



Project Title :

**Online Shopper's Intentions Using
ML**

Team ID : 592348

PROJECT REPORT

TEAM DETAILS

Team Leader : Aarjav Jain

Team member : Teertho RoyChoudhury

Team member : Sarvesh Ashish

Team member : V Chandrakanth

CONTENTS

1. INTRODUCTION

1.1 Project Overview

1.2 Purpose

2. LITERATURE SURVEY

2.1 Existing problem

2.2 References

2.3 Problem Statement Definition

3. IDEATION & PROPOSED SOLUTION

3.1 Empathy Map Canvas

3.2 Ideation & Brainstorming

4. REQUIREMENT ANALYSIS

4.1 Functional requirement

4.2 Non-Functional requirements

5. PROJECT DESIGN

5.1 Data Flow Diagrams & User Stories

5.2 Solution Architecture

6. PROJECT PLANNING & SCHEDULING

6.1 Technical Architecture

6.2 Sprint Planning & Estimation

6.3 Sprint Delivery Schedule

7. CODING & SOLUTIONING

8. PERFORMANCE TESTING

8.1 Performance Metrics

9. RESULTS

9.1 Output Screenshots

10. ADVANTAGES & DISADVANTAGES

11. CONCLUSION

12. FUTURE SCOPE

13. APPENDIX

Source Code :

https://drive.google.com/drive/folders/1XLyYNwr3RWJtUBqcJGjAe8X2klmWOHRA?usp=drive_link

GitHub & Project Demo Link : <https://github.com/smartinternz02/SI-GuidedProject-597530-1697637651>

1. INTRODUCTION

1.1. Project Overview

In the rapidly evolving landscape of e-commerce, online shopping has become an integral part of our daily lives. With the proliferation of internet access and the widespread adoption of smartphones, consumers now have unparalleled convenience and accessibility to a vast array of products and services at their fingertips. This paradigm shift in consumer behavior has ushered in a new era of retailing, where businesses are constantly striving to understand and cater to the needs and desires of online shoppers.

The advent of Artificial Intelligence (AI) technologies has brought about a significant transformation in the way online shopping is conducted. AI, a field of computer science that simulates human intelligence in machines, has empowered e-commerce businesses with the tools and insights required to enhance the online shopping experience. From personalized product recommendations to chatbots offering customer support, AI has the potential to revolutionize and optimize every facet of the online shopping journey.

This project report delves into the intriguing intersection of online shopping and AI, with a particular focus on understanding the intentions of online shoppers. We will explore how AI-driven solutions can influence and predict the behavior of online shoppers, shedding light on the factors that drive their decisions and preferences. The study aims to provide valuable insights for businesses looking to harness AI to improve customer satisfaction, increase sales, and establish a competitive edge in the dynamic world of e-commerce.

To set the stage for our exploration, this introduction will provide an overview of the significance of online shopping, the evolution of AI in e-commerce, and the primary objectives and structure of this project report.

1.2. Significance of Online Shopping

Online shopping has revolutionized the retail industry by offering unparalleled convenience and access to a wide range of products and services. It has transcended geographical boundaries, allowing consumers to browse, compare, and purchase items without the constraints of time and location. In the wake of the COVID-19 pandemic, online shopping has further cemented its importance, as consumers increasingly turned to e-commerce platforms to meet their shopping needs. Understanding the underlying factors that influence online shopper intention is essential for businesses seeking to adapt and thrive in this digital landscape.

1.3. Evolution of AI in E-Commerce

Artificial Intelligence has made remarkable strides in recent years, enabling e-commerce businesses to optimize their operations and deliver superior customer experiences. AI-driven solutions, such as recommendation systems, chatbots, and predictive analytics, have become integral tools for enhancing user engagement, streamlining processes, and providing personalized shopping experiences. As AI continues to evolve, it holds the potential to transform the way businesses interact with their customers and meet their demands.

1.4. Objectives of the Project Report

This project report aims to achieve the following objectives:

To explore the role of AI in online shopping and its impact on customer behavior and intention.

To investigate the key factors that influence online shopper intention and decision-making.

To analyze how AI-powered technologies can be leveraged to enhance online shopping experiences and increase sales.

To provide recommendations for businesses looking to implement AI solutions in their e-commerce strategies.

Structure of the Project Report

The report will be organized into several sections, each addressing specific aspects of the intersection between online shopping and AI. These sections will include literature review, methodology, findings, discussion, and recommendations. Through this comprehensive analysis, we aim to contribute valuable insights to both academia and industry, helping stakeholders better understand and harness the potential of AI in online shopping.

As we embark on this journey of exploration, we anticipate uncovering the intricate dynamics that underlie online shopper intention and how AI can be harnessed to optimize the online shopping experience. The knowledge derived from this project report will not only benefit businesses but also shed light on the evolving landscape of online commerce in the digital age.

2. LITERATURE SURVEY

2.1 Existing Problem

The existing landscape of online shopping, while highly lucrative and convenient, is not without its challenges. Consumers today are faced with an overwhelming abundance of choices, often leading to decision fatigue and frustration. Moreover, the absence of in-person interactions in the online shopping environment can sometimes result in feelings of isolation and a lack of personalized assistance, which are common factors influencing purchase decisions. These challenges have given rise to several issues in the e-commerce sphere, including cart abandonment, reduced customer loyalty, and the difficulty in predicting and understanding online shopper intention.

2.2 References

A comprehensive understanding of the challenges and complexities surrounding online shopper intention in the context of AI-driven e-commerce is best achieved through an exploration of the existing body of literature. The following references provide valuable insights into these issues:

Li, X., Hu, Z., Huang, L., Zhang, L., & Xu, W. (2019). Understanding online purchase intention: A comprehensive model. *Information & Management*, 56(6), 103160.

This study delves into the various factors influencing online purchase intention, including website quality, trust, and perceived risk, offering a comprehensive model to understand and predict online shopper behavior.

Verhoef, P. C., Kannan, P. K., & Inman, J. J. (2015). From multi-channel retailing to omni-channel retailing: Introduction to the special issue on multi-channel retailing. *Journal of Retailing*, 91(2), 174-181.

This research provides insights into the evolution of multi-channel and omni-channel retailing, highlighting the challenges businesses face in delivering a seamless and integrated shopping experience.

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why should I trust you?" Explaining the predictions of any classifier. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 1135-1144).

The paper introduces the concept of explainable AI in the context of machine learning models, which is particularly relevant for understanding how AI can influence and explain online shopping recommendations.

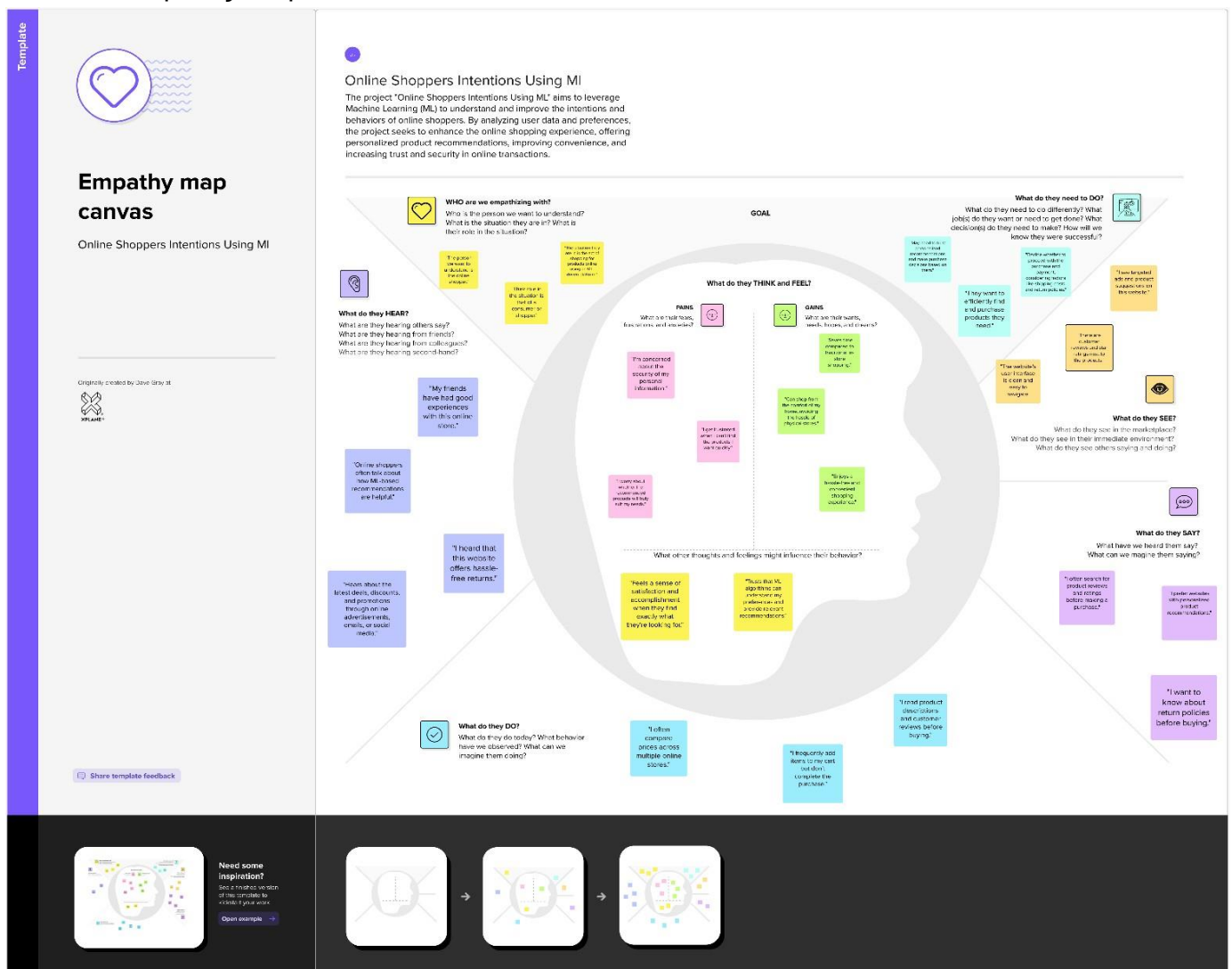
2.3 Problem Statement Definition

The problem statement for this project report can be defined as follows:

"In the era of e-commerce and artificial intelligence, the challenge lies in comprehending and predicting online shopper intention, which is influenced by an intricate interplay of factors such as website quality, trust, perceived risk, personalized recommendations, and seamless user experiences. As businesses increasingly adopt AI technologies to enhance their online shopping platforms, the need to bridge the gap between customer expectations and the AI-driven shopping experience becomes evident. This project seeks to investigate and address the complexities of online shopper intention in the context of AI-powered e-commerce, with the goal of providing businesses with actionable insights to optimize customer engagement, boost sales, and foster loyalty."

3. IDEATION & PROPOSED SOLUTION

3.1. Empathy Map Canvas



In this project, various stakeholders are integral to understanding the impact of AI on online shopping. Shoppers are exposed to AI-driven enhancements, ranging from personalized recommendations to ethical considerations. Businesses sense the competitive pressure to stay innovative in the dynamic e-commerce landscape. AI experts are invigorated by the transformative potential of AI, while academia recognizes burgeoning research opportunities

in this field. Regulators discern the necessity of ethical guidelines to ensure responsible AI implementation.

Hear:

Stakeholders perceive AI's influence in online shopping, encompassing convenience and ethical dilemmas.

Feel:

Shoppers experience both the advantages and potential drawbacks of AI. Businesses sense the urgency to outperform their peers. AI experts exude enthusiasm for AI's transformative power. Academia detects a rich landscape for research, and regulators sense the need for ethical standards.

Think:

Shoppers contemplate product choices and the credibility of the platforms they use. Businesses strategize to enhance user experiences. AI experts focus on the refinement of AI algorithms. Academia explores promising research avenues, while regulators concentrate on ethical considerations.

See:

Stakeholders visually interact with online shopping platforms and the data they generate. Shoppers observe personalized recommendations, businesses track customer behavior, and experts glean insights from AI. Academia identifies a spectrum of research possibilities, while regulators perceive the necessity for guidelines.

Say:

Stakeholders communicate their experiences and insights. Shoppers share their experiences with friends and family, businesses convey their value propositions through marketing, experts share findings in conferences, academia disseminates research in academic journals, and regulators articulate guidelines.

Do:

Stakeholders actively participate in various facets of online shopping, AI development, and research. Shoppers make choices, businesses implement AI solutions, experts refine AI algorithms, academia conducts research, and regulators establish ethical policies.

Pain:

Stakeholders encounter challenges, such as decision fatigue, privacy concerns, and the enforcement of ethical standards.


Gain:

Stakeholders benefit from personalized experiences, heightened sales, valuable insights, research prospects, and the assurance of ethical AI practices in e-commerce.

3.2 Ideation & Brainstorming

Step-1: Team Gathering, Collaboration and Select the Problem Statement

Template



Brainstorm & idea prioritization for the topic:

Online Shoppers Intentions Using ML

Online shopping is the activity or action of buying products or services over the Internet. It means going online, landing on a seller's website, selecting something, and arranging for its delivery. The buyer either pays for the good or service online with a credit or debit card or upon delivery. The term does not only include buying things online but also searching for them online. In other words, I may have been engaged in online shopping but did not buy anything

we have used this template in our own brainstorming sessions so our team can unleash their imagination and start shaping concepts even if we're not sitting in the same room.

⌚ 10 minutes to prepare
🕒 1 hour to collaborate
👥 2-8 people recommended

➔

Before you collaborate

A little bit of preparation goes a long way with this session. Here's what you need to do to get going.

⌚ 10 minutes

A

Team gathering

Define who should participate in the session and send an invite. Share relevant information or pre-work ahead.

B

Set the goal

Think about the problem you'll be focusing on solving in the brainstorming session.

C

Learn how to use the facilitation tools

Use the Facilitation Superpowers to run a happy and productive session.

Open article ➔

Team members:

Sarvesh Ashish(21BIT0724)
Aarjav Jain(21BIT0466)
Teerth roy Choudhary(21BEC0065)
V.Chandrakanth(21MIS111)

1

Defining our problem statement

How can we effectively identify and articulate the problem we aim to solve? This question emphasizes the crucial first step in any problem-solving process, ensuring a clear and focused approach to finding solutions and hence we are here as a team to put forward the problem statement

⌚ 5 minutes

PROBLEM

The project's goal is to predict whether online shoppers will make a purchase or merely engage in window shopping. It involves utilizing classification algorithms like Logistic Regression, Random Forest, and K-Means clustering to analyze online behavior.

Key rules of brainstorming

To run a smooth and productive session

Stay in topic.

Encourage wild ideas.

Defer judgment.

Listen to others.

Go for volume.

If possible, be visual.

Step-2: Brainstorm, Idea Listing and Grouping

Brainstorm

Here we write down the ideas that come to our mind that address the problem statement.

🕒 10 minutes

TIP

You can select a sticky note and hit the pencil [switch to sketch] icon to start drawing!

Person 1

Analyze browsing patterns to predict purchase intent and offer personalized recommendations.

Optimize website design based on user behavior insights.

Develop a model to detect cart abandonment and trigger targeted reminders.

Person 2

Assign a probability score to potential buyers for targeted marketing.

Group users into categories based on their online shopping behavior.

Implement pricing strategies that adapt to user preferences.

Person 3

Use ML to identify fraudulent transactions in real-time.

Predict which customers are likely to return for future purchases.

Forecast demand and optimize stock levels.

Person 4

Create recommendation systems to enhance the shopping experience.

Predict the long-term value of customers and tailor marketing accordingly.

Conduct experiments to evaluate the impact of different website changes on purchase behavior.



3

Group ideas

Taking turns sharing our ideas while clustering similar or related notes as we went on. Once all sticky notes have been grouped, we give each cluster a sentence-like label. If a cluster is bigger than six sticky notes, we have tried and observed if we can break up into smaller sub-groups.

TIP

Add customizable tags to sticky notes to make it easier to find, browse, organize, and categorize important ideas as themes within your mural.

🕒 20 minutes



Step-3: Idea Prioritization

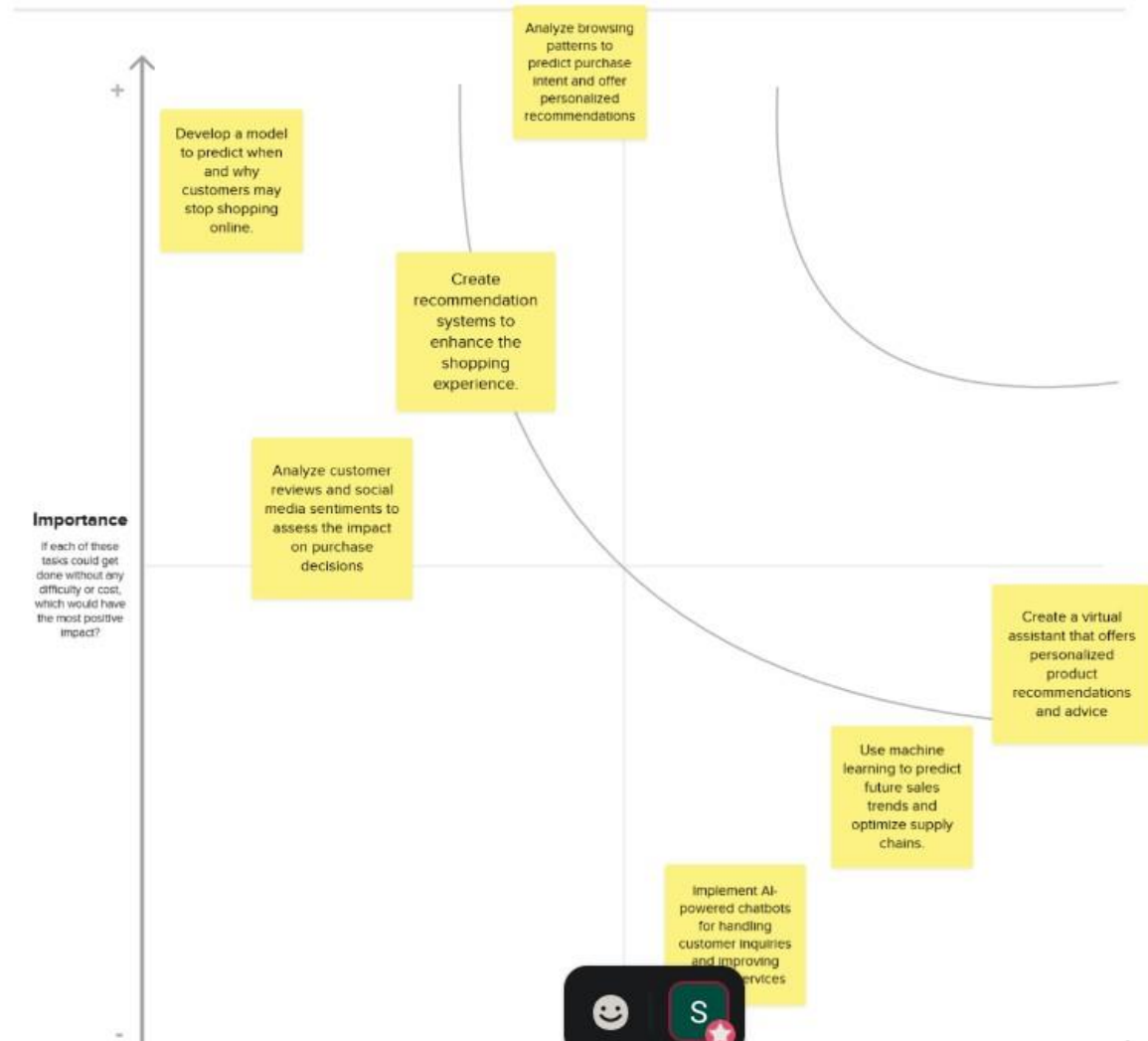
Prioritize

our team should all be on the same page about what's important moving forward. we have tried place our ideas on this grid to determine which ideas are important and which are feasible.

🕒 20 minutes

TIP

Participants can use their cursors to point at where sticky notes should go on the grid. The facilitator can confirm the spot by using the laser pointer holding the **H** key on the keyboard.



4. REQUIREMENT ANALYSIS

In this section, we will delve into the critical aspects of requirement analysis for the project "Online Shopper Intention Using AI." This analysis involves the identification of both functional and non-functional requirements to guide the project's development and ensure its success.

4.1 Functional Requirements

Functional requirements encompass the specific features and functionalities that the project must deliver. These are essential for achieving the project's objectives and meeting the needs of its stakeholders. For "Online Shopper Intention Using AI," the following functional

requirements are identified:

4.1.1 User Profiling: The system must be capable of creating and maintaining user profiles, including shopping preferences, purchase history, and personal information.

4.1.2 AI-Powered Recommendations: Implement an AI-based recommendation system that offers personalized product recommendations to online shoppers based on their browsing and purchasing behavior.

4.1.3 Predictive Analytics: Develop predictive models that analyze online shopper data to forecast their intentions and behaviors, enabling businesses to adapt their strategies accordingly.

4.1.4 Real-time Chat Support: Integrate a real-time chatbot feature to provide immediate assistance and support to online shoppers during their shopping journey.

4.1.5 Privacy and Security Measures: Implement robust security measures to protect user data and ensure compliance with data privacy regulations, such as GDPR.

4.1.6 Analytics Dashboard: Create an analytics dashboard for businesses, allowing them to monitor and evaluate shopper behavior, the success of recommendations, and overall e-commerce performance.

4.1.7 Feedback Mechanism: Establish a mechanism for shoppers to provide feedback on the shopping experience, AI-driven recommendations, and the system's functionality.

4.2 Non-Functional Requirements

Non-functional requirements pertain to the qualities and characteristics that the project must exhibit to ensure its reliability, performance, and user satisfaction. For "Online Shopper Intention Using AI," the following non-functional requirements are outlined:

4.2.1 Performance: The system must provide low-latency responses, ensuring smooth and responsive interactions for online shoppers.

4.2.2 Scalability: The platform should be capable of handling an increasing number of users and data without compromising performance.

4.2.3 Reliability: The system must be reliable, minimizing downtime and ensuring uninterrupted service availability.

4.2.4 Usability: The user interface should be intuitive, user-friendly, and accessible to a wide range of users, regardless of technical proficiency.

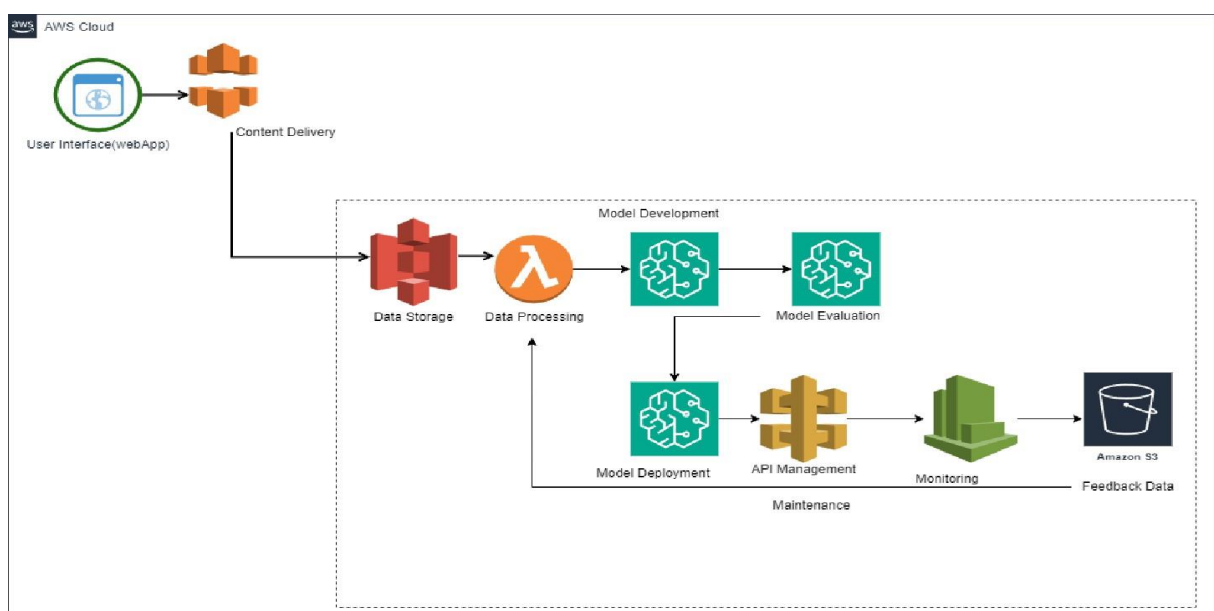
4.2.5 Data Security: Robust security measures, encryption, and access controls must be in place to safeguard user data and protect against potential breaches.

4.2.6 Compliance: Ensure that the system complies with data privacy regulations, industry standards, and best practices.

4.2.7 Maintainability: Implement a system that is easy to maintain, update, and evolve as new AI technologies and e-commerce trends emerge.

By delineating these functional and non-functional requirements, we establish a foundation for the development and implementation of the "Online Shopper Intention Using AI" project. These requirements serve as guidelines to create a system that is both feature-rich and capable of delivering a seamless, secure, and efficient online shopping experience.

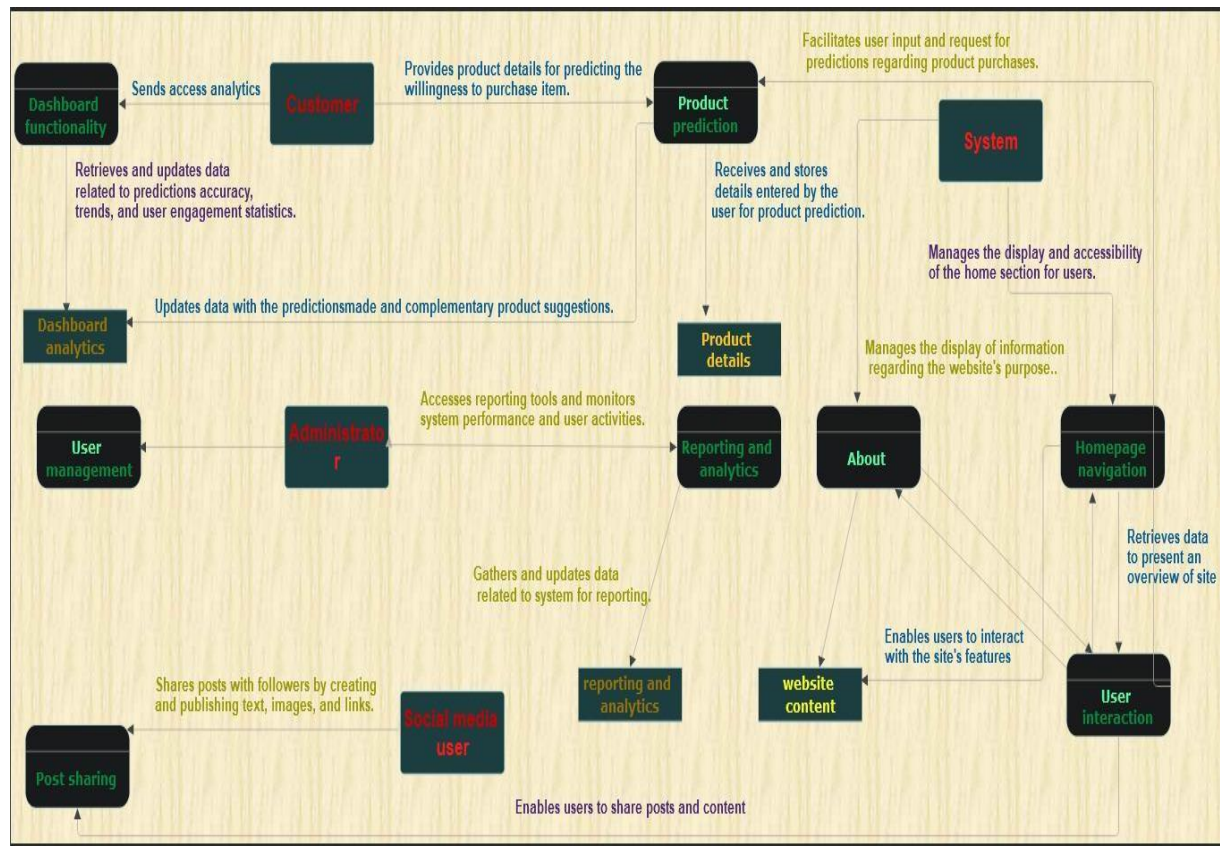
PROJECT PLANNING & SCHEDULING



5. PROJECT DESIGN

5.1 Data Flow Diagrams & User Stories

DATA FLOW DIAGRAM



User Stories:

User Stories

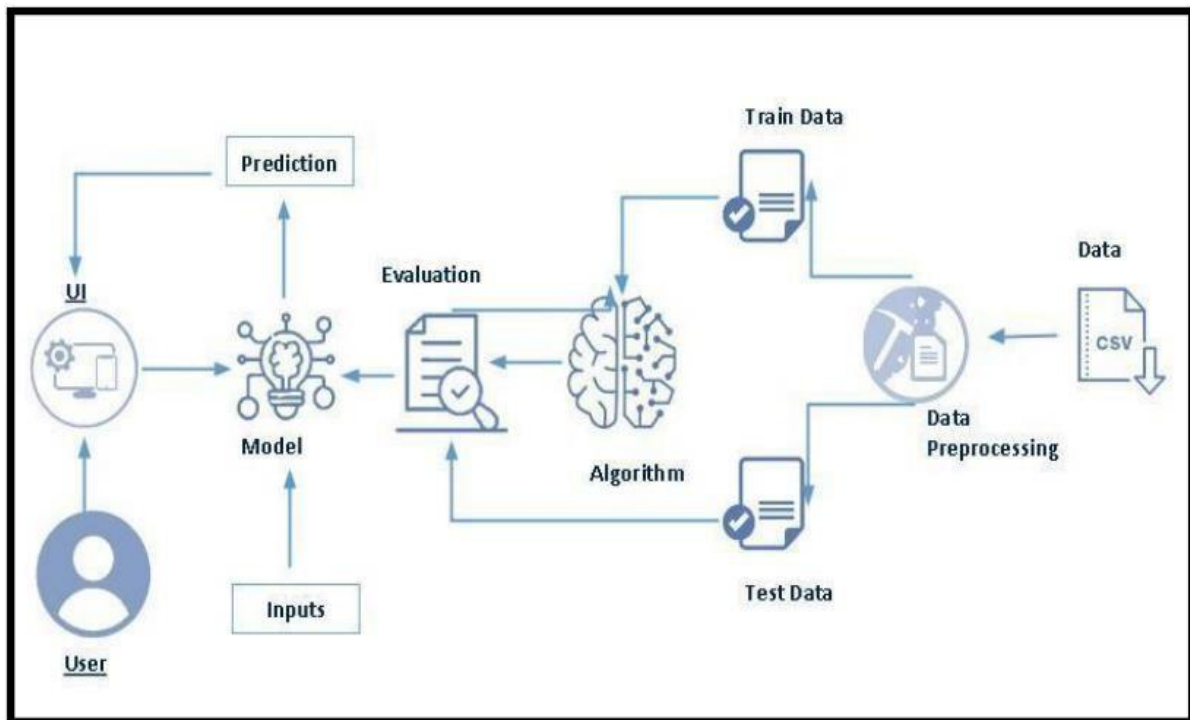
Use the below template to list all the user stories for the product.

User Type	Functional Requirement (Epic)	User Story Number	User Story / Task	Acceptance criteria	Priority	Release
Customer (web user)	Dashboard	USN-1	As a web user, I want to view my recent purchase history on the dashboard.	The purchase history should be updated in real-time as new purchases are made.	High	Sprint-1
	Dashboard analytics	USN-2	As a user, I want to view analytics on dashboard, including predictions accuracy, personal shopping trends, and user engagement statistics.	Dashboard shows top recommended products tailored to the user's browsing behavior.	High	Sprint-1
	Product Prediction	USN-3	As a user, I want to enter product details and click "Predict" to determine if potential buyer is willing to purchase the product or not.	System predicts and displays complementary products based on the selected item. - Suggestions update dynamically as the user explores different items.	High	Sprint -1
	Homepage Navigation	USN-4	As a new user, I want to navigate the home section to explore the website's offerings.	The Home section is easily accessible and offers an overview of site features	Medium	Sprint-2
	About Page Access	USN-5	As a visitor, I want to access the About section to understand the website's purpose and mission.	The About section presents information about the website's history, goals, and team.	Low	Sprint-2
	User Interaction	USN-6	As a user, I want to interact with the website's user interface (UI) to input data and request predictions.	Ability to add items to a Wishlist and create lists.	High	Sprint-1
Administrator	User management	USN-7	As an administrator, I want the ability to add and manage user accounts.	Administrators should be able to reset user	High	Sprint-1

				passwords and assign roles.		
	Reporting and analytics	USN-8	As an administrator, I want access to reporting and analytics tools to monitor system performance and user activity.	The system should generate automated reports and provide real-time data updates.	High	Sprint-2
Social Media user	Post Sharing	USN-9	As a social media user, I want to share posts with my followers.	Users should be able to create and publish posts with text, images, and links.	Medium	Sprint-2

6.PROJECT PLANNING & SCHEDULING

6.1 Technical Architecture



6.2 Sprint Planning & Estimation

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-2	Dashboard	USN-1	As a registered user, I want to have a personalized dashboard that displays my order history, recommended products, and account settings for a better shopping experience.	3	High	Sarvesh Ashish, Aarjav Jain
Sprint -3	User Preferences	USN-2	As a user, I want to set my preferences on the dashboard, such as language, region, and notification settings.	2	Medium	Sarvesh Ashish, Aarjav Jain
Sprint-4	Dashboard Analytics	USN-3	As a user, I want to view analytics on my dashboard, including predictions accuracy, personal shopping trends, and user engagement statistics.	3	High	
Sprint-2	Product Prediction	USN-4	. As a user, I want to enter product details and click "Predict" to determine if potential buyers are willing to purchase the product based on AI analysis.	5	High	Sarvesh, Aarjav Jain
Sprint-5	Product Search	USN-5	As a shopper, I want to search for products by keywords and filter search results based on various criteria to find and purchase items more efficiently.	2	High	Teertho Roy Choudhary
Sprint-1	User Interaction	USN-6	As a user, I want to interact with the website's user interface (UI) to input data and request predictions.	4	High	V chandrakanth
Sprint-2	Input Validation	USN-7	As a user, I want the input data to be validated to ensure it meets required format and quality standards before model analysis.	3	medium	Aarjav jain

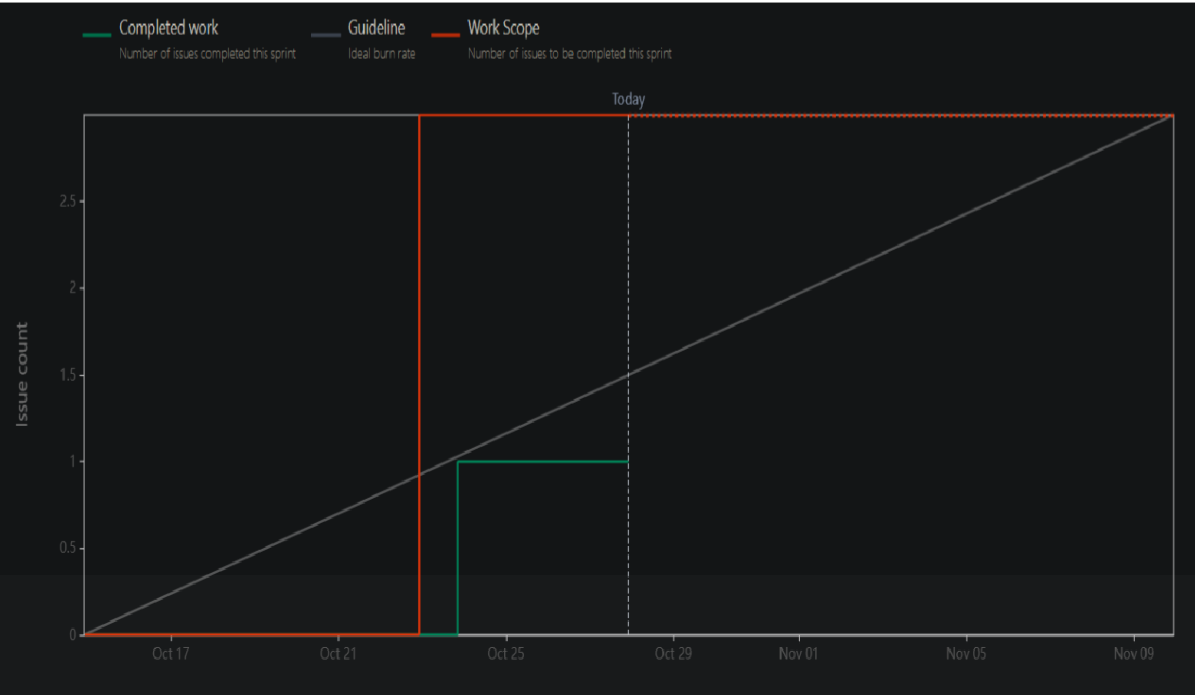
6.3 Sprint Delivery Schedule

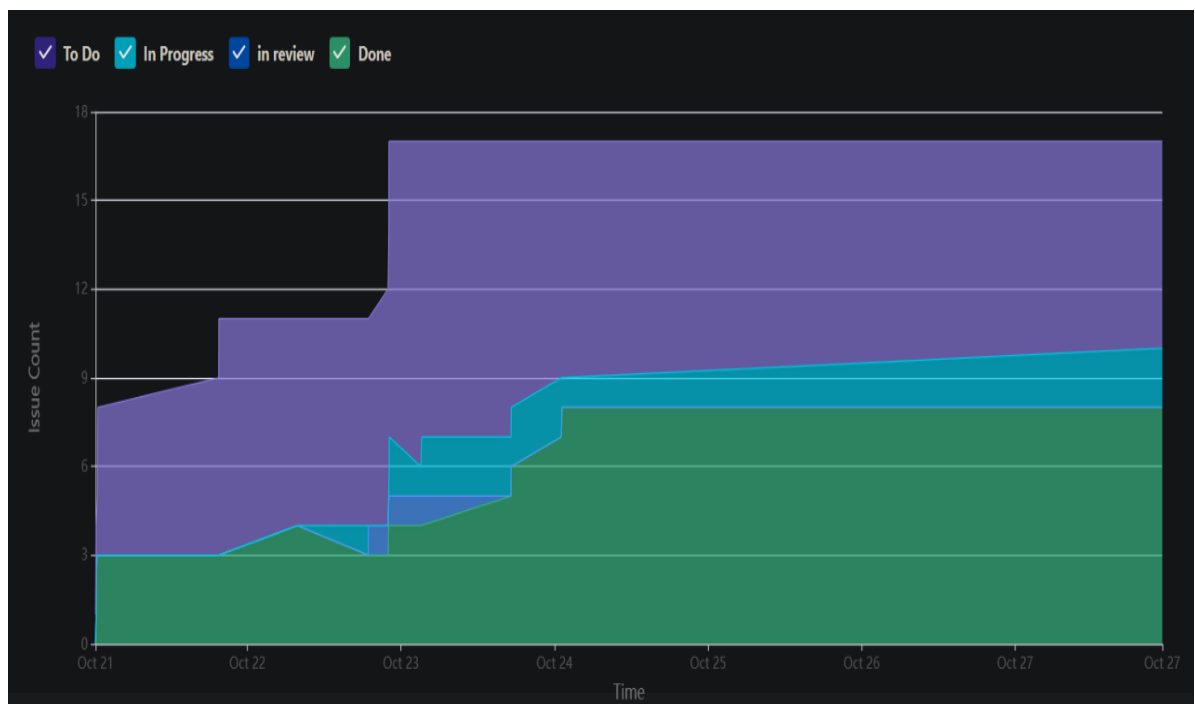
Sprint	Total Story Points	Duration	Sprint Start Date	Sprint End Date (Planned)	Story Points Completed (as on Planned End Date)	Sprint Release Date (Actual)
Sprint-1	7	4 Days	18 Oct 2023	20 Oct 2023	7	20 Oct 2023
Sprint-2	10	4 Days	21 Oct 2023	23 Oct 2023	10	23 Oct 2023
Sprint-3	5	4 Days	24 Oct 2023	27 Oct 2023	5	26 Oct 2023
Sprint-4	5	4 Days	27 Oct 2023	29 Oct 2023	5	29 Oct 2023
Sprint-5	2	4 Days	30 Oct 2022	01 Nov 2023	2	01 Nov 2023
Sprint-6	4	4 Days	02 Nov 2023	04 Nov 2023	4	04 Nov 2023

Burndown Chart:

A burn down chart is a graphical representation of work left to do versus time. It is often used in agile software development methodologies such as Scrum. However, burn down charts can be applied to any project containing measurable progress over time.

Sprint	Total Story Points	Story Points Completed (as on Planned End Date)	Remaining Story Points
Sprint-1	7	7	0
Sprint-2	10	10	0
Sprint-3	5	5	0
Sprint-4	5	5	0
Sprint-5	2	2	0
Sprint-6	4	4	0





7. CODING & SOLUTIONING

Feature 1 - Homepage (For User Login and Sign-up)

```
!DOCTYPE html>
<html lang="en">

<head>
  <meta charset="utf-8">
  <meta content="width=device-width, initial-scale=1.0" name="viewport">

  <title>Online Shoppers Intention</title>
  <meta content="" name="description">
  <meta content="" name="keywords">

  <!-- Favicons -->
  <link href="static/assets/img/favicon.ico" rel="icon">
  <link href="static/assets/img/apple-touch-icon.png" rel="apple-touch-icon">

  <!-- Google Fonts -->
  <link
href="https://fonts.googleapis.com/css?family=Open+Sans:300,300i,400,400i,600,600i,700,7
00i|Raleway:300,300i,400,400i,500,500i,600,600i,700,700i|Poppins:300,300i,400,400i,500,5
00i,600,600i,700,700i"
    rel="stylesheet">

  <!-- Vendor CSS Files -->
  <link href="static/assets/vendor/aos/aos.css" rel="stylesheet">
  <link href="static/assets/vendor/bootstrap/css/bootstrap.min.css" rel="stylesheet">
  <link href="static/assets/vendor/bootstrap-icons/bootstrap-icons.css"
rel="stylesheet">
  <link href="static/assets/vendor/boxicons/css/boxicons.min.css" rel="stylesheet">
  <link href="static/assets/vendor/glightbox/css/glightbox.min.css" rel="stylesheet">
  <link href="static/assets/vendor/remixicon/remixicon.css" rel="stylesheet">
  <link href="static/assets/vendor/swiper/swiper-bundle.min.css" rel="stylesheet">

  <!-- Main CSS File -->
  <link href="static/assets/css/style.css" rel="stylesheet">
</head>

<body>

  <!-- ===== Header ===== -->
  <header id="header" class="fixed-top ">
    <div class="container d-flex align-items-center justify-content-lg-between">

      <h1 class="logo me-auto me-lg-0"><a href="#"
onclick="redirectToPage('')">OSIs<span>.</span></a></h1>

      <nav id="navbar" class="navbar order-last order-lg-0">
```

```

        <ul>
            <li><a class="nav-link scrollto active" href="#hero">Home</a></li>
            <li><a class="nav-link scrollto" href="#about">About</a></li>
            <li><a class="nav-link scrollto" href="#"
onclick="redirectToPage('inner')">Predict</a></li>
            <li><a class="nav-link scrollto" href="#team">Team</a></li>
        </ul>
        <i class="bi bi-list mobile-nav-toggle"></i>
    </nav><!-- .navbar -->

    <a href="#about" class="get-started-btn scrollto">Get Started</a>

</div>
</header><!-- End Header -->

<!-- ===== Hero Section ===== -->
<section id="hero" class="d-flex align-items-center justify-content-center">
    <div class="container" data-aos="fade-up">

        <div class="row justify-content-center" data-aos="fade-up" data-aos-delay="150">
            <div class="col-xl-6 col-lg-8">
                <h1>Online Shoppers Intentions<span>.</span></h1>
                <h2>A platform for understanding customer desires.</h2>
            </div>
        </div>
    </div>
</section><!-- End Hero -->

<main id="main">

    <!-- ===== About Section ===== -->
    <section id="about" class="about">
        <div class="container" data-aos="fade-up">

            <div class="row">
                <div class=" pt-4 pt-lg-0 order-2 order-lg-1 content" data-aos="fade-right"
data-aos-delay="100">
                    <h3>Project Description</h3>
                    <p class="fst-italic">
                        In the "Online Shopper's Intention" project, user data provides key
                        insights into website interactions, helping businesses improve user experience and
                        conversion rates.
                    </p>
                    <ul>
                        <li><i class="ri-check-double-line"></i> <B>Id</B>: unique ID of the
                        website visitor</li>
                        <li><i class="ri-check-double-line"></i> <B>Administrative</B>: Number of
                        times the user visited the administrative section</li>

```

```

        <li><i class="ri-check-double-line"></i> <B>Administrative Duration</B>:
Total time the user spent in the administrative section</li>
        <li><i class="ri-check-double-line"></i> <B>Informational</B>: Number of
times the user visited the informational section</li>
        <li><i class="ri-check-double-line"></i> <B>Informational Duration</B>:
Total time the user spent in the informational section</li>
        <li><i class="ri-check-double-line"></i> <B>Productrelated</B>: Number of
times the user visited the related products section</li>
        <li><i class="ri-check-double-line"></i> <B>Productrelated Duration</B>:
Total time the user spent in the related products section</li>
        <li><i class="ri-check-double-line"></i> <B>Bouncerrates</B>: This is the
percentage of visitors who enter the page and immediately "bounce" without interacting
with it. This metric is only taken into account if it is the first page they visit
within the website.</li>
        <li><i class="ri-check-double-line"></i> <B>Exitrates</B>: From the total
number of visits to the pages of the website, the percentage of visitors who logged out
through this specific page is obtained, that is, it indicates the percentage of users
whose last visit to the website was this specific page.</li>
        <li><i class="ri-check-double-line"></i> <B>Pagevalues</B>: This is the
average value of the website, it indicates the contribution that this website made to
the visitor arriving at the final purchase page or section.</li>
        <li><i class="ri-check-double-line"></i> <B>Special Day</B>: Is the value
that indicates the proximity to a special date such as Valentine's Day. The range of
this variable is 0 to 1, with 1 being the exact day of the special date and 0 if there
is no range near that date.</li>
        <li><i class="ri-check-double-line"></i> <B>Month</B>: Month of the visit
to the website.</li>
        <li><i class="ri-check-double-line"></i> <B>Operating Systems</B>: Type of
operating system</li>
        <li><i class="ri-check-double-line"></i> <B>Browser</B>: Name of the web
browser</li>
        <li><i class="ri-check-double-line"></i> <B>Region</B>: Visitor's
geographic region</li>
        <li><i class="ri-check-double-line"></i> <B>Traffic Type</B>: Type of web
traffic</li>
        <li><i class="ri-check-double-line"></i> <B>Visitor Type</B>: Whether you
are a new visitor or a returning visitor</li>
        <li><i class="ri-check-double-line"></i> <B>Weekend</B>: 0 indicates that
it is not a weekend day and 1 indicates that it is a weekend day.</li>
    </ul>
    <p>
        Explore this data to optimize online shopping platforms.
    </p>
</div>
</div>
</section><!-- End About Section -->

<!-- ===== Clients Section ===== -->

```

```

<section id="clients" class="clients">
  <div class="container" data-aos="zoom-in">

    <div class="clients-slider swiper">
      <div class="swiper-wrapper align-items-center">
        <div class="swiper-slide"></div>
        <div class="swiper-slide"></div>
        <div class="swiper-slide"></div>
        <div class="swiper-slide"></div>
        <div class="swiper-slide"></div>
        <div class="swiper-slide"></div>
        <div class="swiper-slide"></div>
        <div class="swiper-slide"></div>
        <div class="swiper-slide"></div>
      </div>
      <div class="swiper-pagination"></div>
    </div>

  </div>
</section><!-- End Clients Section -->

<!-- ===== Features Section ===== -->
<section id="features" class="features">
  <div class="container" data-aos="fade-up">

    <div class="row">
      <div class="image col-lg-6" style='background-image:
url("static/assets/img/features.jpg");'
        data-aos="fade-right">
      </div>
      <div class="col-lg-6 data-aos="fade-left" data-aos-delay="100">
        <div class="icon-box mt-5 mt-lg-0" data-aos="zoom-in" data-aos-delay="150">
          <i class="bx bx-receipt"></i>
          <h4>Enhanced User Experience</h4>
          <p>Optimize the website based on user behavior, leading to a smoother
shopping experience.</p>
        </div>
        <div class="icon-box mt-5" data-aos="zoom-in" data-aos-delay="150">
          <i class="bx bx-cube-alt"></i>
          <h4>Improved Conversions</h4>

```

```
        <p>Data-driven decisions can increase conversion rates and boost business revenue.</p>
```

```
    </div>
```

```
    <div class="icon-box mt-5" data-aos="zoom-in" data-aos-delay="150">
```

```
        <i class="bx bx-images"></i>
```

```
        <h4>Personalized Marketing</h4>
```

```
        <p>Special day tracking and user segmentation enable more targeted and effective marketing campaigns.</p>
```

```
    </div>
```

```
    <div class="icon-box mt-5" data-aos="zoom-in" data-aos-delay="150">
```

```
        <i class="bx bx-shield"></i>
```

```
        <h4>Timely Engagement</h4>
```

```
        <p>Weekend detection helps in scheduling promotions when users are most active, maximizing impact.</p>
```

```
    </div>
```

```
</div>
```

```
</div>
```

```
</div>
```

```
</section><!-- End Features Section -->
```

```
<!-- ===== predict Section ===== -->
```

```
<section id="predict" class="cta">
```

```
    <div class="container" data-aos="zoom-in">
```

```
        <div class="text-center">
```

```
            <h3>Discover Personalized Shopping with AI!</h2>
```

```
            <p>Experience the future of online shopping with our Machine Learning-powered platform.</p>
```

```
            <p>Start your journey now and enhance your shopping experience.</p>
```

```
            <a class="cta-btn" href="#" onclick="redirectToPage('inner')">Predict</a>
```

```
        </div>
```

```
    </div>
```

```
</section><!-- End predict Section -->
```

```
<!-- ===== Counts Section ===== -->
```

```
<section id="counts" class="counts">
```

```
    <div class="container" data-aos="fade-up">
```

```
        <div class="row no-gutters">
```

```
            <div class="image col-xl-5 d-flex align-items-stretch justify-content-center justify-content-lg-start"
```

```
                data-aos="fade-right" data-aos-delay="100"></div>
```

```
            <div class="col-xl-7 ps-4 ps-lg-5 pe-4 pe-lg-1 d-flex align-items-stretch" data-aos="fade-left"
```

```
                data-aos-delay="100">
```

```

<div class="content d-flex flex-column justify-content-center">
  <h3>Voluptatem dignissimos provident quasi</h3>
  <p>
    Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod
    tempor incididunt ut labore et
    dolore magna aliqua. Duis aute irure dolor in reprehenderit
  </p>
  <div class="row">
    <div class="col-md-6 d-md-flex align-items-md-stretch">
      <div class="count-box">
        <i class="bi bi-emoji-smile"></i>
        <span data-purecounter-start="0" data-purecounter-end="10" data-
purecounter-duration="1"
          class="purecounter"></span>
        <p><strong>Cart Abandonment:</strong> By identifying and addressing
cart abandonment issues, businesses can potentially achieve a 10% increase in conversion
rates, resulting in a significant boost in sales.</p>
      </div>
    </div>

    <div class="col-md-6 d-md-flex align-items-md-stretch">
      <div class="count-box">
        <i class="bi bi-journal-richtext"></i>
        <span data-purecounter-start="0" data-purecounter-end="15" data-
purecounter-duration="2"
          class="purecounter"></span>
        <p><strong>User Experience Enhancement:</strong> Improving the user
experience can lead to a 15% reduction in bounce rates and a 20% increase in average
session durations, contributing to better customer engagement.
      </p>
      </div>
    </div>

    <div class="col-md-6 d-md-flex align-items-md-stretch">
      <div class="count-box">
        <i class="bi bi-clock"></i>
        <span data-purecounter-start="0" data-purecounter-end="10" data-
purecounter-duration="1"
          class="purecounter"></span>
        <p><strong>Inventory Management Optimization:</strong> Optimizing
inventory management can potentially reduce excess inventory costs by 10% and avoid
stockouts, leading to increased sales and cost savings.</p>
      </div>
    </div>

    <div class="col-md-6 d-md-flex align-items-md-stretch">
      <div class="count-box">
        <i class="bi bi-award"></i>

```



```

        <span data-purecounter-start="0" data-purecounter-end="20" data-
purecounter-duration="2"
        class="purecounter"></span>
        <p><strong>Personalized Marketing:</strong> Implementing
personalized marketing strategies based on user segmentation may result in a 15%
increase in click-through rates and a 20% improvement in conversion rates, driving
higher revenue.</p>
    </div>
</div>
</div>
</div><!-- End .content-->
</div>
</div>

</div>
</section><!-- End Counts Section -->

<!-- ===== Testimonials Section ===== -->
<section id="testimonials" class="testimonials">
    <div class="container" data-aos="zoom-in">

        <div class="testimonials-slider swiper" data-aos="fade-up" data-aos-delay="100">
            <div class="swiper-wrapper">

                <div class="swiper-slide">
                    <div class="testimonial-item">
                        
                        <h3>Carly Fiorina</h3>
                        <h4>CEO of Hewlett-Packard</h4>
                        <p>
                            <i class="bx bxs-quote-alt-left quote-icon-left"></i>
                            The goal is to turn data into information, and information into
insight.
                            <i class="bx bxs-quote-alt-right quote-icon-right"></i>
                        </p>
                    </div>
                </div><!-- End testimonial item -->

                <div class="swiper-slide">
                    <div class="testimonial-item">
                        
                        <h3>Bill Gates</h3>
                        <h4>CEO of Microsoft</h4>
                        <p>
                            <i class="bx bxs-quote-alt-left quote-icon-left"></i>
                            Your most unhappy customers are your greatest source of learning.
                            <i class="bx bxs-quote-alt-right quote-icon-right"></i>
                        </p>
                    </div>
                </div>
            </div>
        </div>
    </div>

```

```

        </p>
    </div>
</div><!-- End testimonial item -->

    <div class="swiper-slide">
        <div class="testimonial-item">
            
            <h3>Joel Anderson</h3>
            <h4>CEO of Walmart</h4>
            <p>
                <i class="bx bxs-quote-alt-left quote-icon-left"></i>
                You can't just open a website and expect people to flood in. If you
really want to succeed, you have to create traffic.
                <i class="bx bxs-quote-alt-right quote-icon-right"></i>
            </p>
        </div>
    </div><!-- End testimonial item -->

    <div class="swiper-slide">
        <div class="testimonial-item">
            
            <h3>Mary Meeker</h3>
            <h4>American venture capitalist</h4>
            <p>
                <i class="bx bxs-quote-alt-left quote-icon-left"></i>
                The Internet has changed everything; it's changed the way we shop, the
way we sell, and the way we interact.
                <i class="bx bxs-quote-alt-right quote-icon-right"></i>
            </p>
        </div>
    </div><!-- End testimonial item -->

    <div class="swiper-slide">
        <div class="testimonial-item">
            
            <h3>Sam Walton</h3>
            <h4>American business magnate</h4>
            <p>
                <i class="bx bxs-quote-alt-left quote-icon-left"></i>
                The goal as a company is to have customer service that is not just the
best but legendary.
                <i class="bx bxs-quote-alt-right quote-icon-right"></i>
            </p>
        </div>
    </div><!-- End testimonial item -->
</div>

```

```

        <div class="swiper-pagination"></div>
    </div>

</div>
</section><!-- End Testimonials Section -->

<!-- ===== Team Section ===== -->
<section id="team" class="team">
    <div class="container" data-aos="fade-up">

        <div class="section-title">
            <h2>Team</h2>
            <p>Check our Team</p>
        </div>

        <div class="row">

            <div class="col-lg-3 col-md-6 d-flex align-items-stretch">
                <div class="member" data-aos="fade-up" data-aos-delay="100">
                    <div class="member-img">
                        
                    <div class="social">
                        <a href="https://github.com/Aarjav423"><i class="bi bi-
github"></i></a>
                        <a href="https://instagram.com/aarjav423?igshid=OGQ5ZDc2ODk2ZA=="><i
class="bi bi-instagram"></i></a>
                        <a href="https://www.linkedin.com/in/aarjav-jain-08b061227/"><i
class="bi bi-linkedin"></i></a>
                    </div>
                </div>
                <div class="member-info">
                    <h4>Aarjav Jain</h4>
                    <span>Team Lead</span>
                </div>
            </div>

            <div class="col-lg-3 col-md-6 d-flex align-items-stretch">
                <div class="member" data-aos="fade-up" data-aos-delay="200">
                    <div class="member-img">
                        
                    <div class="social">
                        <a href="https://github.com/chandu-cpz"><i class="bi bi-
github"></i></a>
                        <a href="https://www.instagram.com/chandu_cpz/"><i class="bi bi-
instagram"></i></a>
                        <a href="https://www.linkedin.com/in/chandu-cpz/"><i class="bi bi-
linkedin"></i></a>
                    </div>
                </div>
            </div>
        </div>
    </div>

```

```

        </div>
        <div class="member-info">
            <h4>V.Chandrakanth</h4>
            <span>Member</span>
        </div>
    </div>
</div>

<div class="col-lg-3 col-md-6 d-flex align-items-stretch">
    <div class="member" data-aos="fade-up" data-aos-delay="300">
        <div class="member-img">
            
            <div class="social">
                <a href="http://github.com/Teertho65"><i class="bi bi-github"></i></a>
                <a href="https://instagram.com/teertho__rc?igshid=N3JyYTdrMnF6cWk4&utm_source=qr"><i class="bi bi-instagram"></i></a>
                <a href="https://www.linkedin.com/in/teertho-roychoudhury-824129260?utm_source=share&utm_campaign=share_via&utm_content=profile&utm_medium=ios_app"><i class="bi bi-linkedin"></i></a>
            </div>
        </div>
        <div class="member-info">
            <h4>Teertho</h4>
            <span>Member</span>
        </div>
    </div>
</div>

<div class="col-lg-3 col-md-6 d-flex align-items-stretch">
    <div class="member" data-aos="fade-up" data-aos-delay="400">
        <div class="member-img">
            
            <div class="social">
                <a href="https://github.com/Sarvesh-pr"><i class="bi bi-github"></i></a>
                <a href="https://www.instagram.com/sarvesh_ashish_51/"><i class="bi bi-instagram"></i></a>
                <a href="https://www.linkedin.com/in/sarvesh-ashish-7aa53a222/"><i class="bi bi-linkedin"></i></a>
            </div>
        </div>
        <div class="member-info">
            <h4>Sarvesh Ashish</h4>
            <span>Member</span>
        </div>
    </div>
</div>

```

```

        </div>

    </div>
</section><!-- End Team Section -->

<!-- ===== Contact Section ===== -->

</main><!-- End #main -->

<!-- ===== Footer ===== -->
<footer id="footer">
</footer>
<!-- End Footer -->

<div id="preloader"></div>
<a href="#" class="back-to-top d-flex align-items-center justify-content-center"><i
    class="bi bi-arrow-up-short"></i></a>

<!-- Vendor JS Files -->
<script src="static/assets/vendor/purecounter/purecounter_vanilla.js"></script>
<script src="static/assets/vendor/aos/aos.js"></script>
<script src="static/assets/vendor/bootstrap/js/bootstrap.bundle.min.js"></script>
<script src="static/assets/vendor/glightbox/js/glightbox.min.js"></script>
<script src="static/assets/vendor/isotope-layout/isotope.pkgd.min.js"></script>
<script src="static/assets/vendor/swiper/swiper-bundle.min.js"></script>
<script src="static/assets/vendor/php-email-form/validate.js"></script>

<!-- Template Main JS File -->
<script src="static/assets/js/main.js"></script>

<script>
    function redirectToPage(page) {
        window.location.href = "/" + page;
    }
</script>

</body>

</html>

```

7.2 Feature 2 – Predict page

```

8. <!DOCTYPE html>
9. <html lang="en">
10.
11. <head>
12.     <meta charset="utf-8">

```

```
13. <meta content="width=device-width, initial-scale=1.0" name="viewport">
14.
15. <!-- <title>Inner Page - Gp Bootstrap Template</title> -->
16. <meta content="" name="description">
17. <meta content="" name="keywords">
18.
19. <!-- Favicons -->
20. <link href="static/assets/img/favicon.ico" rel="icon">
21. <link href="static/assets/img/apple-touch-icon.png" rel="apple-touch-icon">
22.
23. <!-- Google Fonts -->
24. <link
25.     href="https://fonts.googleapis.com/css?family=Open+Sans:300,300i,400,400i,600,600i
        ,700,700i|Raleway:300,300i,400,400i,500,500i,600,600i,700,700i|Poppins:300,300i,40
        0,400i,500,500i,600,600i,700,700i"
26.     rel="stylesheet">
27.
28. <!-- Vendor CSS Files -->
29. <link href="static/assets/vendor/aos/aos.css" rel="stylesheet">
30. <link href="static/assets/vendor/bootstrap/css/bootstrap.min.css"
    rel="stylesheet">
31. <link href="static/assets/vendor/bootstrap-icons/bootstrap-icons.css"
    rel="stylesheet">
32. <link href="static/assets/vendor/boxicons/css/boxicons.min.css"
    rel="stylesheet">
33. <link href="static/assets/vendor/glightbox/css/glightbox.min.css"
    rel="stylesheet">
34. <link href="static/assets/vendor/remixicon/remixicon.css" rel="stylesheet">
35. <link href="static/assets/vendor/swiper/swiper-bundle.min.css" rel="stylesheet">
36.
37. <!-- Template Main CSS File -->
38. <link href="static/assets/css/style.css" rel="stylesheet">
39. </head>
40.
41. <body>
42.
43. <!-- ===== Header ===== -->
44. <header id="header" class="fixed-top header-inner-pages">
45.     <div class="container d-flex align-items-center justify-content-lg-between">
46.
47. <h1 class="logo me-auto me-lg-0"><a href="#"
    onclick="redirectToPage('')">OSIs</span></a></h1>
48.     <nav id="navbar" class="navbar order-last order-lg-0">
49.         <ul>
50. <li><a class="nav-link scrollto " href="#hero"
    onclick="redirectToPage('')">Home</a></li>
51. <li><a class="nav-link scrollto" href="#about"
    onclick="redirectToPage('')">About</a></li>
```

[illegible]

```

97. <input type="number" name="Administrative" class="form-control"
    value="Administrative"
98.         placeholder="Administrative" required>
99.     </div>
100.         <div class="col-md-4 form-group mt-3 mt-md-0">
101.             <h6>Administrative Duration</h6>
102.             <input type="number" step="0.1" class="form-control"
    name="Administrative_Duration"
103.                 value="Administrative_Duration"
    placeholder="Administrative Duration" required>
104.         </div>
105.         <div class="col-md-4 form-group mt-3 mt-md-0">
106.             <h6>Informational</h6>
107.             <input type="number" class="form-control"
    name="Informational" value="Informational"
108.                 placeholder="Informational" required>
109.         </div>
110.     </div>
111.     <div class="row mt-5">
112.         <div class="col-md-4 form-group">
113.             <h6>Informational Duration</h6>
114.             <input type="number" step="0.1"
    name="Informational_Duration" class="form-control"
115.                 value="Informational_Duration"
    placeholder="Informational Duration" required>
116.         </div>
117.         <div class="col-md-4 form-group mt-3 mt-md-0">
118.             <h6>Product Related</h6>
119.             <input type="number" class="form-control"
    name="ProductRelated" value="ProductRelated"
120.                 placeholder="ProductRelated" required>
121.         </div>
122.         <div class="col-md-4 form-group mt-3 mt-md-0">
123.             <h6>ProductRelated Duration</h6>
124.             <input type="number" step="0.1" class="form-control"
    name="ProductRelated_Duration"
125.                 value="ProductRelated_Duration"
    placeholder="ProductRelated Duration" required>
126.         </div>
127.     </div>
128.     <div class="row mt-5">
129.         <div class="col-md-4 form-group">
130.             <h6>Bounce Rates</h6>
131.             <input type="number" step="0.1" name="BounceRates"
    class="form-control" value="BounceRates"
132.                 placeholder="BounceRates" required>
133.         </div>
134.         <div class="col-md-4 form-group mt-3 mt-md-0">
135.             <h6>Exit Rates</h6>

```



```
136.         <input type="number" step="0.1" class="form-control"
    name="ExitRates" value="ExitRates"
137.             placeholder="ExitRates" required>
138.     </div>
139.     <div class="col-md-4 form-group mt-3 mt-md-0">
140.         <h6>Page Values</h6>
141.         <input type="number" step="0.1" class="form-control"
    name="PageValues" value="PageValues"
142.             placeholder="PageValues" required>
143.     </div>
144. </div>
145. <div class="row mt-5">
146.     <div class="col-md-4 form-group">
147.         <h6>Special Day</h6>
148.         <input type="number" step="0.1" name="SpecialDay"
    class="form-control" value="SpecialDay"
149.             placeholder="Special Day" required>
150.     </div>
151.     <div class="col-md-4 form-group mt-3 mt-md-0">
152.         <h6>Month</h6>
153.         <div class="select-wrapper">
154.             <select style="height: 44px;" class="form-control"
    name="Month" id="Month" placeholder="month"
155.                 required>
156.                 <option value="1">January</option>
157.                 <option value="2">February</option>
158.                 <option value="3">March</option>
159.                 <option value="4">April</option>
160.                 <option value="5">May</option>
161.                 <option value="6">June</option>
162.                 <option value="7">July</option>
163.                 <option value="8">August</option>
164.                 <option value="9">September</option>
165.                 <option value="10">October</option>
166.                 <option value="11">November</option>
167.                 <option value="12">December</option>
168.             </select>
169.             <div class="arrow-down"></div>
170.         </div>
171.     </div>
172.     <div class="col-md-4 form-group mt-3 mt-md-0">
173.         <h6>Operating Systems</h6>
174.         <input type="number" class="form-control"
    name="OperatingSystems" value="OperatingSystems"
175.             placeholder="Operating Systems" required>
176.     </div>
177. </div>
178. <div class="row mt-5">
179.     <div class="col-md-4 form-group">
```

```
180.         <h6>Browser</h6>
181.         <input type="number" name="Browser" class="form-
182.             control" value="Browser" placeholder="Browser"
183.             required>
184.     </div>
185.     <div class="col-md-4 form-group mt-3 mt-md-0">
186.         <h6>Region</h6>
187.         <input type="number" class="form-control"
188.             name="Region" value="Region" placeholder="Region"
189.             required>
190.     </div>
191.     <div class="col-md-4 form-group mt-3 mt-md-0">
192.         <h6>TrafficType</h6>
193.         <input type="number" class="form-control"
194.             name="TrafficType" value="TrafficType"
195.             placeholder="TrafficType" required>
196.     </div>
197. </div>
198. <div class="row mt-5">
199.     <div class="col-md-4 form-group">
200.         <h6>VisitorType</h6>
201.         <div class="select-wrapper">
202.             <select style="height: 44px;" class="form-control"
203.                 name="VisitorType" placeholder="Weekend"
204.                 required>
205.                 <option value="0">Returning_Visitor</option>
206.                 <option value="1">New_Visitor</option>
207.                 <option value="2">Other</option>
208.             </select>
209.             <div class="arrow-down"></div>
210.         </div>
211.     </div>
212.     <div class="col-md-4 form-group mt-3 mt-md-0">
213.         <h6>Weekend</h6>
214.         <div class="select-wrapper">
215.             <select style="height: 44px;" class="form-control"
216.                 name="Weekend" placeholder="Weekend"
217.                 required>
218.                 <option value="1">True</option>
219.                 <option value="0">False</option>
220.             </select>
221.             <div class="arrow-down"></div>
222.         </div>
223.     </div>
224. </div>
225. <div class="my-3">
226.     <div class="loading">Loading</div>
227.     <div class="error-message"></div>
```

```

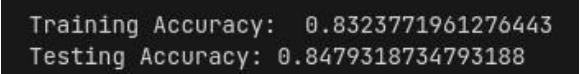
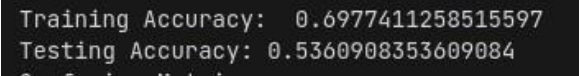
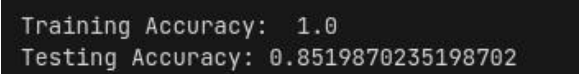
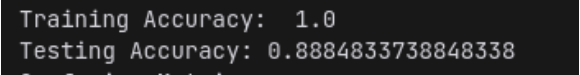
223.             <div class="sent-message">Your response has been sent.
                Thank you!</div>
224.             </div>
225.             <div class="text-center"><button type="submit"
                value="submit">SUBMIT</button></div>
226.
227.             <a id="predict"></a>
228.             <div class="output">
229.                 <h2><b>{{prediction_text}}</b></h2>
230.             </div>
231.         </form>
232.     </div>
233. </div>
234. </div>
235. </section>
236. </div>
237. </section><!-- End Prediction Section -->
238.
239. <!-- ===== Footer ===== -->
240.     <footer id="footer">
241. </footer>
242. <!-- End Footer -->
243.
244.     <div id="preloader"></div>
245.     <a href="#" class="back-to-top d-flex align-items-center justify-
        content-center"><i
246.         class="bi bi-arrow-up-short"></i></a>
247.
248.     <!-- Vendor JS Files -->
249.     <script
        src="static/assets/vendor/purecounter/purecounter_vanilla.js"></script>
250.     <script src="static/assets/vendor/aos/aos.js"></script>
251.     <script
        src="static/assets/vendor/bootstrap/js/bootstrap.bundle.min.js"></script>
252.     <script
        src="static/assets/vendor/glightbox/js/glightbox.min.js"></script>
253.     <script src="static/assets/vendor/isotope-
        layout/isotope.pkgd.min.js"></script>
254.     <script src="static/assets/vendor/swiper/swiper-bundle.min.js"></script>
255.     <script src="static/assets/vendor/predict-form/validate.js"></script>
256.
257.     <!-- Template Main JS File -->
258.     <script src="static/assets/js/main.js"></script>
259.     <script>
260.         function redirectToPage(page) {
261.             window.location.href = "/" + page;
262.         }
263.     </script>
264. </body>

```

265.
266. </html>

1. PERFORMANCE TESTING

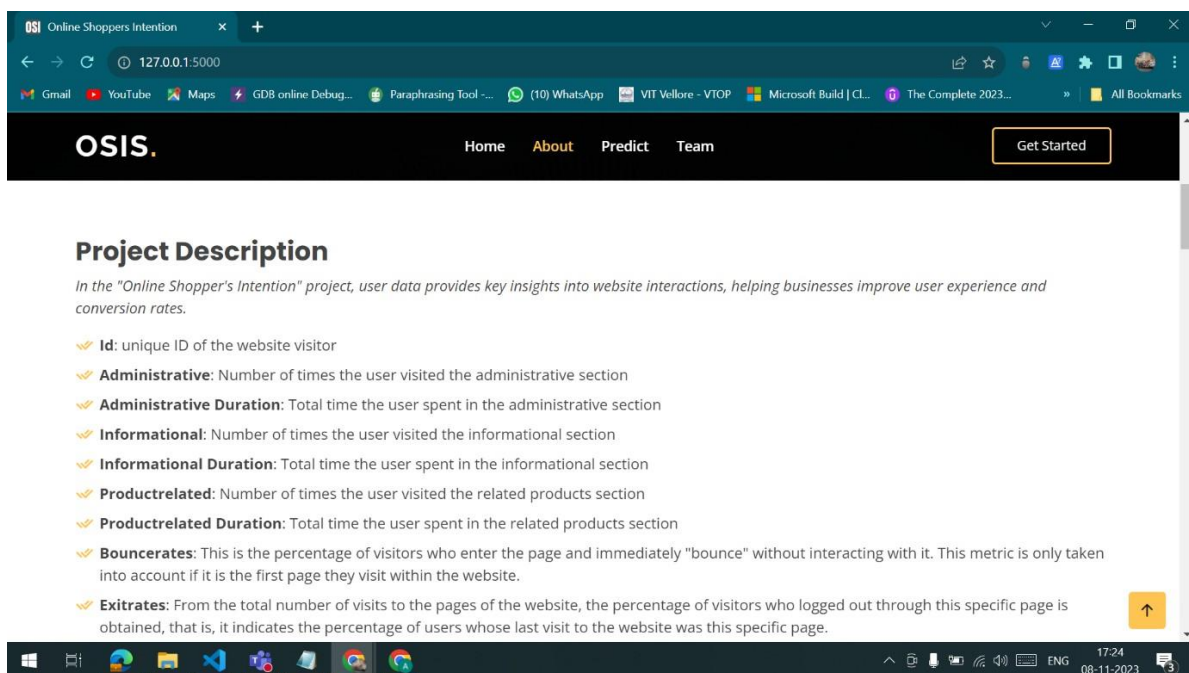
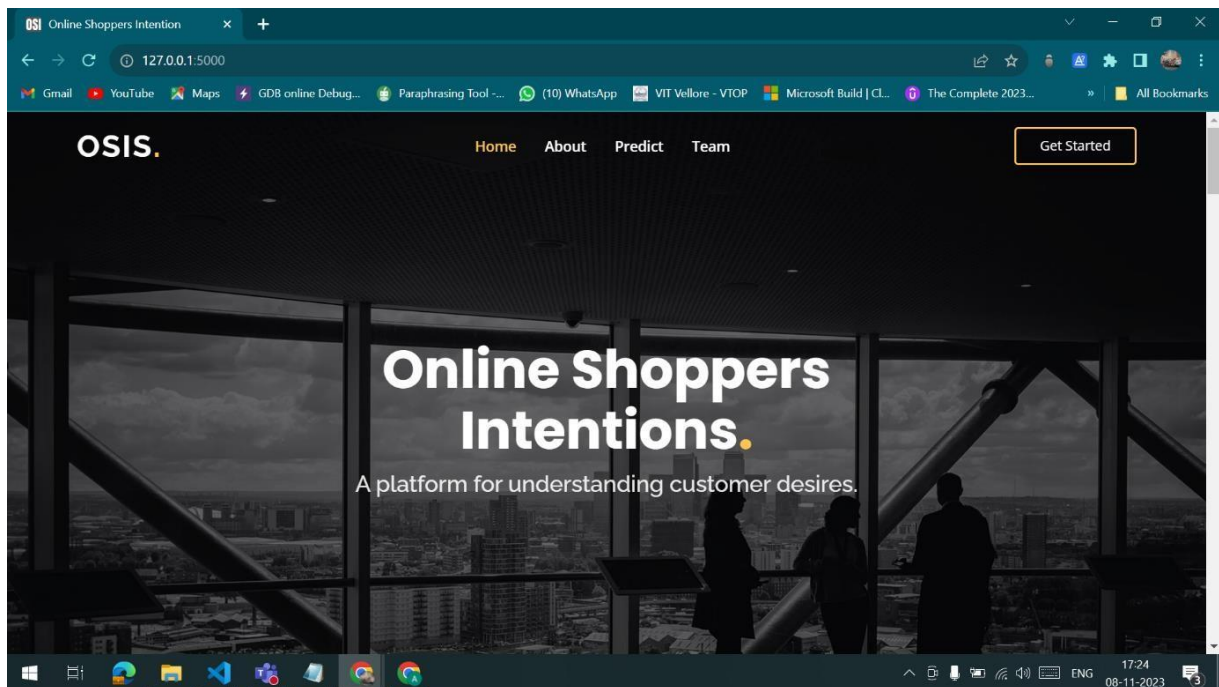
8.1 Performance Metrics

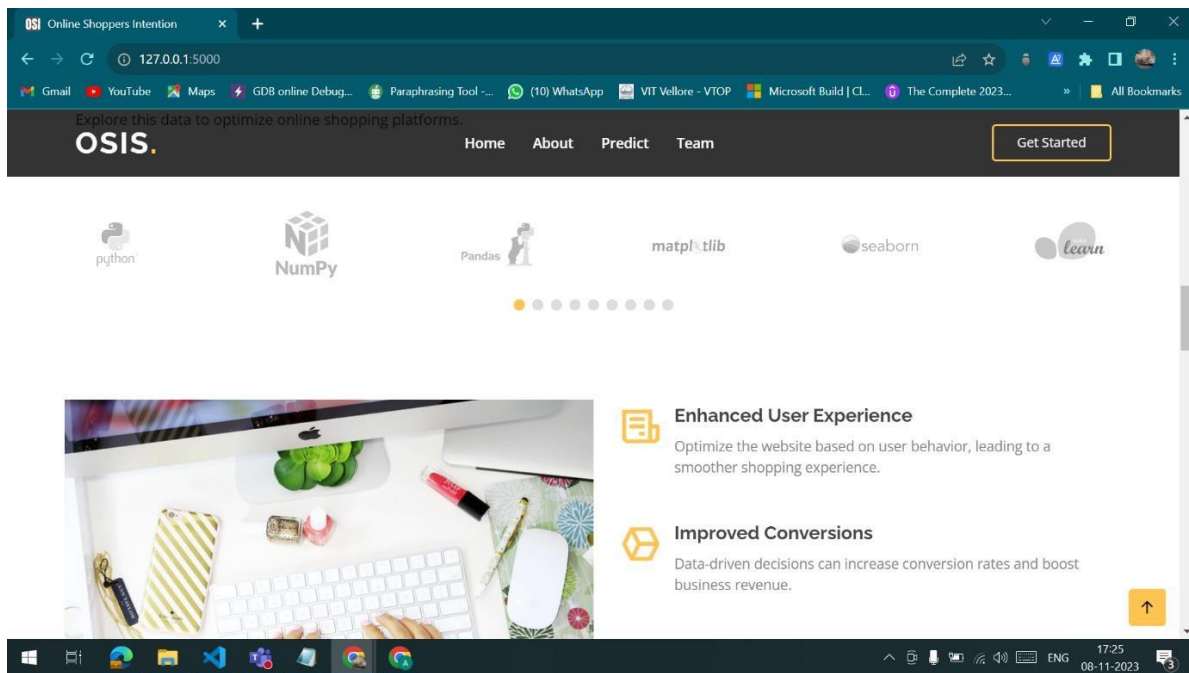
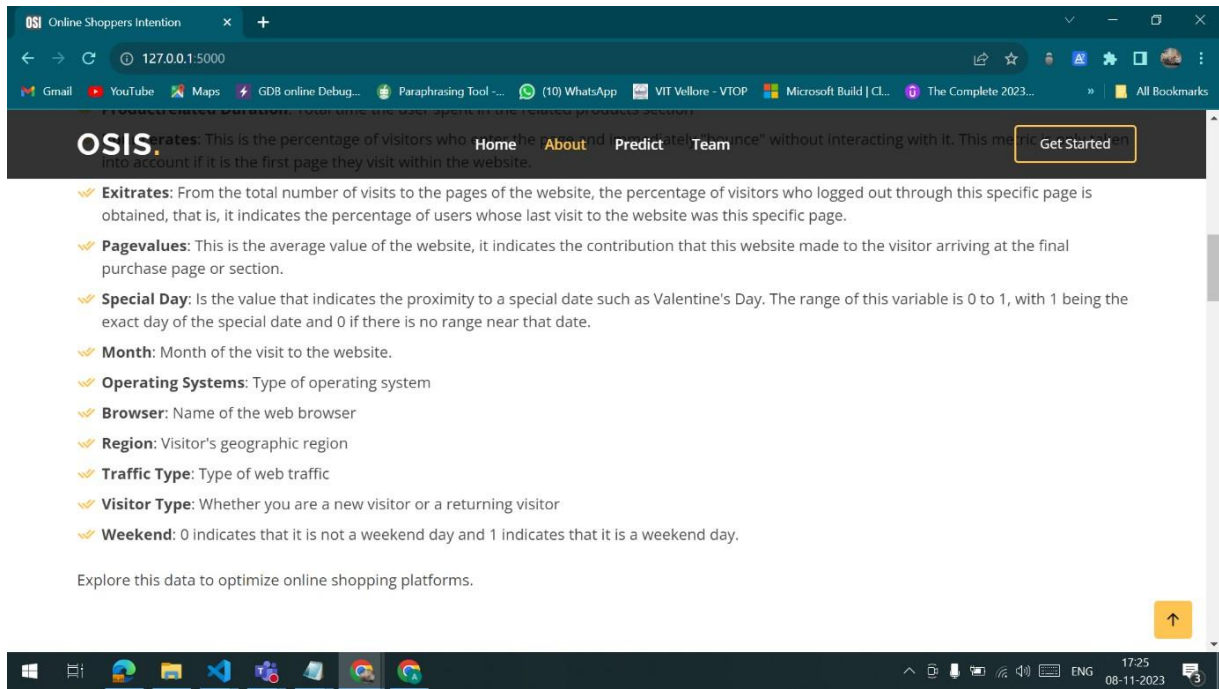
S.No	Parameter	Values	Screenshot
1	Accuracy score, classification report, confusion matrix	Logistic Regression Model: Training Accuracy: 0.8323771961276443 Testing Accuracy: 0.8479318734793188	
		GuassianNB Model: Training Accuracy: 0.6977411258515597 Testing Accuracy: 0.5360908353609084	
		Decision Tree Classifier Model: Training Accuracy: 1.0 Testing Accuracy: 0.8519870235198702	
		Random Forest Classifier: Training Accuracy: 1.0 Testing Accuracy: 0.8884833738848338 Confusion Matrix: [[1899 156] [119 292]] Classification Report: precision recall f1-score support 0.0 0.94 0.92 0.93 2055 1.0 0.65 0.71 0.68 411 accuracy 0.89 2466	

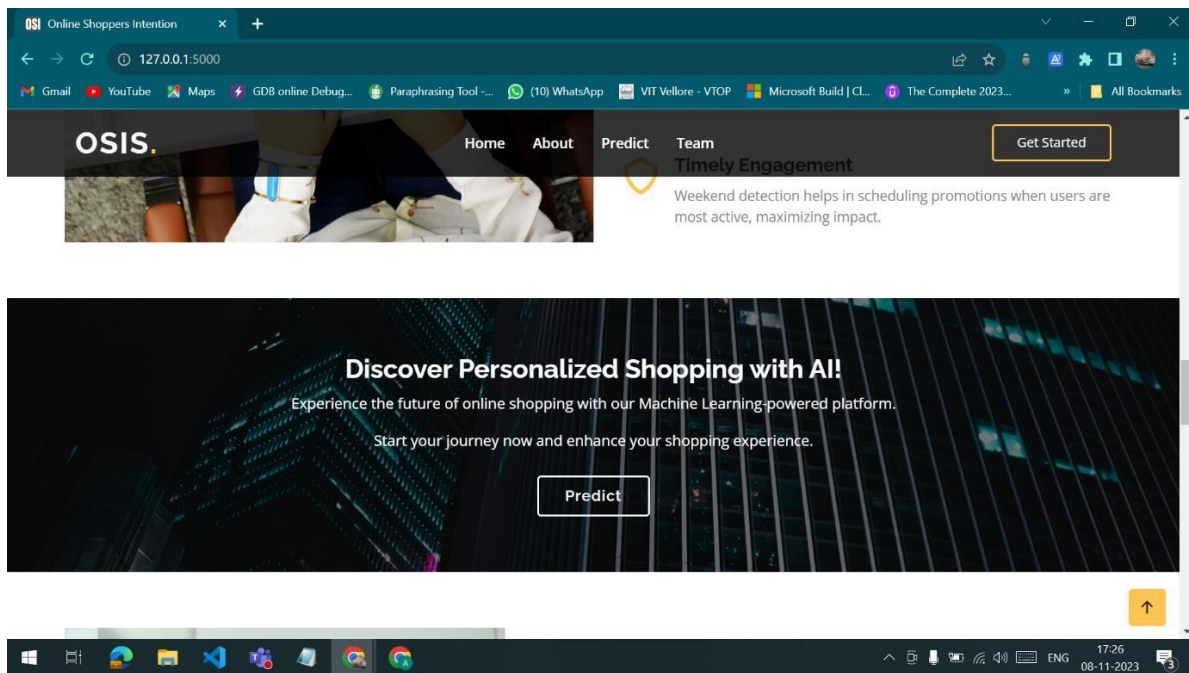
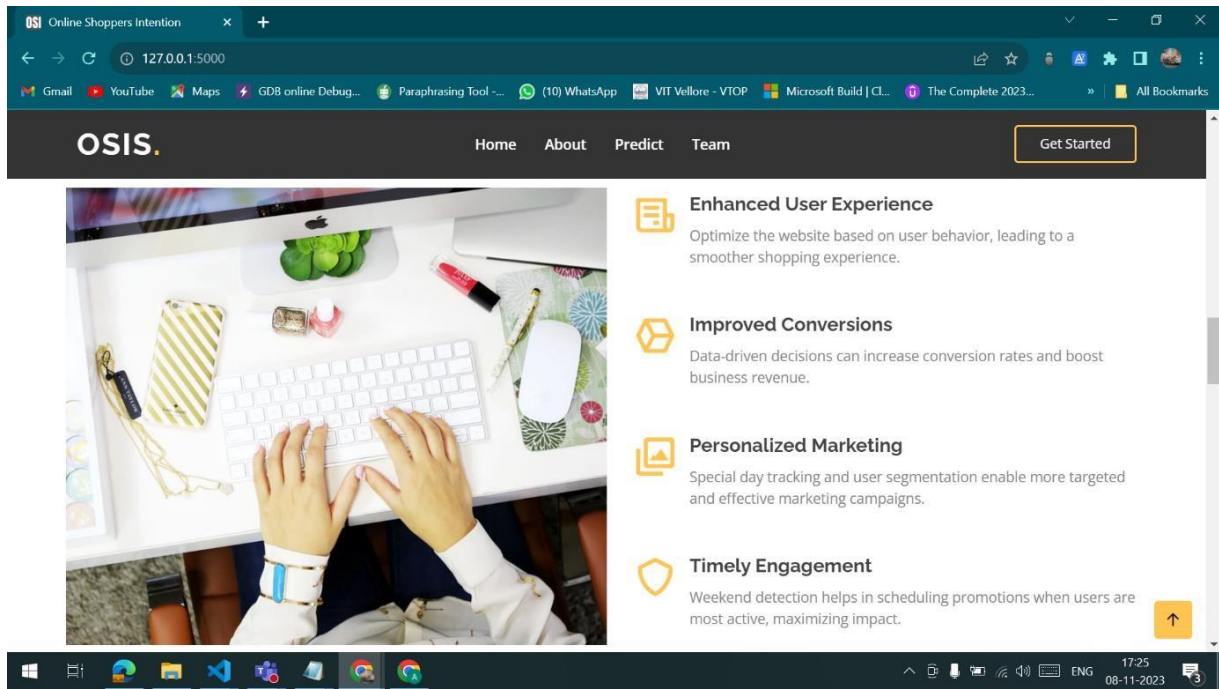
		<p>macro</p> <p>avg 0.80 0.82 0.81 2466</p> <p>weighted</p> <p>avg 0.89 0.89 0.89 2466</p>	
		<p>Support Vector Classifier</p> <p>Model: Training Accuracy: 0.8572367634755588 Testing Accuracy: 0.8690186536901865</p>	<p>Training Accuracy: 0.8572367634755588</p> <p>Testing Accuracy: 0.8690186536901865</p>
		<p>KNN Classifier: Training Accuracy: 0.904027727978965</p> <p>Testing Accuracy: 0.7181670721816708</p>	<p>Training Accuracy: 0.904027727978965</p> <p>Testing Accuracy: 0.7181670721816708</p>

		<p>Random Forest Classifier - HyperParameter Tuned:</p> <pre>{'max_depth': 8, 'min_samples_leaf': 5, 'min_samples_split': 2, 'n_estimators': 300} Training Accuracy: 0.9117365841998327 Testing Accuracy: 0.8771289537712895 Confusion Matrix: [[1832 223] [80 331]] Classification Report: precision recall f1-score support 0.0 0.96 0.89 0.92 2055 1.0 0.60 0.81 0.69 411 accuracy 0.88 2466 macro avg 0.78 0.85 0.80 2466 weighted avg 0.90 0.88 0.88 2466</pre>	<pre>{'max_depth': 8, 'min_samples_leaf': 5, 'min_samples_split': 2, 'n_estimators': 300} Training Accuracy: 0.9117365841998327 Testing Accuracy: 0.8771289537712895 Confusion Matrix: [[1832 223] [80 331]] Classification Report: precision recall f1-score support 0.0 0.96 0.89 0.92 2055 1.0 0.60 0.81 0.69 411 accuracy 0.88 2466 macro avg 0.78 0.85 0.80 2466 weighted avg 0.90 0.88 0.88 2466</pre>
2	Classification Report	<p>Validation – The no of wrong classification's compared to correct predictions make it a good model to predict the user intention to shop.</p>	<p>Hence the model is valid.</p>

9. RESULTS







OSI Online Shoppers Intention


127.0.0.1:5000

Gmail YouTube Maps GDB online Debug... Paraphrasing Tool ... (10) WhatsApp VIT Vellore - VTOP Microsoft Build | CL... The Complete 2023... All Bookmarks

OSIS.


Home About Predict Team

Get Started




Problems Solved


Four major problems that can be addressed by the "Online Shopper's Intention" project, along with their potential impact in numerical terms:

**10**


Cart Abandonment: By identifying and addressing cart abandonment issues, businesses can potentially achieve a 10% increase in conversion rates, resulting in a significant boost in sales.

**15**

User Experience Enhancement: Improving the user experience can lead to a 15% reduction in bounce rates and a 20% increase in average session durations, contributing to better customer engagement.

**10**

Inventory Management Optimization: Optimizing inventory management can potentially reduce excess inventory costs by 10% and avoid stockouts, leading to increased sales and cost savings.

**20**

Personalized Marketing: Implementing personalized marketing strategies based on user segmentation may result in a 15% increase in click-through rates and a 20% improvement in conversion rates, driving higher revenue.

↑

Windows Taskbar

17:28 08-11-2023

OSI Online Shoppers Intention

127.0.0.1:5000

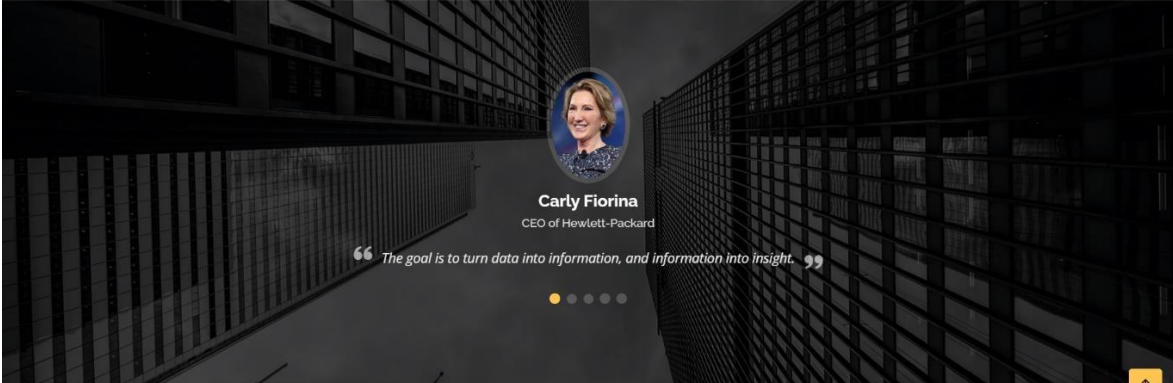
Gmail YouTube Maps GDB online Debug... Paraphrasing Tool ... (10) WhatsApp VIT Vellore - VTOP Microsoft Build | CL... The Complete 2023... All Bookmarks


OSIS.

Home About Predict Team

higher revenue

Get Started





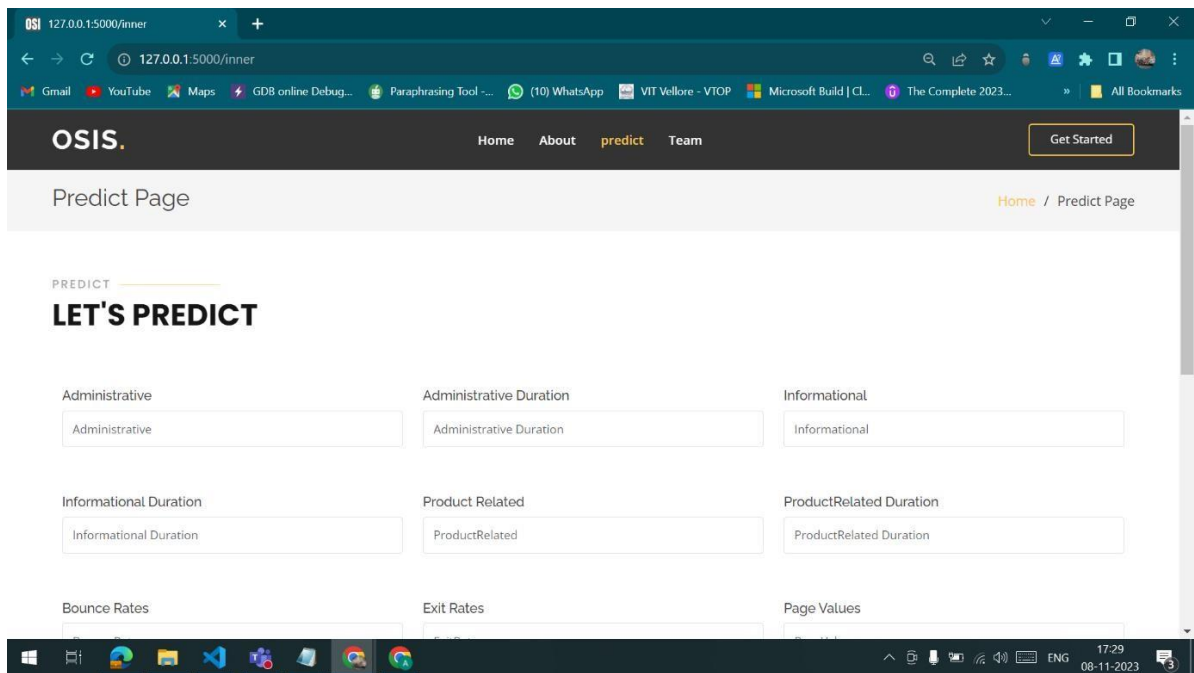
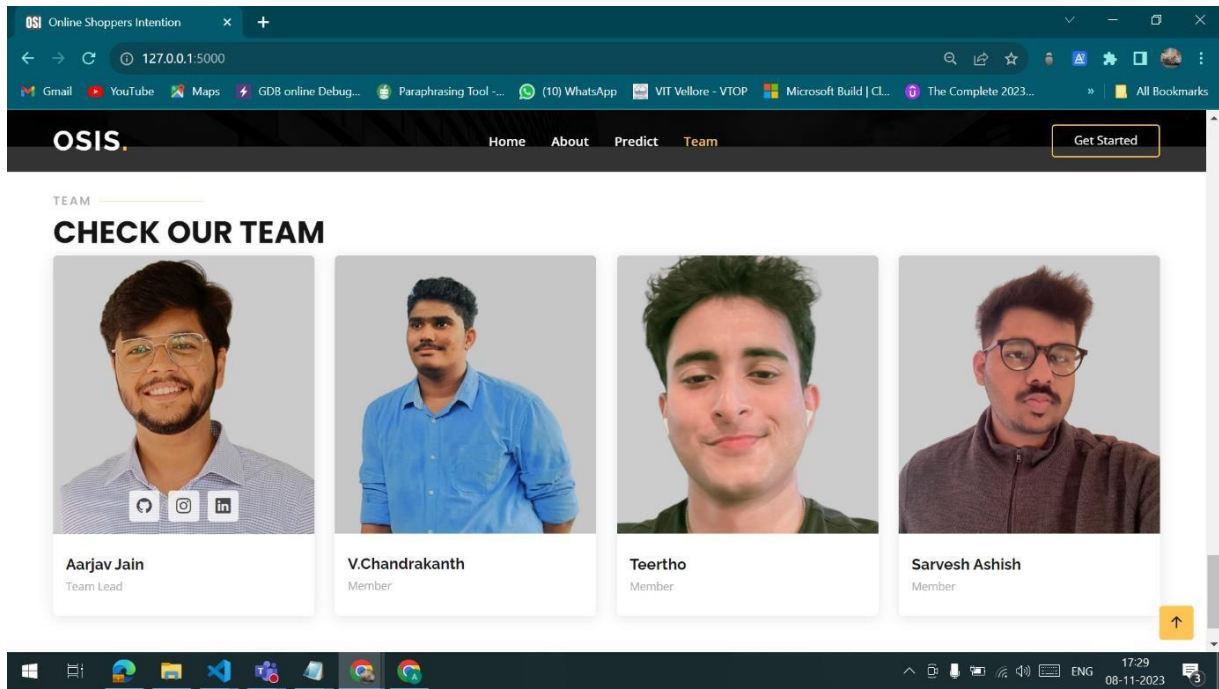
Carly Fiorina
CEO of Hewlett-Packard

“ The goal is to turn data into information, and information into insight. ”

↑

Windows Taskbar

17:29 08-11-2023



The screenshot shows a web browser window with the URL 127.0.0.1:5000/inner. The page has a dark header with the OSIS logo and navigation links: Home, About, predict (highlighted), and Team. A 'Get Started' button is in the top right. The main content area contains a form with the following fields:

- Bounce Rates:
- Exit Rates:
- Page Values:
- Special Day:
- Month: - Operating Systems:
- Browser:
- Region: - Traffic Type:
- Visitor Type: - Weekend:

A yellow SUBMIT button is located at the bottom center of the form. The browser's taskbar at the bottom shows various icons and the system clock indicating 17:30 on 08-11-2023.

10. ADVANTAGES & DISADVANTAGES

Advantages of "Online Shopper Intention Using AI" Project:

- 1) **Personalized Shopping Experience:** The implementation of an AI-powered recommendation system allows for highly personalized product suggestions based on individual user preferences and behaviors, enhancing the overall shopping experience.
- 2) **Increased Sales and Customer Loyalty:** By accurately predicting and catering to online shopper intentions, businesses can optimize their product offerings, leading to increased sales and fostering customer loyalty through tailored recommendations.
- 3) **Efficient Predictive Analytics:** The integration of predictive analytics enables businesses to forecast trends and behaviors, allowing for proactive decision-making and the adaptation of marketing and sales strategies in real-time.
- 4) **Real-time Customer Support:** The incorporation of a real-time chatbot feature provides immediate assistance to online shoppers, addressing queries and concerns promptly, thereby improving customer satisfaction and trust.
- 5) **Data-Driven Business Insights:** The analytics dashboard offers businesses valuable insights into shopper behavior, allowing for informed decision-making, targeted marketing campaigns, and continuous improvement of the online shopping platform.

Disadvantages of "Online Shopper Intention Using AI" Project:

- 1) **Privacy Concerns:** The collection and analysis of user data for personalized recommendations raise privacy concerns among some users. Striking a balance between personalization and respecting user privacy is crucial.
- 2) **Algorithmic Bias:** AI algorithms may inadvertently exhibit bias, leading to skewed recommendations or discriminatory outcomes. Continuous monitoring and adjustments are necessary to ensure fairness and prevent unintended consequences.
- 3) **Implementation Challenges:** Integrating AI technologies into existing e-commerce platforms can be complex and may require significant resources, including time, expertise, and financial investment.
- 4) **User Resistance:** Some users may resist or feel uneasy about AI-driven systems, especially when it comes to automated decision-making. Clear communication and user education are essential to mitigate such resistance.
- 5) **Dependency on Data Quality:** The effectiveness of the AI system heavily relies on the quality and accuracy of the data used for training. Inaccurate or biased data may lead to suboptimal recommendations and predictions.
- 6) **Ethical Considerations:** Ethical dilemmas may arise concerning the use of AI in influencing consumer behavior. Striking a balance between business interests and ethical considerations is crucial for the project's success and societal acceptance.

Online Shoppers Intention

Data Collection

- Import libraries and load dataset

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
%matplotlib inline
plt.style.use("fivethirtyeight")
import seaborn as sns
import warnings
warnings.filterwarnings('ignore')

df = pd.read_csv("online_shoppers_intention.csv")
# copy of the data
df_copy = df.copy()
df.head()
```

	Administrative	Administrative_Duration	Informational	\
0	0	0.0	0	
1	0	0.0	0	
2	0	0.0	0	
3	0	0.0	0	
4	0	0.0	0	

	Informational_Duration	ProductRelated	ProductRelated_Duration	\
0	0.0	1	0.000000	
1	0.0	2	64.000000	
2	0.0	1	0.000000	
3	0.0	2	2.666667	
4	0.0	10	627.500000	

	BounceRates	ExitRates	PageValues	SpecialDay	Month
OperatingSystems \					
0	0.20	0.20	0.0	0.0	Feb
1					
1	0.00	0.10	0.0	0.0	Feb
2					
2	0.20	0.20	0.0	0.0	Feb
4					
3	0.05	0.14	0.0	0.0	Feb
3					
4	0.02	0.05	0.0	0.0	Feb
3					

	Browser	Region	TrafficType	VisitorType	Weekend	Revenue
0	1	1	1	Returning_Visitor	False	False
1	2	1	2	Returning_Visitor	False	False
2	1	9	3	Returning_Visitor	False	False
3	2	2	4	Returning_Visitor	False	False
4	3	1	4	Returning_Visitor	True	False

Visualising and Analysing data

```
df.shape

(12330, 18)

df.info()

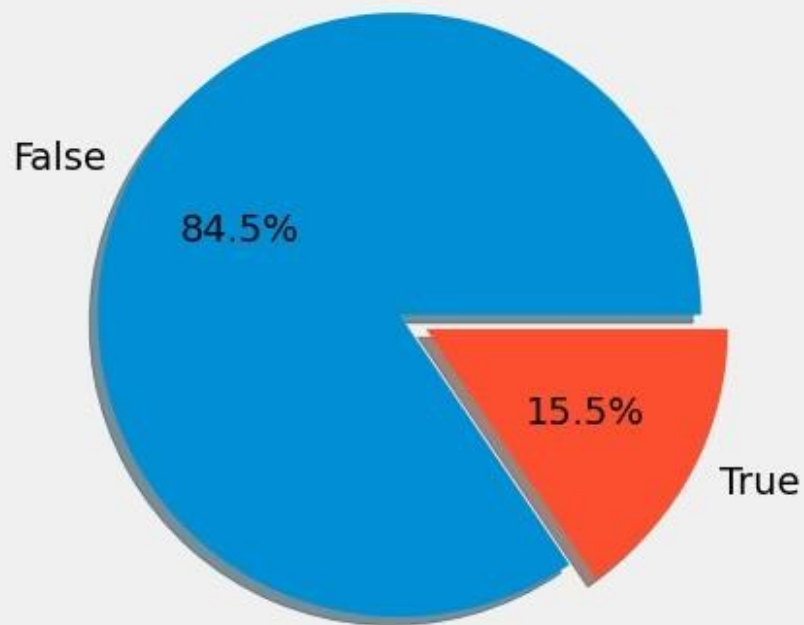
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 12330 entries, 0 to 12329
Data columns (total 18 columns):
#   Column                                Non-Null Count  Dtype
---  -
0   Administrative                        12330 non-null  int64
1   Administrative_Duration              12330 non-null  float64
2   Informational                        12330 non-null  int64
3   Informational_Duration               12330 non-null  float64
4   ProductRelated                      12330 non-null  int64
5   ProductRelated_Duration             12330 non-null  float64
6   BounceRates                         12330 non-null  float64
7   ExitRates                          12330 non-null  float64
8   PageValues                         12330 non-null  float64
9   SpecialDay                         12330 non-null  float64
10  Month                              12330 non-null  object
11  OperatingSystems                   12330 non-null  int64
12  Browser                           12330 non-null  int64
13  Region                            12330 non-null  int64
14  TrafficType                       12330 non-null  int64
15  VisitorType                       12330 non-null  object
16  Weekend                           12330 non-null  bool
17  Revenue                           12330 non-null  bool
dtypes: bool(2), float64(7), int64(7), object(2)
memory usage: 1.5+ MB
```

Data Visualisation

- Univariate Analysis

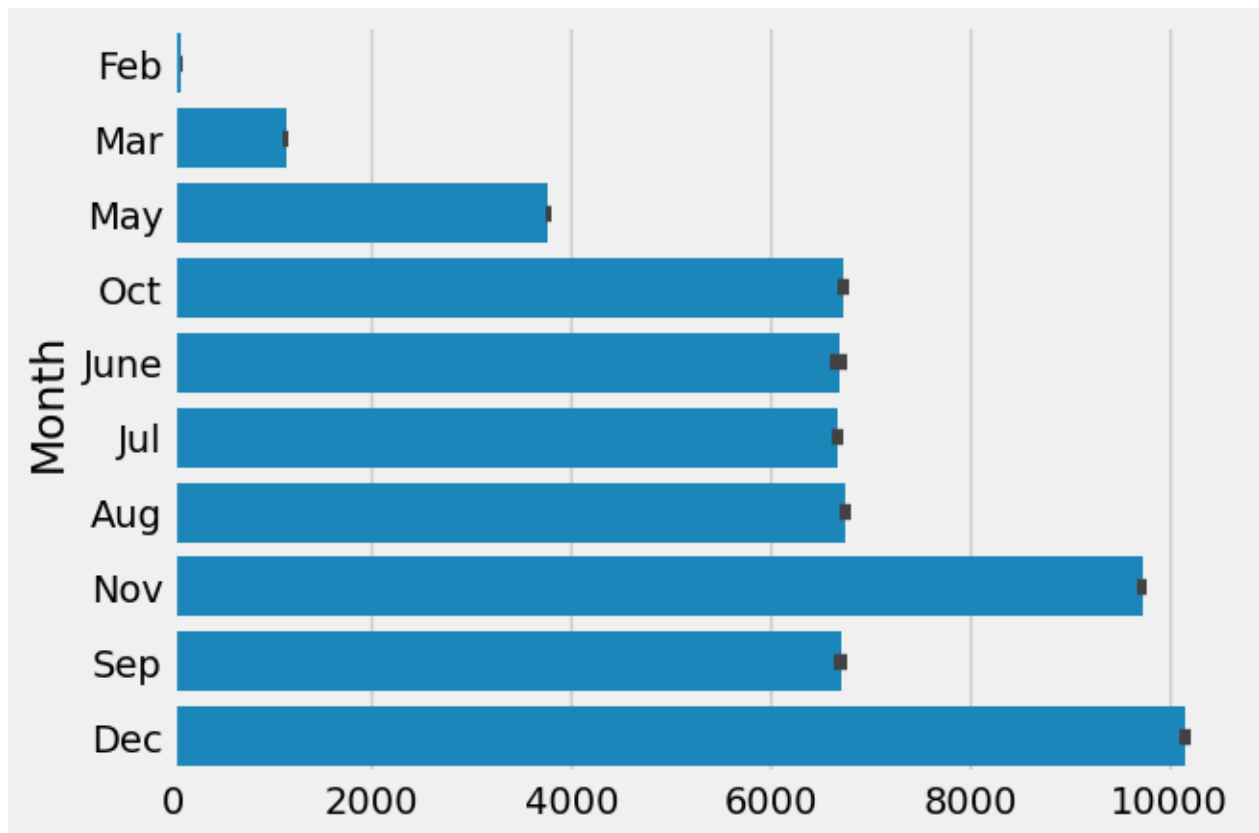
```
plt.pie(df.Revenue.value_counts(), explode= [0,0.1], labels=
df.Revenue.value_counts().index, shadow= True, autopct= "%1.1f%%" )
plt.title("Revenue Generated by visitors")
plt.show()
```

Revenue Generated by visitors

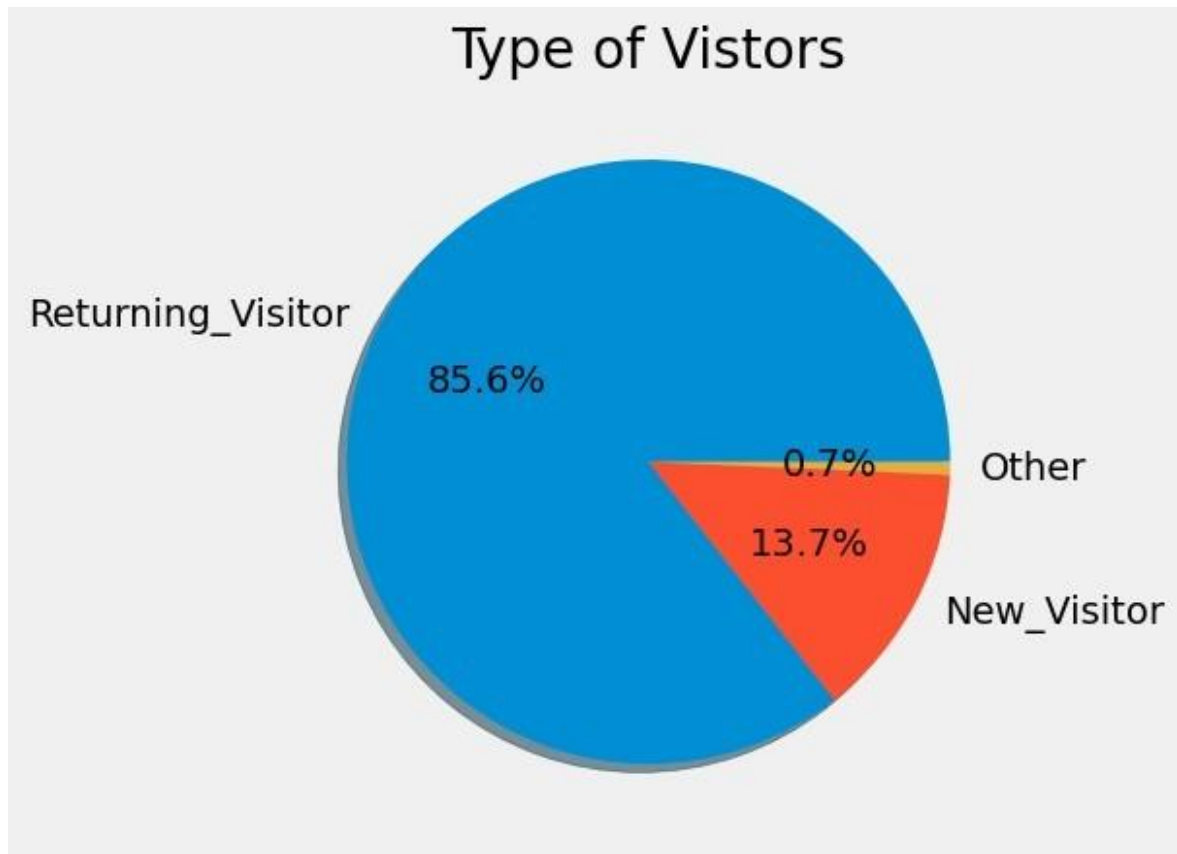


```
sns.barplot(df.Month)
```

```
<AxesSubplot: ylabel='Month'>
```

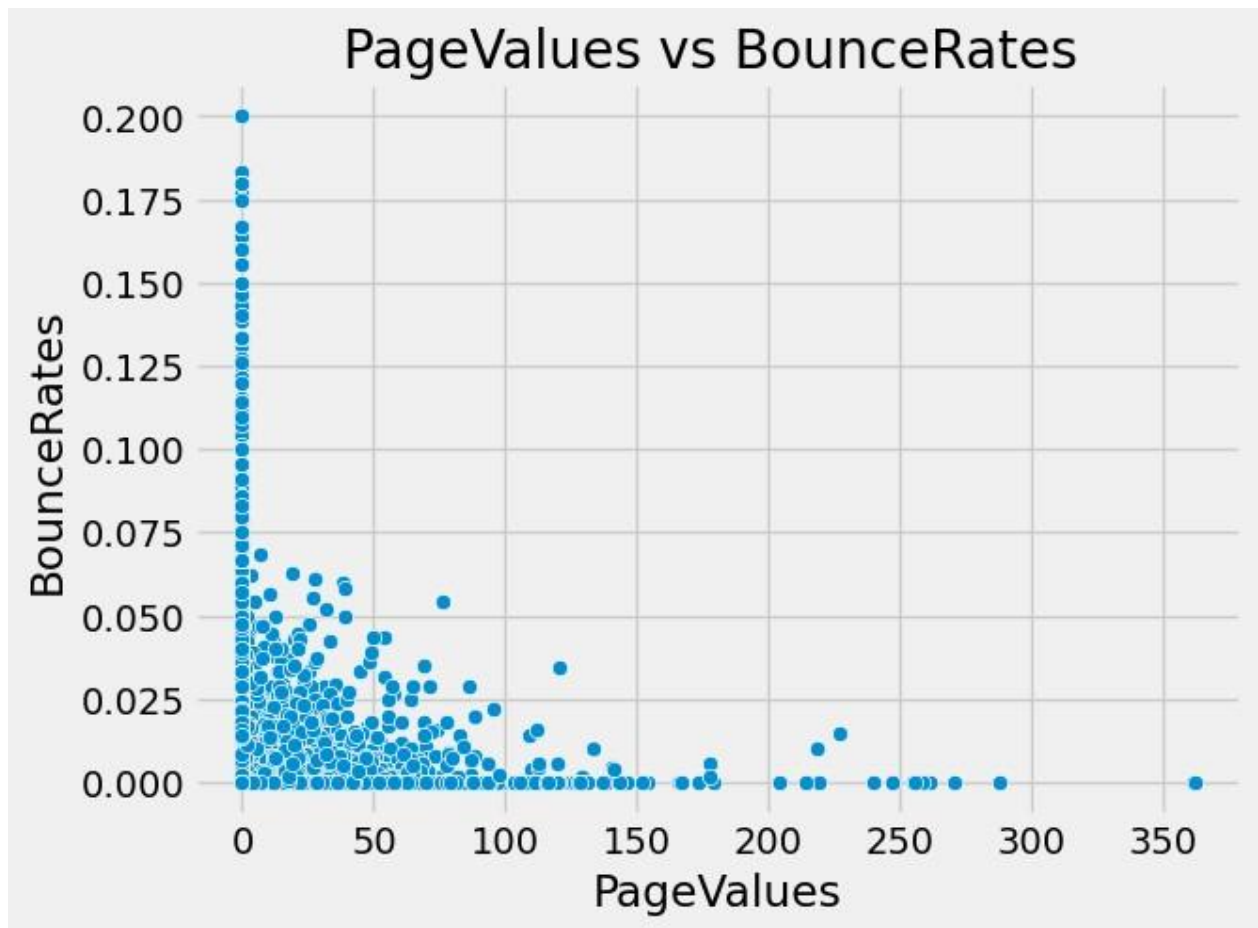



```
plt.pie(df.VisitorType.value_counts(), labels=
df.VisitorType.value_counts().index, shadow= True, autopct= "%1.1f%
%")
plt.title("Type of Vistors")
plt.show()
```

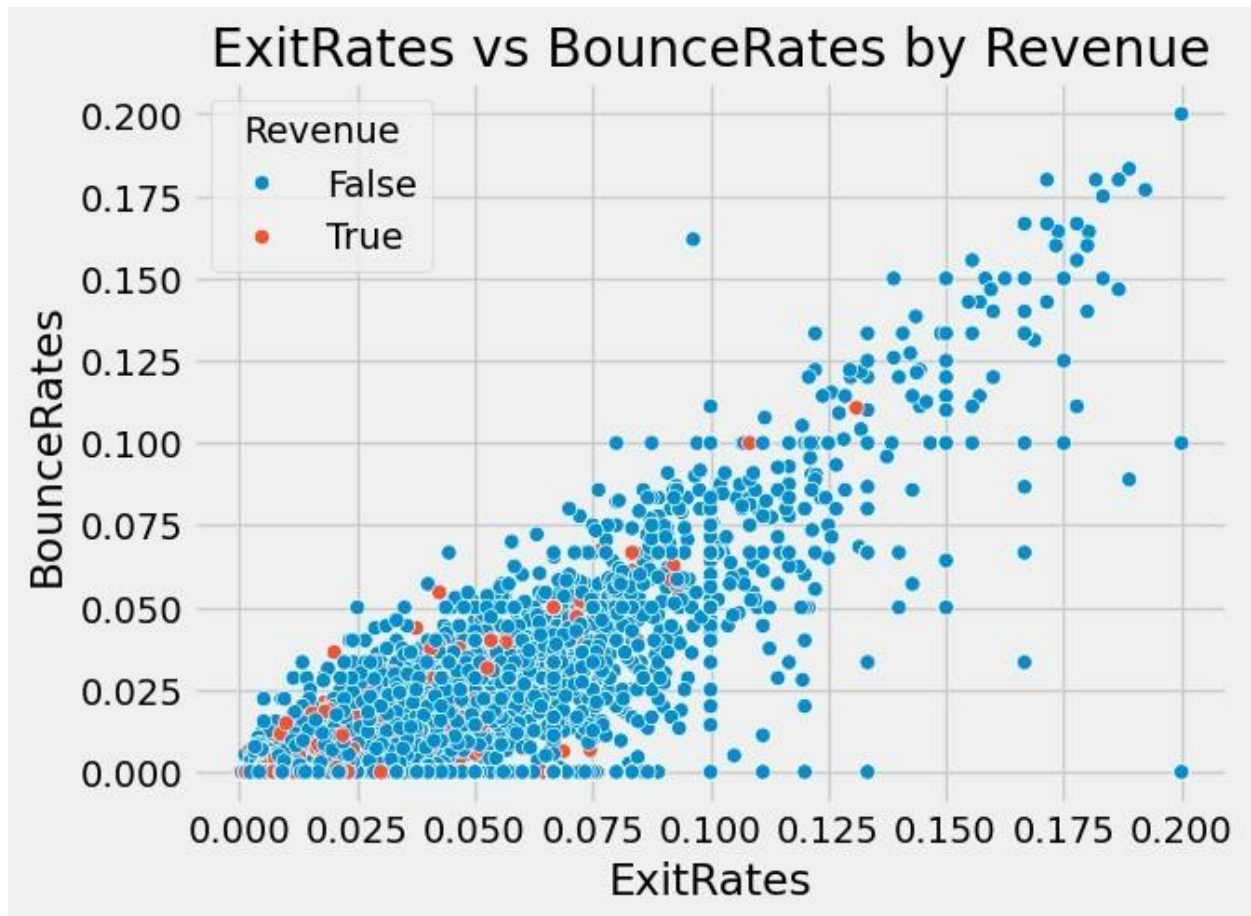


- Bivariate Analysis

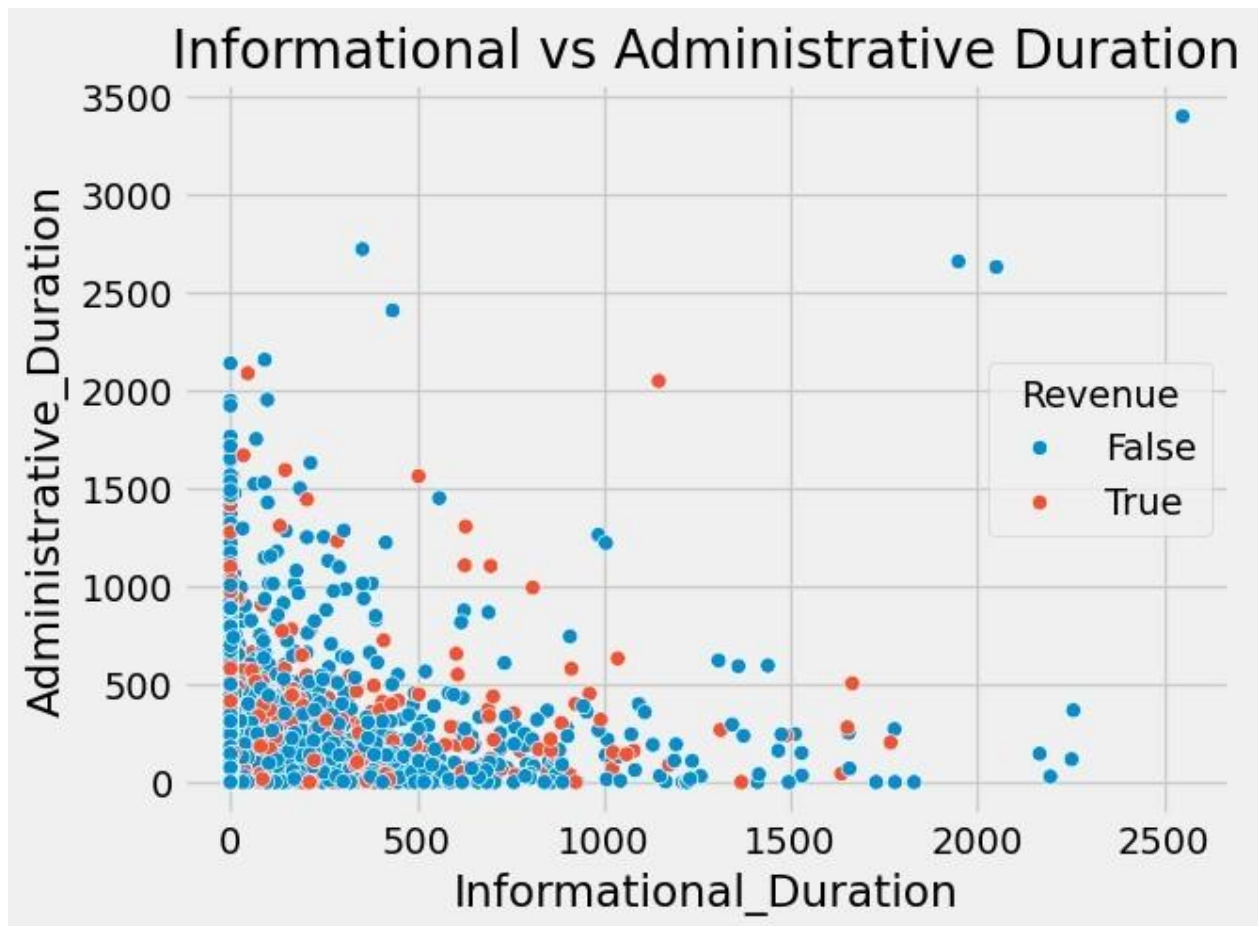
```
sns.scatterplot(data=df, x='PageValues', y='BounceRates')  
plt.xlabel('PageValues')  
plt.ylabel('BounceRates')  
plt.title('PageValues vs BounceRates ')  
plt.show()
```



```
sns.scatterplot(data=df, x='ExitRates', y='BounceRates',  
hue='Revenue')  
plt.xlabel('ExitRates')  
plt.ylabel('BounceRates')  
plt.title('ExitRates vs BounceRates by Revenue')  
plt.show()
```

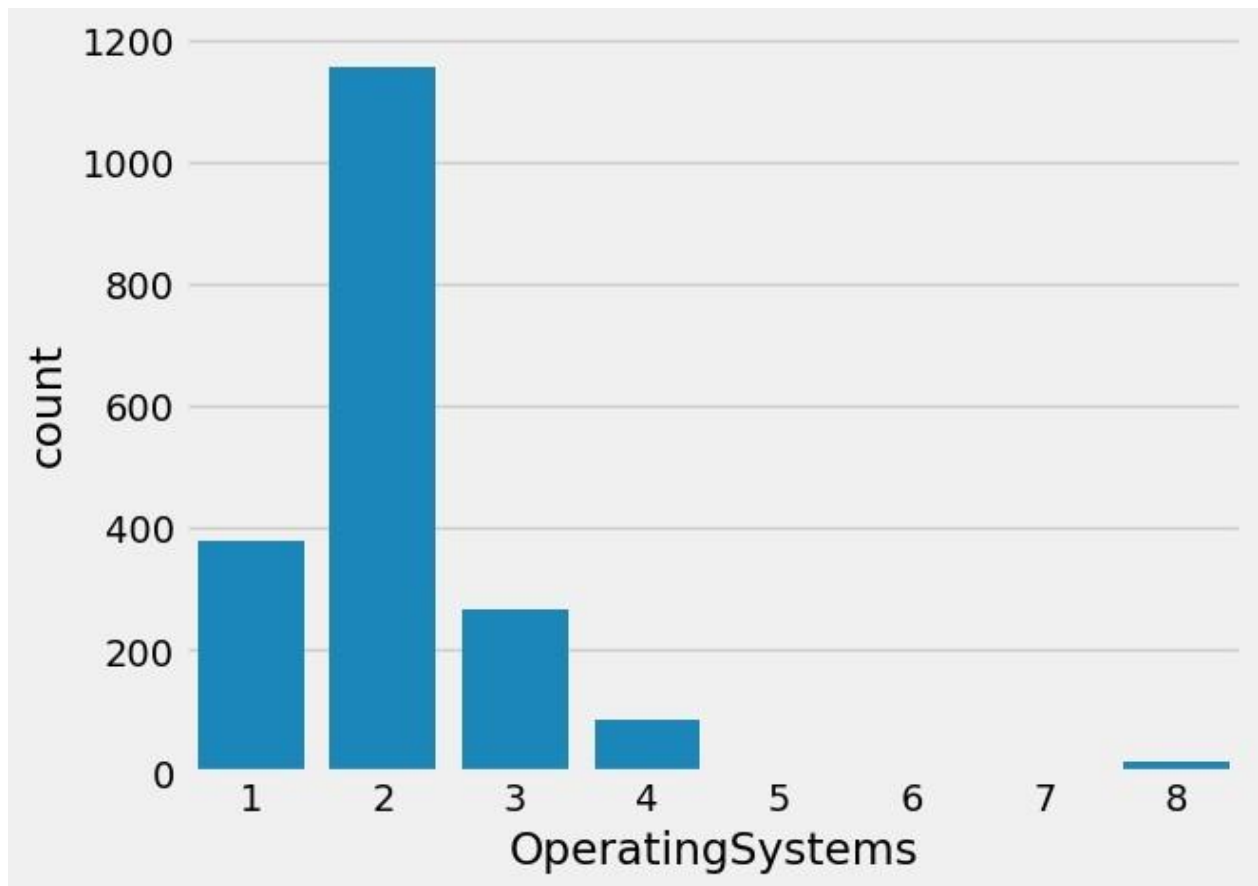


```
sns.scatterplot(data=df, x='Informational_Duration',  
y='Administrative_Duration', hue= 'Revenue')  
plt.xlabel('Informational_Duration')  
plt.ylabel('Administrative_Duration')  
plt.title('Informational vs Administrative Duration')  
plt.show()
```



```
sns.countplot(data=df[df['Revenue']==True], x='OperatingSystems')
```

```
<AxesSubplot: xlabel='OperatingSystems', ylabel='count'>
```

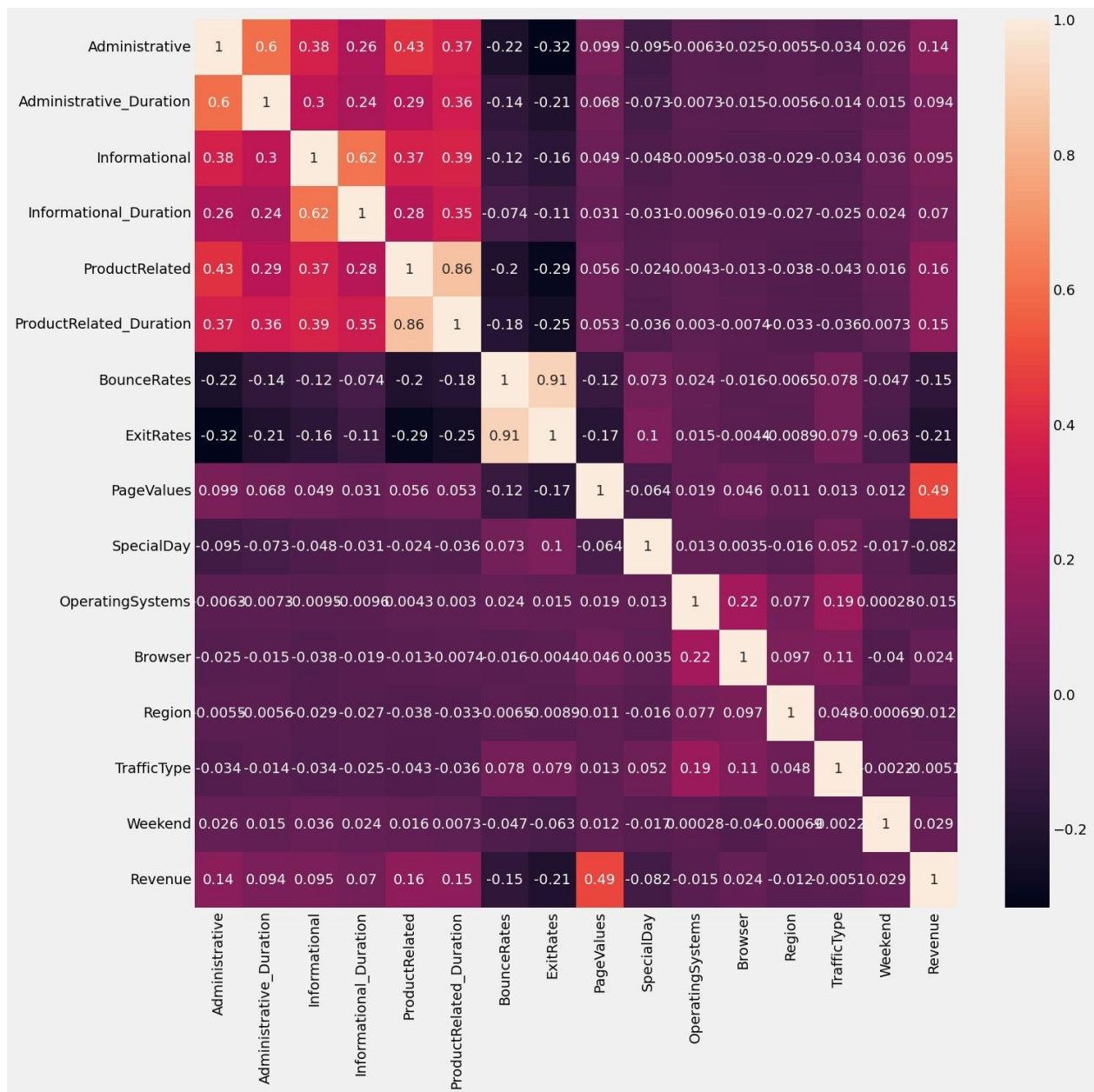


- Multivarait Analysis

```
# sns.pairplot(df)

plt.figure(figsize= (15, 15))
sns.heatmap(df.corr(numeric_only= True), annot= True)

<AxesSubplot: >
```



- Descriptive Analysis

```
df.describe(include= 'all')
```

	Administrative	Administrative_Duration	Informational	\
count	12330.000000	12330.000000	12330.000000	
unique	NaN	NaN	NaN	
top	NaN	NaN	NaN	
freq	NaN	NaN	NaN	
mean	2.315166	80.818611	0.503569	
std	3.321784	176.779107	1.270156	
min	0.000000	0.000000	0.000000	
25%	0.000000	0.000000	0.000000	

50%	1.000000	7.500000	0.000000
75%	4.000000	93.256250	0.000000
max	27.000000	3398.750000	24.000000

Informational_Duration ProductRelated		
ProductRelated_Duration \		
count	12330.000000	12330.000000
12330.000000		
unique	NaN	NaN
NaN		
top	NaN	NaN
NaN		
freq	NaN	NaN
NaN		
mean	34.472398	31.731468
1194.746220		
std	140.749294	44.475503
1913.669288		
min	0.000000	0.000000
0.000000		
25%	0.000000	7.000000
184.137500		
50%	0.000000	18.000000
598.936905		
75%	0.000000	38.000000
1464.157214		
max	2549.375000	705.000000
63973.522230		

	BounceRates	ExitRates	PageValues	SpecialDay	Month
\					
count	12330.000000	12330.000000	12330.000000	12330.000000	12330
unique	NaN	NaN	NaN	NaN	10
top	NaN	NaN	NaN	NaN	May
freq	NaN	NaN	NaN	NaN	3364
mean	0.022191	0.043073	5.889258	0.061427	NaN
std	0.048488	0.048597	18.568437	0.198917	NaN
min	0.000000	0.000000	0.000000	0.000000	NaN
25%	0.000000	0.014286	0.000000	0.000000	NaN
50%	0.003112	0.025156	0.000000	0.000000	NaN
75%	0.016813	0.050000	0.000000	0.000000	NaN

max	0.200000	0.200000	361.763742	1.000000	NaN
-----	----------	----------	------------	----------	-----

	OperatingSystems	Browser	Region	TrafficType	\
count	12330.000000	12330.000000	12330.000000	12330.000000	
unique	NaN	NaN	NaN	NaN	
top	NaN	NaN	NaN	NaN	
freq	NaN	NaN	NaN	NaN	
mean	2.124006	2.357097	3.147364	4.069586	
std	0.911325	1.717277	2.401591	4.025169	
min	1.000000	1.000000	1.000000	1.000000	
25%	2.000000	2.000000	1.000000	2.000000	
50%	2.000000	2.000000	3.000000	2.000000	
75%	3.000000	2.000000	4.000000	4.000000	
max	8.000000	13.000000	9.000000	20.000000	

	VisitorType	Weekend	Revenue
count	12330	12330	12330
unique	3	2	2
top	Returning_Visitor	False	False
freq	10551	9462	10422
mean	NaN	NaN	NaN
std	NaN	NaN	NaN
min	NaN	NaN	NaN
25%	NaN	NaN	NaN
50%	NaN	NaN	NaN
75%	NaN	NaN	NaN
max	NaN	NaN	NaN

Data Preprocessing

- Checking for null values

```
df.isnull().sum()
```

Administrative	0
Administrative_Duration	0
Informational	0
Informational_Duration	0
ProductRelated	0
ProductRelated_Duration	0
BounceRates	0
ExitRates	0
PageValues	0
SpecialDay	0
Month	0
OperatingSystems	0
Browser	0

```
Region          0
TrafficType     0
VisitorType     0
Weekend         0
Revenue         0
dtype: int64
```

- Handling Categorical Values

```
categorical_Columns =
df.select_dtypes(include=['object','bool']).columns
categorical_Columns

Index(['Month', 'VisitorType', 'Weekend', 'Revenue'], dtype='object')
```

Lets do some manual encoding because month should 1,2,3in order rather just being label encoded and also will be easier in encoding again in UI.

```
df.Month.value_counts()

Month
May      3364
Nov      2998
Mar      1907
Dec      1727
Oct       549
Sep       448
Aug       433
Jul       432
June      288
Feb       184
Name: count, dtype: int64
```

```
df.Month = df.Month.map({
    "Jan" : 1,
    "Feb" : 2,
    "Mar" : 3,
    "April" : 4,
    "May" : 5,
    "June" : 6,
    "Jul" : 7,
    "Aug" : 8,
    "Sep" : 9,
    "Oct" : 10,
    "Nov" : 11,
    "Dec" : 12
})
```

```
df.Month.value_counts()
```

```

Month
5      3364
11     2998
3      1907
12     1727
10      549
9       448
8       433
7       432
6       288
2       184
Name: count, dtype: int64

df.Revenue = df.Revenue.map({
    True : 1,
    False : 0
})

df.Weekend = df.Weekend.map({
    True : 1,
    False : 0
})

df.Revenue.value_counts()

Revenue
0      10422
1       1908
Name: count, dtype: int64

df.VisitorType.value_counts()

VisitorType
Returning_Visitor    10551
New_Visitor         1694
Other                85
Name: count, dtype: int64

df.VisitorType = df.VisitorType.map({
    "Returning_Visitor" : 0,
    "New_Visitor" : 1,
    "Other" : 2
})

df.VisitorType.value_counts()

VisitorType
0      10551
1       1694
2          85
Name: count, dtype: int64

```

- Remove Unwanted Columns

```
df.columns

Index(['Administrative', 'Administrative_Duration', 'Informational',
      'Informational_Duration', 'ProductRelated',
      'ProductRelated_Duration',
      'BounceRates', 'ExitRates', 'PageValues', 'SpecialDay',
      'Month',
      'OperatingSystems', 'Browser', 'Region', 'TrafficType',
      'VisitorType',
      'Weekend', 'Revenue'],
      dtype='object')

KMeansdf = df.drop(columns= ['Revenue'])
```

4. Model Building

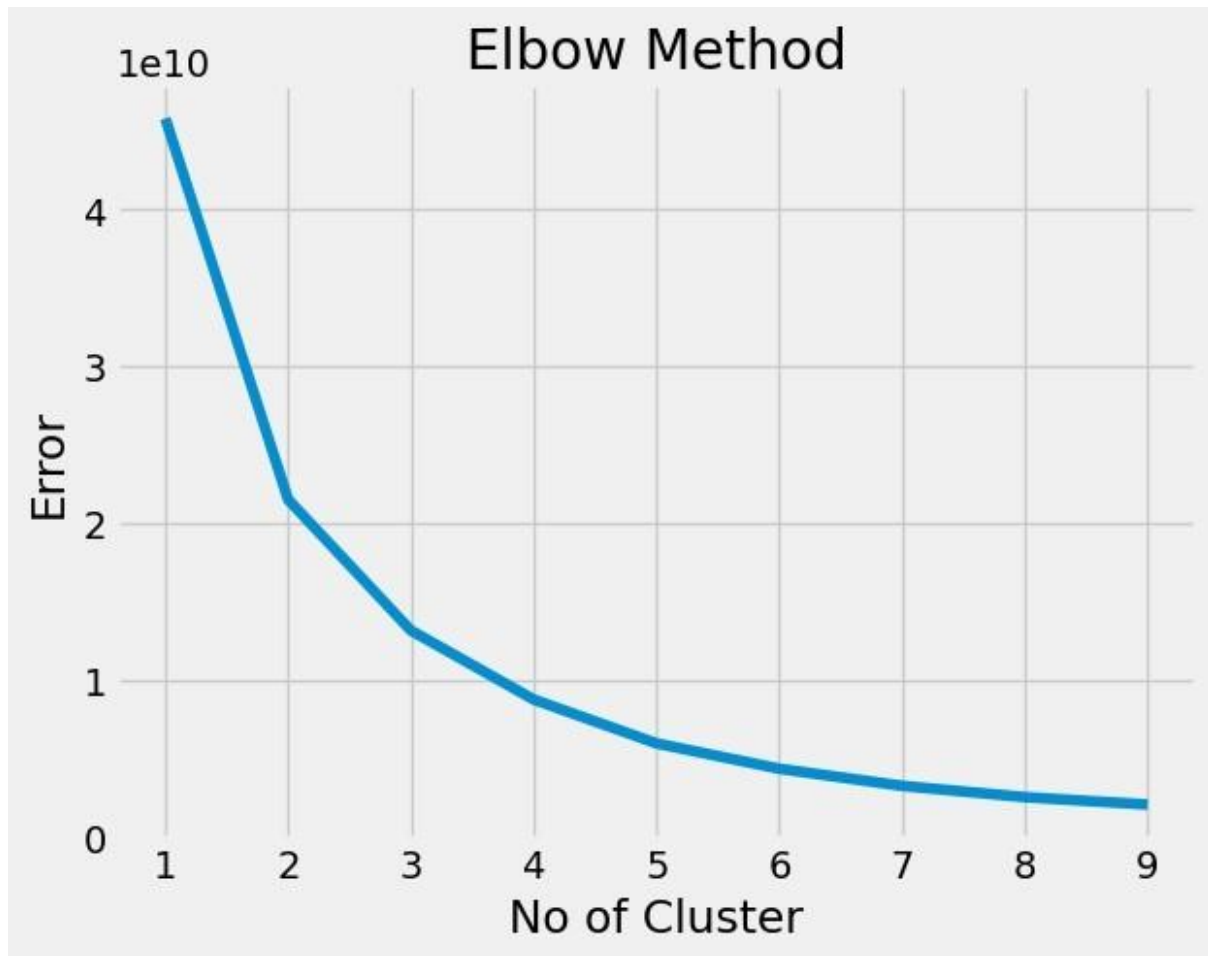
Unsupervised Machine Learning

- Elbow Method

```
from sklearn import cluster
error = []
for i in range(1,10):
    kmeans = cluster.KMeans(n_clusters= i, init= 'k-means++',
random_state=0)
    kmeans.fit(KMeansdf)
    error.append(kmeans.inertia_)
error

[45809194600.98817,
 21549400786.602127,
 13215043574.155777,
 8819549922.344086,
 6028084961.338605,
 4412518157.6448145,
 3323646299.250299,
 2604703161.924039,
 2133824458.1113849]

plt.plot(range(1, 10), error)
plt.title("Elbow Method")
plt.xlabel("No of Cluster")
plt.ylabel("Error")
plt.show()
```



- Initialize Model

```
km = cluster.KMeans(n_clusters= 4)
y_pred = km.fit_predict(KMeansdf)
```

- Dimensionality Reduction

```
from sklearn.decomposition import PCA
pca = PCA(n_components= 2)
dfPCA = pca.fit_transform(KMeansdf)
dfPCA = pd.DataFrame(dfPCA, columns= ['PCA 1', 'PCA 2'])
dfPCA['Cluster'] = y_pred
dfPCA.head()
```

	PCA 1	PCA 2	Cluster
0	-1197.636670	-40.824759	0
1	-1133.685674	-43.291841	0
2	-1197.637151	-40.824772	0
3	-1194.953161	-40.934340	0
4	-570.632670	-65.006054	0

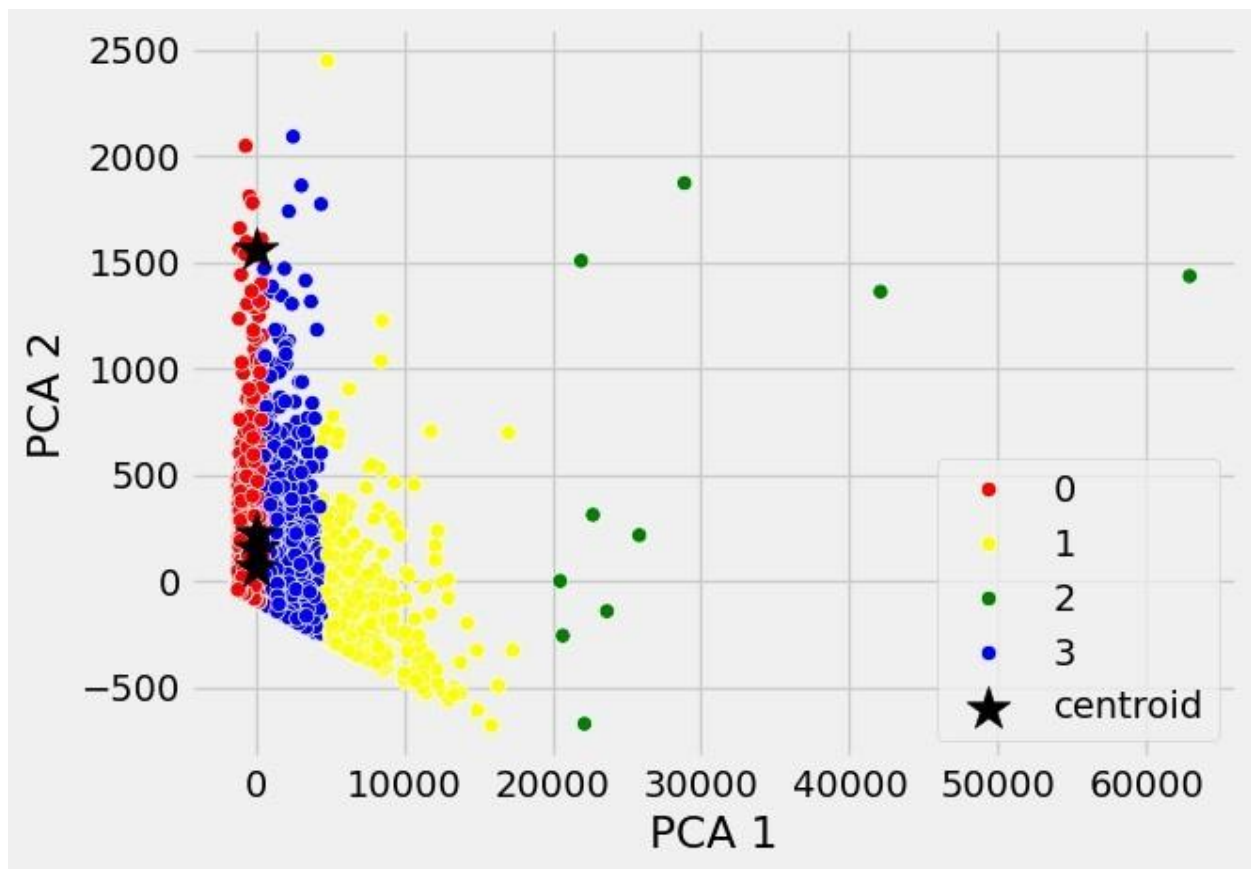
```

# Plot with seaborn
sns.scatterplot(
    data=dfPCA,
    x='PCA 1',
    y='PCA 2',
    hue='Cluster',
    palette=['red', 'yellow', 'green', 'blue']
)

# Add centroids
plt.scatter(
    km.cluster_centers_[ :,0],
    km.cluster_centers_[ :,1],
    color='black',
    s=300,
    marker='*',
    label='centroid'
)

plt.legend()
<matplotlib.legend.Legend at 0x7f2b180e06d0>

```



Supervised Machine Learning

```
# Use the previous copy made in the start of this file
df = df_copy.copy()
```

- Handling Outliers

```
def removeOutliers(df):
    for col in df.select_dtypes(include=[np.number]).columns:
        q3 = df[col].quantile(0.75)
        q1 = df[col].quantile(0.25)
        iqr = q3 - q1
        upper_limit = q3 + 1.5 * iqr
        lower_limit = q1 - 1.5 * iqr
        median = df[col].median()
        df[col] = np.where((df[col] > upper_limit) | (df[col] <
lower_limit), median, df[col])
        print("Finished Removing Outliers")
    return df
# df = removeOutliers(df) Not usign cause : the model accuracy
drastically drops upon outlier removal
```

- Handling Categorical Values

```
categorical_Columns =
df.select_dtypes(include=['object', 'bool']).columns
categorical_Columns

Index(['Month', 'VisitorType', 'Weekend', 'Revenue'], dtype='object')
```

Lets do some manual encoding because month should 1,2,3in order rather just being label encoded and also will be easier in encoding again in UI.

```
df.Month.value_counts()

Month
May      3364
Nov      2998
Mar      1907
Dec      1727
Oct       549
Sep       448
Aug       433
Jul       432
June      288
Feb       184
Name: count, dtype: int64

df.Month = df.Month.map({
    "Jan" : 1,
```

```

    "Feb" : 2,
    "Mar" : 3,
    "April" : 4,
    "May" : 5,
    "June" : 6,
    "Jul" : 7,
    "Aug" : 8,
    "Sep" : 9,
    "Oct" : 10,
    "Nov" : 11,
    "Dec" : 12
})

df.Month.value_counts()

Month
5      3364
11     2998
3      1907
12     1727
10      549
9       448
8       433
7       432
6       288
2       184
Name: count, dtype: int64

df.Revenue = df.Revenue.map({
    True : 1,
    False : 0
})

df.Weekend = df.Weekend.map({
    True : 1,
    False : 0
})

df.Revenue.value_counts()

Revenue
0      10422
1       1908
Name: count, dtype: int64

df.VisitorType.value_counts()

VisitorType
Returning_Visitor    10551
New_Visitor         1694

```



```

Other      85
Name: count, dtype: int64

df.VisitorType = df.VisitorType.map({
    "Returning_Visitor" : 0,
    "New_Visitor" : 1,
    "Other" : 2
})

df.VisitorType.value_counts()

VisitorType
0      10551
1       1694
2         85
Name: count, dtype: int64

```

- **Scaling the features**

```

from sklearn.preprocessing import MinMaxScaler
scaler = MinMaxScaler()
df = pd.DataFrame(scaler.fit_transform(df), columns= df.columns)
df.head()

```

	Administrative	Administrative_Duration	Informational	\
0	0.0	0.0	0.0	
1	0.0	0.0	0.0	
2	0.0	0.0	0.0	
3	0.0	0.0	0.0	
4	0.0	0.0	0.0	

	Informational_Duration	ProductRelated	ProductRelated_Duration	\
0	0.0	0.001418	0.000000	
1	0.0	0.002837	0.001000	
2	0.0	0.001418	0.000000	
3	0.0	0.002837	0.000042	
4	0.0	0.014184	0.009809	

OperatingSystems	\
0	1.00 1.00 0.0 0.0 0.0
0.000000	
1	0.00 0.50 0.0 0.0 0.0
0.142857	
2	1.00 1.00 0.0 0.0 0.0
0.428571	
3	0.25 0.70 0.0 0.0 0.0
0.285714	
4	0.10 0.25 0.0 0.0 0.0
0.285714	

	Browser	Region	TrafficType	VisitorType	Weekend	Revenue
0	0.000000	0.000	0.000000	0.0	0.0	0.0
1	0.083333	0.000	0.052632	0.0	0.0	0.0
2	0.000000	1.000	0.105263	0.0	0.0	0.0
3	0.083333	0.125	0.157895	0.0	0.0	0.0
4	0.166667	0.000	0.157895	0.0	1.0	0.0

- **Splitting Dataset**

```
from sklearn.model_selection import train_test_split

x = df.drop(columns= 'Revenue')
y = df.Revenue

x_train, x_test, y_train, y_test = train_test_split(x, y, test_size=
0.2, random_state=42)

# Lets define a evaluator that does metrics

def evaluator(model, x_test, y_test, x_train, y_train):
    from sklearn.metrics import accuracy_score, confusion_matrix,
classification_report
    y_pred = model.predict(x_test)
    y_pred_train = model.predict(x_train)
    print("Training Accuracy: ", accuracy_score(y_train,
y_pred_train))
    print(f"Testing Accuracy: {accuracy_score(y_test, y_pred)}")
    print(f"Confusion Matrix:\n {confusion_matrix(y_test, y_pred)}")
    print(f"Classification Report: \n {classification_report(y_test,
y_pred)}")
```

- **Logistic Regression**

```
from sklearn.linear_model import LogisticRegression

lr = LogisticRegression()
lr.fit(x_train, y_train)
evaluator(lr,x_test, y_test, x_train, y_train)

Training Accuracy: 0.8790551500405515
Testing Accuracy: 0.8605028386050284
Confusion Matrix:
[[2021  34]
 [ 310 101]]
Classification Report:
```

	precision	recall	f1-score	support
0.0	0.87	0.98	0.92	2055
1.0	0.75	0.25	0.37	411
accuracy			0.86	2466

macro avg	0.81	0.61	0.65	2466
weighted avg	0.85	0.86	0.83	2466

- Decision Tree

```
from sklearn.tree import DecisionTreeClassifier

dt = DecisionTreeClassifier()
dt.fit(x_train, y_train)
evaluator(dt, x_test, y_test, x_train, y_train)

Training Accuracy: 1.0
Testing Accuracy: 0.8556366585563666
Confusion Matrix:
[[1876  179]
 [ 177  234]]
Classification Report:
              precision    recall  f1-score   support

    0.0               0.91       0.91       0.91       2055
    1.0               0.57       0.57       0.57        411

 accuracy               0.86               0.86       2466
 macro avg              0.74       0.74       0.74       2466
 weighted avg           0.86       0.86       0.86       2466
```

- Random Forest Classifier

```
from sklearn.ensemble import RandomForestClassifier

rfc = RandomForestClassifier()
rfc.fit(x_train, y_train)
evaluator(rfc, x_test, y_test, x_train, y_train)

Training Accuracy: 0.9998986212489862
Testing Accuracy: 0.8917274939172749
Confusion Matrix:
[[1971   84]
 [ 183  228]]
Classification Report:
              precision    recall  f1-score   support

    0.0               0.92       0.96       0.94       2055
    1.0               0.73       0.55       0.63        411

 accuracy               0.89               0.89       2466
 macro avg              0.82       0.76       0.78       2466
 weighted avg           0.88       0.89       0.89       2466
```

- **Support Vector Classifier**

```
from sklearn.svm import SVC
svc = SVC()
svc.fit(x_train, y_train)
evaluator(svc, x_test, y_test, x_train, y_train)
```

Training Accuracy: 0.8866585563665855
 Testing Accuracy: 0.8661800486618005
 Confusion Matrix:

```
[[2016  39]
 [ 291 120]]
```

Classification Report:

	precision	recall	f1-score	support
0.0	0.87	0.98	0.92	2055
1.0	0.75	0.29	0.42	411
accuracy			0.87	2466
macro avg	0.81	0.64	0.67	2466
weighted avg	0.85	0.87	0.84	2466

- **Gaussian Naive Bayes Classifier**

```
from sklearn.naive_bayes import GaussianNB
gnb = GaussianNB()
gnb.fit(x_train, y_train)
evaluator(gnb, x_test, y_test, x_train, y_train)
```

Training Accuracy: 0.8033252230332523
 Testing Accuracy: 0.7980535279805353
 Confusion Matrix:

```
[[1690  365]
 [ 133  278]]
```

Classification Report:

	precision	recall	f1-score	support
0.0	0.93	0.82	0.87	2055
1.0	0.43	0.68	0.53	411
accuracy			0.80	2466
macro avg	0.68	0.75	0.70	2466
weighted avg	0.84	0.80	0.81	2466

- **K Nearest Neighbours Classifier**

```
from sklearn.neighbors import KNeighborsClassifier
knn = KNeighborsClassifier()
knn.fit(x_train, y_train)
evaluator(knn, x_test, y_test, x_train, y_train)
```

Training Accuracy: 0.8864557988645579

Testing Accuracy: 0.85117599351176

Confusion Matrix:

```
[[2012  43]
```

```
[ 324  87]]
```

Classification Report:

	precision	recall	f1-score	support
0.0	0.86	0.98	0.92	2055
1.0	0.67	0.21	0.32	411
accuracy			0.85	2466
macro avg	0.77	0.60	0.62	2466
weighted avg	0.83	0.85	0.82	2466

- Cross Validation

```
from sklearn.model_selection import cross_val_score
print(f"The cross validation score: {np.mean(cross_val_score(rfc, x,
y, cv= 5))}")
```

The cross validation score: 0.8941605839416058

- Save the Model

```
import pickle
pickle.dump(rfc, open("rfc.pkl", 'wb'))

!pip freeze > requirements.txt
```

Implement SMOTE to increase model performace

```
from imblearn.over_sampling import SMOTE
smote = SMOTE()
x_train, y_train = smote.fit_resample(x_train, y_train)
y_train.value_counts()
```

Revenue

```
0.0    8367
```

```
1.0    8367
```

Name: count, dtype: int64

```
from sklearn.linear_model import LogisticRegression
```

```
lr = LogisticRegression()
```

```
lr.fit(x_train, y_train)
evaluator(lr, x_test, y_test, x_train, y_train)
```

Training Accuracy: 0.8323771961276443

Testing Accuracy: 0.8479318734793188

Confusion Matrix:

```
[[1801 254]
 [ 121 290]]
```

Classification Report:

	precision	recall	f1-score	support
0.0	0.94	0.88	0.91	2055
1.0	0.53	0.71	0.61	411
accuracy			0.85	2466
macro avg	0.74	0.79	0.76	2466
weighted avg	0.87	0.85	0.86	2466

```
from sklearn.tree import DecisionTreeClassifier
```

```
dt = DecisionTreeClassifier()
dt.fit(x_train, y_train)
evaluator(dt, x_test, y_test, x_train, y_train)
```

Training Accuracy: 1.0

Testing Accuracy: 0.8519870235198702

Confusion Matrix:

```
[[1848 207]
 [ 158 253]]
```

Classification Report:

	precision	recall	f1-score	support
0.0	0.92	0.90	0.91	2055
1.0	0.55	0.62	0.58	411
accuracy			0.85	2466
macro avg	0.74	0.76	0.75	2466
weighted avg	0.86	0.85	0.86	2466

```
from sklearn.ensemble import RandomForestClassifier
```

```
rfc = RandomForestClassifier()
rfc.fit(x_train, y_train)
evaluator(rfc, x_test, y_test, x_train, y_train)
```

Training Accuracy: 1.0

Testing Accuracy: 0.8884833738848338

Confusion Matrix:

```
[[1899 156]
 [ 119 292]]
```

Classification Report:

	precision	recall	f1-score	support
0.0	0.94	0.92	0.93	2055
1.0	0.65	0.71	0.68	411
accuracy			0.89	2466
macro avg	0.80	0.82	0.81	2466
weighted avg	0.89	0.89	0.89	2466

```
from sklearn.svm import SVC
svc = SVC()
svc.fit(x_train, y_train)
evaluator(svc, x_test, y_test, x_train, y_train)
```

Training Accuracy: 0.8572367634755588

Testing Accuracy: 0.8690186536901865

Confusion Matrix:

```
[[1844 211]
 [ 112 299]]
```

Classification Report:

	precision	recall	f1-score	support
0.0	0.94	0.90	0.92	2055
1.0	0.59	0.73	0.65	411
accuracy			0.87	2466
macro avg	0.76	0.81	0.78	2466
weighted avg	0.88	0.87	0.87	2466

```
from sklearn.naive_bayes import GaussianNB
gnb = GaussianNB()
gnb.fit(x_train, y_train)
evaluator(gnb, x_test, y_test, x_train, y_train)
```

Training Accuracy: 0.6977411258515597

Testing Accuracy: 0.5360908353609084

Confusion Matrix:

```
[[ 968 1087]
 [  57 354]]
```

Classification Report:

	precision	recall	f1-score	support
0.0	0.94	0.47	0.63	2055
1.0	0.25	0.86	0.38	411
accuracy			0.54	2466
macro avg	0.60	0.67	0.51	2466

weighted avg	0.83	0.54	0.59	2466
--------------	------	------	------	------

```
from sklearn.neighbors import KNeighborsClassifier
knn = KNeighborsClassifier()
knn.fit(x_train, y_train)
evaluator(knn, x_test, y_test, x_train, y_train)
```

Training Accuracy: 0.904027727978965

Testing Accuracy: 0.7181670721816708

Confusion Matrix:

```
[[1535  520]
 [ 175  236]]
```

Classification Report:

	precision	recall	f1-score	support
0.0	0.90	0.75	0.82	2055
1.0	0.31	0.57	0.40	411
accuracy			0.72	2466
macro avg	0.60	0.66	0.61	2466
weighted avg	0.80	0.72	0.75	2466

```
from sklearn.model_selection import cross_val_score
print(f"The cross validation score: {np.mean(cross_val_score(rfc, x,
y, cv= 5))}")
```

The cross validation score: 0.8940794809407947

HyperParameter Tuning

```
from sklearn.model_selection import GridSearchCV
from sklearn.ensemble import RandomForestClassifier

# Define parameter grid
param_grid = {"n_estimators": [100, 200, 300],
              "max_depth": [5, 8],
              "min_samples_split": [2, 5, 10],
              "min_samples_leaf": [1, 5, 8]}

# Create grid search object
rf = RandomForestClassifier(n_jobs=-1)
grid_search = GridSearchCV(rf, param_grid=param_grid, cv=5,
scoring='accuracy')

# Fit grid search
grid_search.fit(x_train, y_train)
```



```
# View best parameters
print(grid_search.best_params_)

evaluator(grid_search,x_test,y_test, x_train, y_train)

{'max_depth': 8, 'min_samples_leaf': 5, 'min_samples_split': 2,
'n_estimators': 300}
Training Accuracy: 0.9117365841998327
Testing Accuracy: 0.8771289537712895
Confusion Matrix:
```

Classification Report:

	precision	recall	f1-score	support
0.0	0.96	0.89	0.92	2055
1.0	0.60	0.81	0.69	411
accuracy			0.88	2466
macro avg	0.78	0.85	0.80	2466
weighted avg	0.90	0.88	0.88	2466

Lets try if removing non corelated col's will increase accuracy .

```
df.corr(numeric_only=True).Revenue.sort_values()
```

ExitRates	-0.207071
BounceRates	-0.150673
SpecialDay	-0.082305
OperatingSystems	-0.014668
Region	-0.011595
TrafficType	-0.005113
Browser	0.023984
Weekend	0.029295
Informational_Duration	0.070345
Administrative_Duration	0.093587
Informational	0.095200
VisitorType	0.098485
Month	0.127372
Administrative	0.138917
ProductRelated_Duration	0.152373
ProductRelated	0.158538
PageValues	0.492569
Revenue	1.000000

```

from sklearn.model_selection import train_test_split

x = df.drop(columns= ['Revenue', 'SpecialDay', 'OperatingSystems',
'Region', 'TrafficType', 'Browser', 'Weekend',
'Informational_Duration', 'Administrative_Duration', 'Informational'])
y = df.Revenue

x_train, x_test, y_train, y_test = train_test_split(x, y, test_size=
0.2, random_state=42)

```

- **Logistic Regression**

```

from sklearn.linear_model import LogisticRegression

lr = LogisticRegression()
lr.fit(x_train, y_train)
evaluator(lr,x_test, y_test, x_train, y_train)

Training Accuracy: 0.8790551500405515
Testing Accuracy: 0.862124898621249
Confusion Matrix:
[[2022  33]
 [ 307 104]]
Classification Report:

```

	precision	recall	f1-score	support
0.0	0.87	0.98	0.92	2055
1.0	0.76	0.25	0.38	411
accuracy			0.86	2466
macro avg	0.81	0.62	0.65	2466
weighted avg	0.85	0.86	0.83	2466

- **Decision Tree**

```

from sklearn.tree import DecisionTreeClassifier

dt = DecisionTreeClassifier()
dt.fit(x_train, y_train)
evaluator(dt, x_test, y_test, x_train, y_train)

Training Accuracy: 0.9996958637469586
Testing Accuracy: 0.8507704785077048
Confusion Matrix:
[[1869  186]
 [ 182  229]]
Classification Report:

```

	precision	recall	f1-score	support
0.0	0.91	0.91	0.91	2055

	1.0	0.55	0.56	0.55	411
accuracy				0.85	2466
macro avg	0.73	0.73	0.73	0.73	2466
weighted avg	0.85	0.85	0.85	0.85	2466

- **Random Forest Classifier**

```
from sklearn.ensemble import RandomForestClassifier
rfc = RandomForestClassifier()
rfc.fit(x_train, y_train)
evaluator(rfc, x_test, y_test, x_train, y_train)
```

Training Accuracy: 0.9996958637469586
Testing Accuracy: 0.8909164639091647
Confusion Matrix:
[[1972 83]
[186 225]]

Classification Report:

	precision	recall	f1-score	support
0.0	0.91	0.96	0.94	2055
1.0	0.73	0.55	0.63	411
accuracy			0.89	2466
macro avg	0.82	0.75	0.78	2466
weighted avg	0.88	0.89	0.88	2466

- **Support Vector Classifier**

```
from sklearn.svm import SVC
svc = SVC()
svc.fit(x_train, y_train)
evaluator(svc, x_test, y_test, x_train, y_train)
```

Training Accuracy: 0.8917274939172749
Testing Accuracy: 0.8726682887266829
Confusion Matrix:
[[2010 45]
[269 142]]

Classification Report:

	precision	recall	f1-score	support
0.0	0.88	0.98	0.93	2055
1.0	0.76	0.35	0.47	411
accuracy			0.87	2466
macro avg	0.82	0.66	0.70	2466

weighted avg	0.86	0.87	0.85	2466
--------------	------	------	------	------

- **Gaussian Naive Bayes Classifier**

```
from sklearn.naive_bayes import GaussianNB
gnb = GaussianNB()
gnb.fit(x_train, y_train)
evaluator(gnb, x_test, y_test, x_train, y_train)
```

Training Accuracy: 0.8353609083536091
Testing Accuracy: 0.8317112733171127
Confusion Matrix:

```
[[1784  271]
 [ 144  267]]
```

Classification Report:

	precision	recall	f1-score	support
0.0	0.93	0.87	0.90	2055
1.0	0.50	0.65	0.56	411
accuracy			0.83	2466
macro avg	0.71	0.76	0.73	2466
weighted avg	0.85	0.83	0.84	2466

- **K Nearest Neighbours Classifier**

```
from sklearn.neighbors import KNeighborsClassifier
knn = KNeighborsClassifier()
knn.fit(x_train, y_train)
evaluator(knn, x_test, y_test, x_train, y_train)
```

Training Accuracy: 0.9103811841038119
Testing Accuracy: 0.8767234387672344
Confusion Matrix:

```
[[1981   74]
 [ 230  181]]
```

Classification Report:

	precision	recall	f1-score	support
0.0	0.90	0.96	0.93	2055
1.0	0.71	0.44	0.54	411
accuracy			0.88	2466
macro avg	0.80	0.70	0.74	2466
weighted avg	0.86	0.88	0.86	2466

- **Cross Validation**

```
from sklearn.model_selection import cross_val_score
print(f"The cross validation score: {np.mean(cross_val_score(rfc, x,
y, cv= 5))}")
```

The cross validation score: 0.8910786699107867

- Save the Model

```
import pickle
pickle.dump(rfc, open("rfc.pkl", 'wb'))

!pip freeze > requirements.txt
```

Conclusion:

- The Random Forest Classifier model that has SMOTE data performs best and has a good classification score and also no of wrong predictions for minority class is good , and that matches with our usecase of wheather the user generates 'revenue' or not.