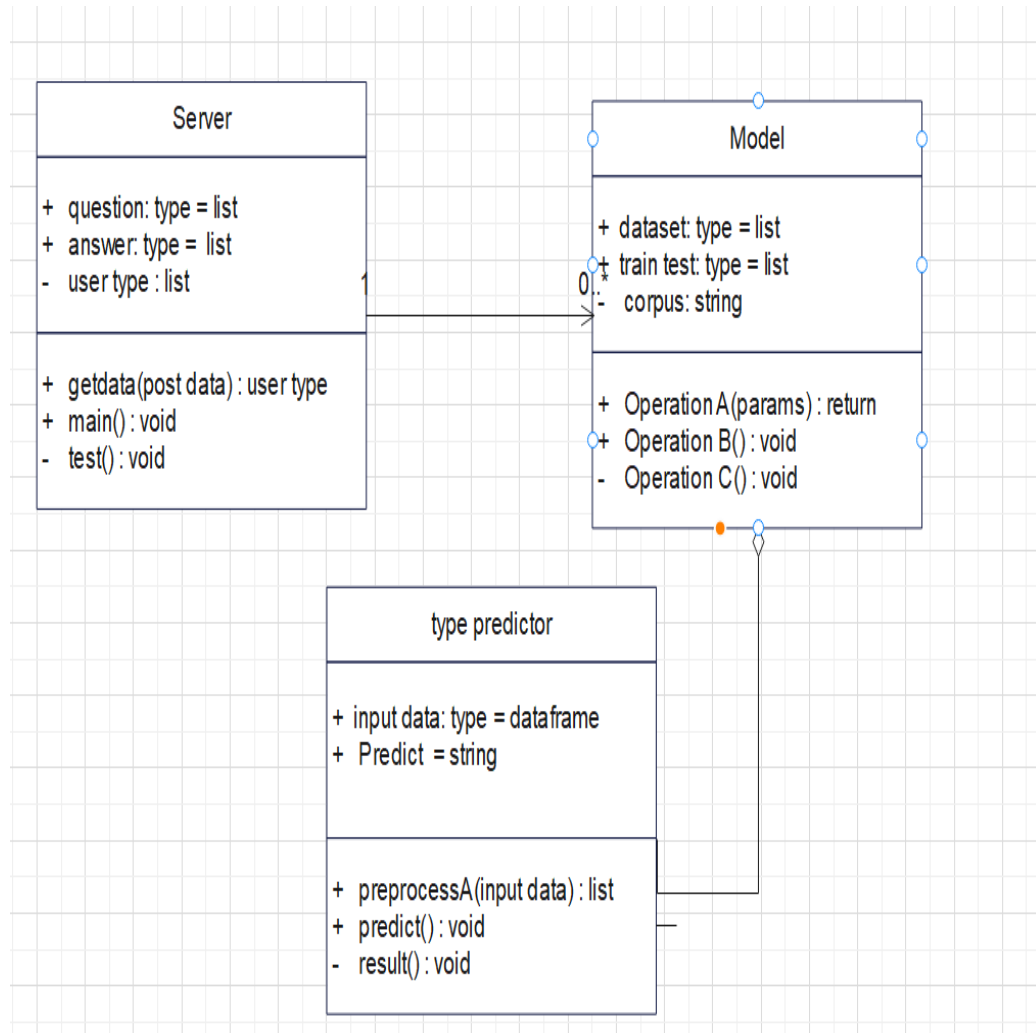

Figure 3.3: Class Diagram for Analyzing Customers
Personality for Business Improvement

3.5 SEQUENCE DIAGRAM

A sequence diagram in Unified Modeling Language (UML) are also called Event Diagrams, and Timing Diagrams. It is a kind of interaction diagram showing how processes operate with one another and in what order.
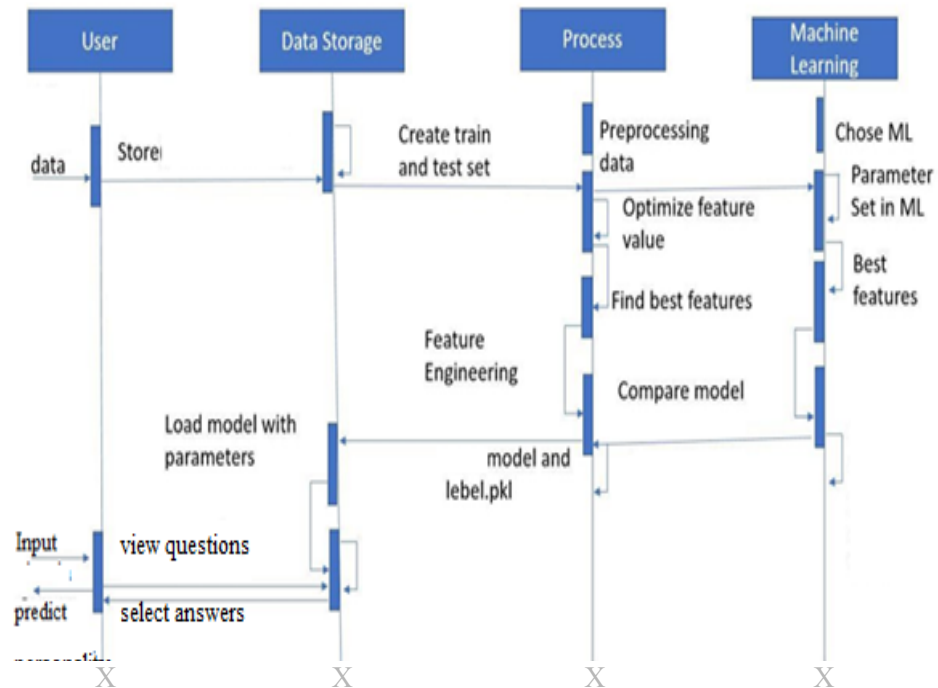

Figure 3.4: Sequence Diagram for Analyzing Customers
Personality for Business Improvement

3.6 ACTIVITY DIAGRAM

Activity diagrams within the Unified Modeling Language, describe the commercial enterprise and operational step-by-step workflows of components in a system. It represents workflows of stepwise activities and moves with assist for preference, iteration and concurrency. An activity diagram suggests the overall flow of control.
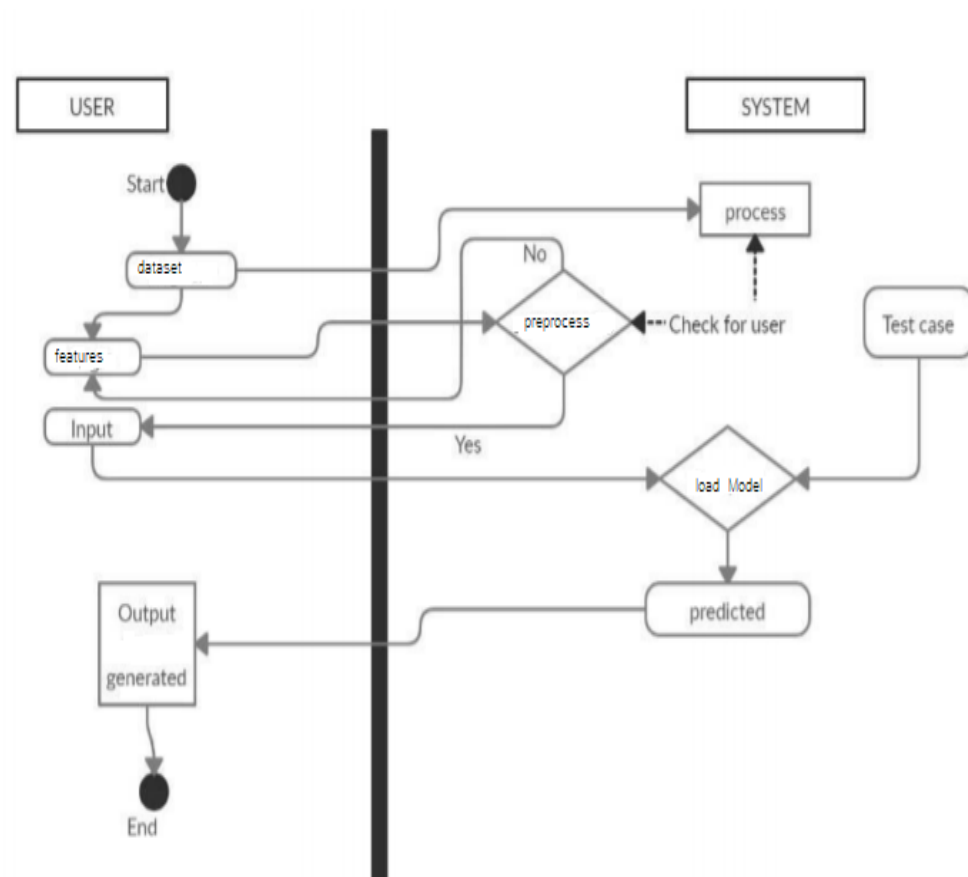

Figure 3.5: Activity Diagram for Analyzing Customers

Personality for Business Improvement

.

9

4. IMPLEMENTATION

4.1 SAMPLE CODE

```python
import webbrowser
import numpy as np
import torch
import torch.nn.functional as F
import torch.optim as optim
import torch.utils.data as data_utils
import torchvision
from PIL import Image, ImageFile
from torch import nn
from torch import optim as optim
from torch.autograd import Variable
from torch.optim import lr_scheduler
from torch.utils.data import DataLoader, Dataset
from torch.utils.data.sampler import SubsetRandomSampler
from torchvision import datasets, models, transform
import cv2
train_on_gpu = torch.cuda.is_available()
if not train_on_gpu:
    print('CUDA is not available.  Training on CPU ...')
else:
    print('CUDA is available!  Training on GPU ...')
ImageFile.LOAD_TRUNCATED_IMAGES = True
#!pip install --upgrade wandb
test_transforms = transforms.Compose([transforms.Resize(255),
                # transforms.CenterCrop(224),
                transforms.ToTensor(),
                ])
model = models.densenet161()
model.classifier = nn.Sequential(nn.Linear(2208, 1000),
                nn.ReLU(),
                nn.Dropout(0.2), nn.Linear(1000, 2),nn.LogSoftmax(dim=1))
criterion = nn.NLLLoss()
```

```python
# Only train the classifier parameters, feature parameters are frozen
optimizer = optim.Adam(model.parameters(), lr=0.001)
scheduler = lr_scheduler.StepLR(optimizer, step_size=7, gamma=0.1)
model = model.cuda()
model.load_state_dict(torch.load('tensorboardexp.pt'))
classes = ["accident", "noaccident"]
# model.load_state_dict(torch.load('tensorboardexp.pt')
)
count = 0
counts = 1
videopath = '12.mp4'
vid = cv2.VideoCapture(videopath)
ret = True
while ret:
    if ret == True:
        ret, frame = vid.read()
        try:
            img = Image.fromarray(frame)
        except ValueError:
            break
        except AttributeError
            break
        img = test_transforms(img)
        img = img.unsqueeze(dim=0)
        img = img.cuda()
        model.eval()
        with torch.no_grad():
            output = model(img)
        _, predicted = torch.max(output, 1)
            index = int(predicted.item())
            if index == 0:
cv2.imwrite(r"C:\xampp\htdocs\img\frame%d.png" % count, frame)
                count += 1
                if counts == 1:
```
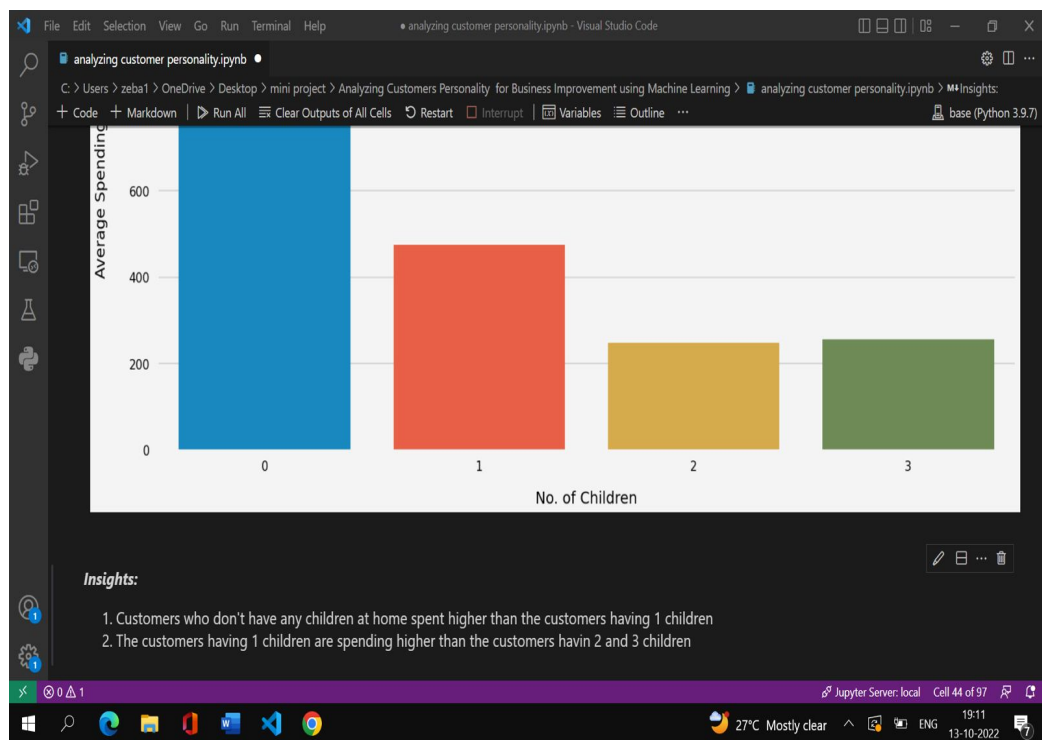
```python
                    webbrowser.open('127.0.0.1', new=2)
                counts += 1
        labels = 'status: ' + classes[index]
     cv2.putText(frame, labels, (10, 100),
             cv2.FONT_HERSHEY_DUPLEX, 2, (0, 0, 255), 5, cv2.LINE_AA)
     cv2.imshow('Frame', frame)
     if cv2.waitKey(1) & 0xFF == ord('q'):
        break
vid.release()
cv2.destroyAllWindows()
```
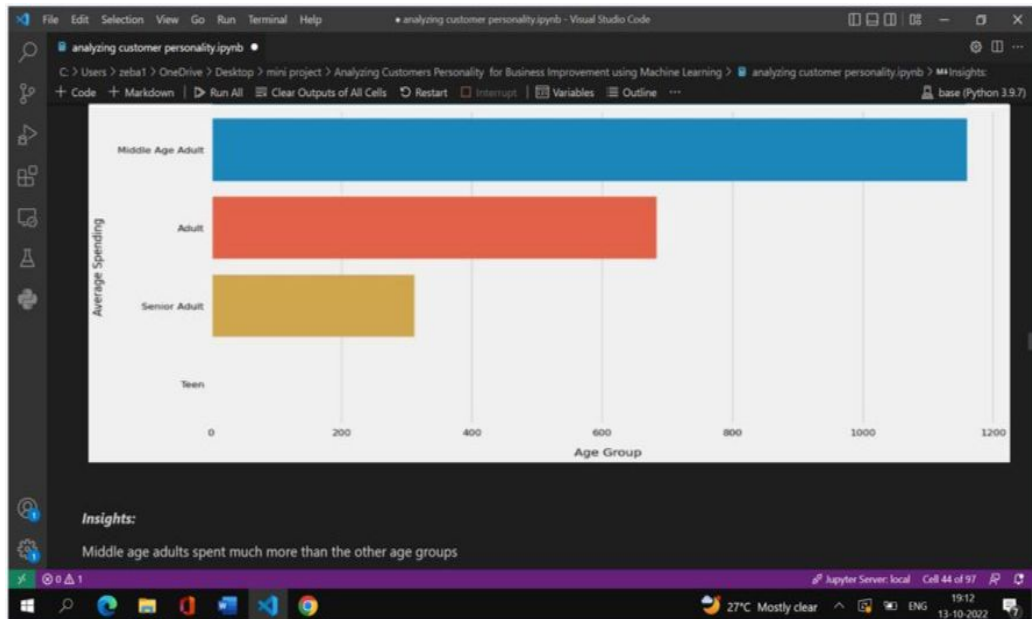
5. SCREENSHOTS
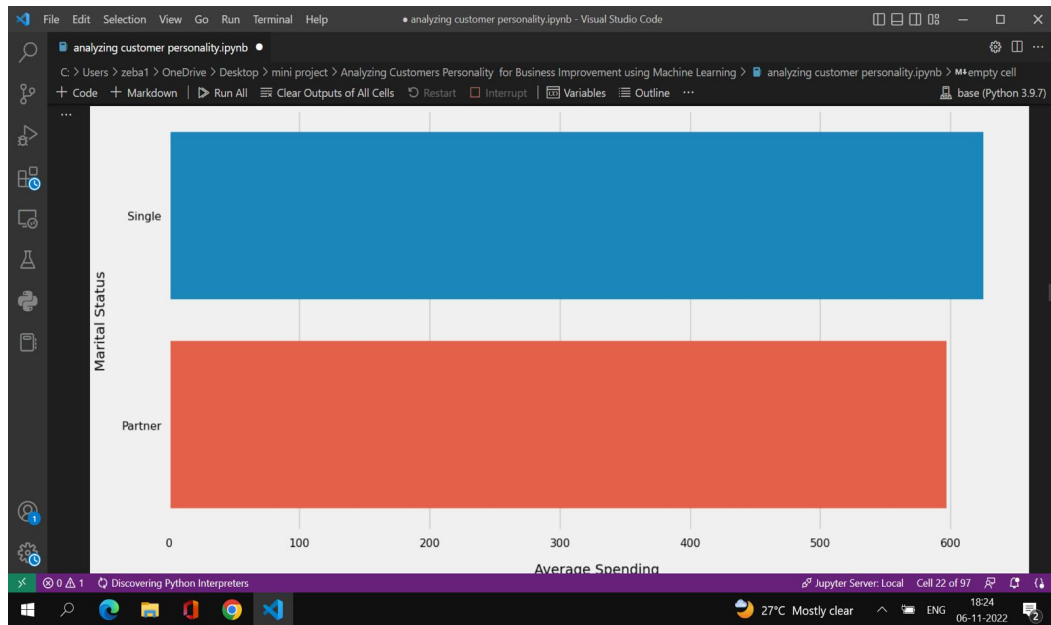
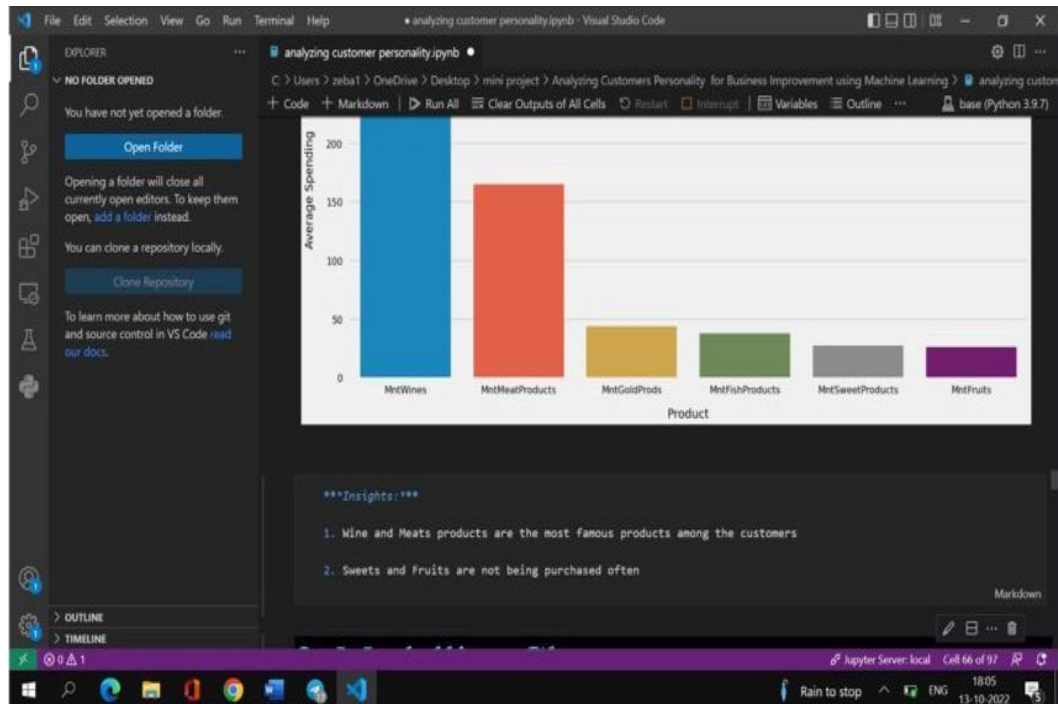Screenshot 5.1: Average Spending vs Marital Status

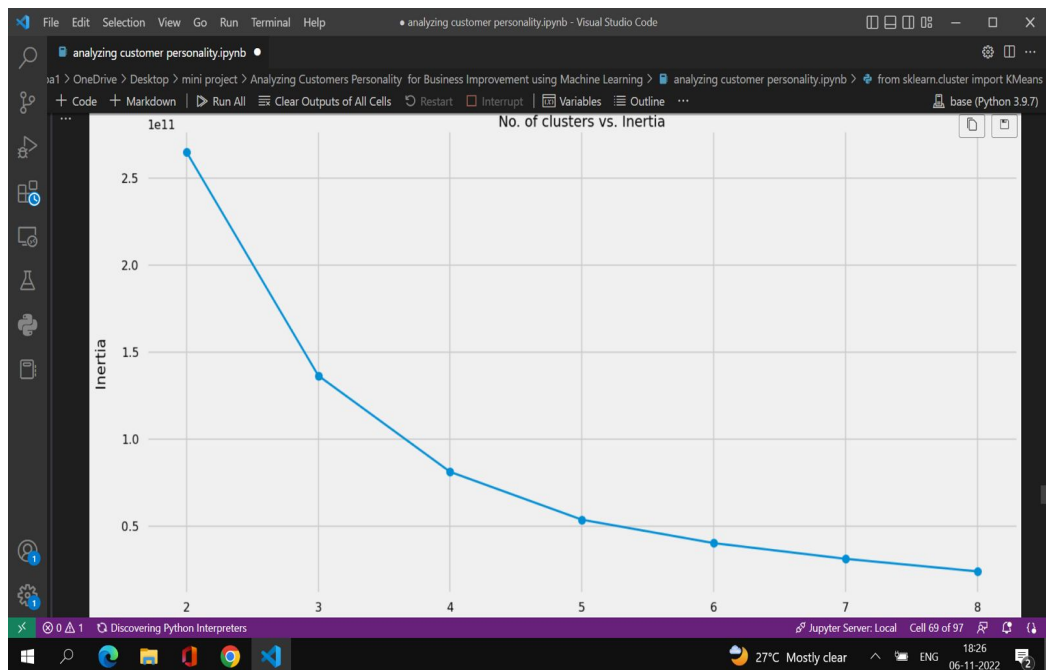Screenshot 5.2: No. of Children vs Average Spending

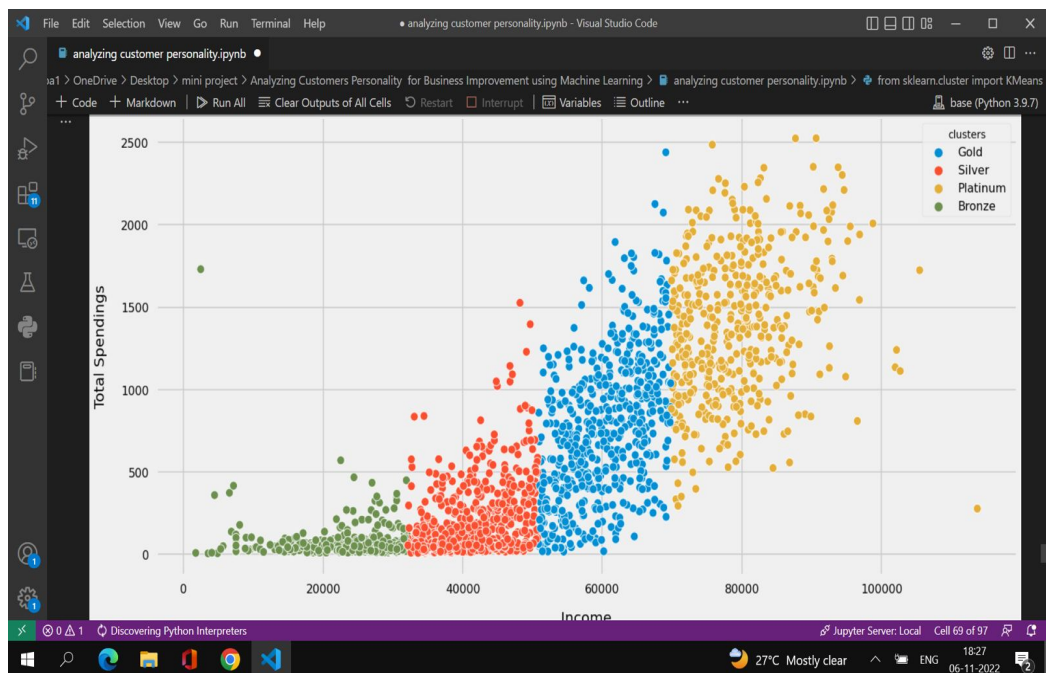Screenshot 5.3: Age Group vs Average Spending

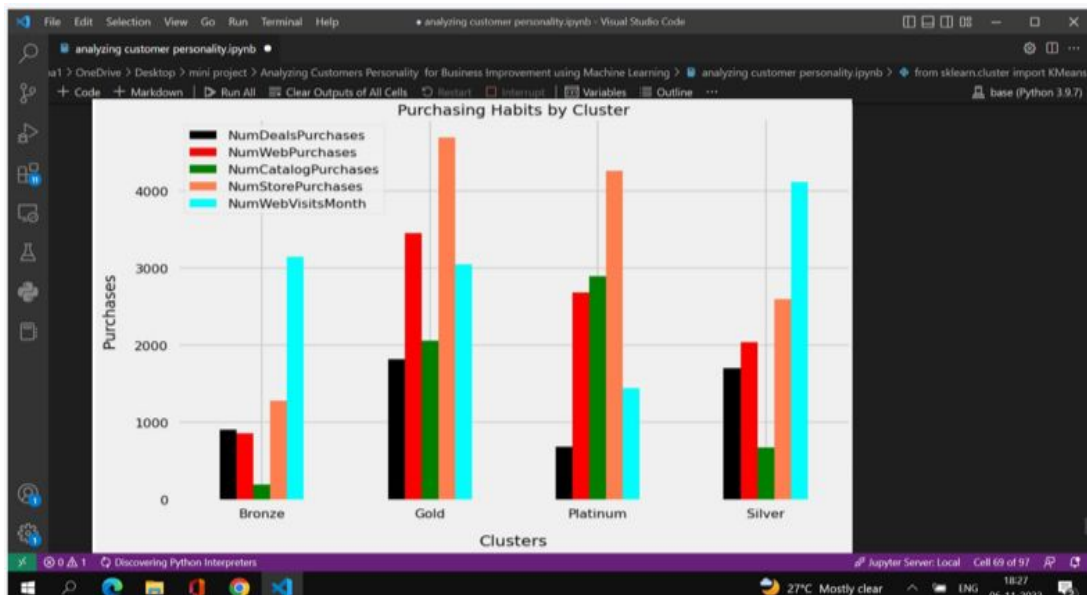

Screenshot 5.4: Relationship Income vs Spendings

Screenshot 5.5: Most Bought Products

Screenshot 5.6: Clustering Methods

Screenshot 5.7: Income vs Spending Clusters

Screenshot 5.8: Purchase Habits

6. TESTING

6. TESTING

6.1 INTRODUCTION

Testing is the process of finding faults and weaknesses in a working product. Testing is pleasantly performed while person development is requested to help in figuring out all errors and bugs. The pattern facts are used for checking out. It isn't the quantity but the quality of the information used for testing. Tested is aimed toward making sure that the system was correct and efficient earlier than live operation instructions.

Testing objectives:

The main objective of testing is to uncover a bunch of errors, systematically and with minimal time and effort. Stating officially, we will say, testing is a procedure of executing a program with the purpose of finding an error.

1. A very good test case is one that has the probability of locating an error, if it exists.
2. The software more or less confirms the quality and reliability standards.

6.2 LEVELS OF TESTING

Code Testing:

It is the examination of the logic of the program.

Speciation Testing:

Executing specification is beginning with what the program is ought to do, and how it should perform under diverse situations. Test cases for various situations, and combinations of conditions in all the modules are tested.

Unit Testing:

This is additionally called Module Testing. In unit testing every module is tested individually and integrated with the overall system. It particularly focuses on verification efforts which are accomplished on the smallest unit of the software layout within the module. The module of the system is tested one at a time. This testing is performed during the programming level itself. In the testing step each

module is observed to work satisfactorily as regard to expected output from the module. There are few validation checks for fields additionally.

Every module can be tested using the following two strategies:

- Black Box Testing
- White Box Testing

6.2.1 BLACK BOX TESTING

Black Box Testing is a technique for software testing wherein the capability of Software Under Test(SUT) is tested without looking at the internal code structure, implementation details and knowledge of internal paths of the software program. This sort of testing is based totally on the software requirements and specifications.

Black Box Testing mainly focuses on the inputs and outputs of the software system without considering the inner knowledge of the software program. The "black box" in "black box testing" symbolizes not being able to see the internal workings of the software, so that only the end-user experience can be tested.

 Types of Black Box Testing

The following are the prominent ones among many

Functional testing: This type of black box testing is performed by the software testers. It is related to the functional requirements of a system.

Non-functional testing: This type of black box testing is related to non-functional requirements such as performance, scalability, usability however not related to testing of a particular functionality.

Regression testing: Regression testing is done after code fixes , upgrades or any other system maintenance to check the new code has not affected the existing code.

6.2.2 WHITE BOX TESTING

White Box Testing is also known as clear, open, structural, and glass box testing. It is the testing of a software solution's internal coding and infrastructure. It focuses primarily on strengthening security, the flow of inputs and outputs through the application, and enhancing layout and usability. The clear box or whitebox name symbolizes the ability to see via the software's outer shell (or "box") into its inner workings.

7. CONCLUSION

7. CONCLUSION AND FUTURE SCOPE

7.1 CONCLUSION

Although there are many tactics proposed for discovering potential customers, a machine learning technique is an uncommon alternative. Our studies is first of its kind in which machine learning is used to observe patron's buy conduct for a retail superstore. The test of potential client classification achieved prediction accuracy as much as 94% with recall 98.9% and precision 99.7%. We engineered functions to capture the relationship between classes, items, amount, measurement unit, and income. The distinction in end result between native and engineered characteristic is 42.6%. A commercial enterprise can be rather benefited by way of identifying their potential client correctly. The potential patron may be approached with a customized advertising plan that could grow the sale of a business.

7.2 FUTURE SCOPE

Future scope would be to understand customers conduct using machine learning to meet changing needs, a product of interest, buying frequency which will help to make appropriate marketing plans, efficient supply chain management, and drives sales growth.

8. BIBLIOGRAPHY

8. BIBLIOGRAPHY

8.1 REFERENCES

[1] Berry, Leonard L." Relationship marketing of servicesgrowing interest, growing interest, emerging perspectives." Journal of the Academy of marketing science23, no. 4, pp. 236-245, (1995)

[2] Reichheld, Frederick F., and Jr WE Sasser. "Zero defections: Quality Comes to services." Harvard business review 68, no. 5, pp. 105-111,(1990).

[3]Collobert, Ronan, and Jason Weston. "A unified architecture for naturallanguage processing: Deep neural networks with multitask learning." In Proceedings of the 25th international conference on Machine learning,pp. 160-167. ACM, 2008.

[4] Goldberg, Yoav. " A primer on neural network models for natural language processing." Journal of Artificial Intelligence Research 57, pp.345-420, (2016).

[5] Ohn-Bar, Eshed, and Mohan Manubhai Trivedi." Looking at humans in the age of self-driving and highly automated vehicles." IEEE Transac-tions on Intelligent Vehicles 1, no. 1, pp. 90-104, (2016).

[6] Zhang, Xiang, Junbo Zhao, and Yann LeCun. " Character-level con-volutional networks for text classification." In Advances in neural information processing systems, pp. 649-657. 2015.

[7] Rosten, Edward, and Tom Drummond. "Machine learning for high-speedcorner detection." In European conference on computer vision, pp. 430-443. Springer, Berlin, Heidelberg, 2006.

[8] Chapelle, Olivier, Patrick Haffner, and Vladimir N. Vapnik. " Supportvector machines for histogram-based image classifictatiom." IEEE trans-actions on Neural Networks 10, no.5,pp. 1055-1064, (1999).

[9] Graves, Alex, Abdel - rahman Mohamed, and Geoffrey Hinton. " Speech recognition with deep recurrent neural networks." In Acoustics, speech andsignal processing (icassp), 2013 IEEE international conference on, pp. 6645-6649. IEEE, 2013.