**Project Report Format**

1. **INTRODUCTION**
   1. Project Overview
   2. Purpose
2. **LITERATURE SURVEY**
   1. Existing problem
   2. References
   3. Problem Statement Definition
3. **IDEATION & PROPOSED SOLUTION**
   1. Empathy Map Canvas
   2. Ideation & Brainstorming
4. **REQUIREMENT ANALYSIS**
   1. Functional requirement
   2. Non-Functional requirements
5. **PROJECT DESIGN**
   1. Data Flow Diagrams & User Stories
   2. Solution Architecture
6. **PROJECT PLANNING & SCHEDULING**
   1. Technical Architecture
   2. Sprint Planning & Estimation
   3. Sprint Delivery Schedule
7. **CODING & SOLUTIONING (Explain the features added in the project along with code)**
   1. Feature 1
   2. Feature 2
   3. Database Schema (if Applicable)
8. **PERFORMANCE TESTING**
   1. Performance Metrics
9. **RESULTS**
   1. Output Screenshots
10. **ADVANTAGES & DISADVANTAGES**
11. **CONCLUSION**
12. **FUTURE SCOPE 13. APPENDIX** Source Code

GitHub & Project Demo Link

1. **Introduction**
   1. **Project Overview**

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 project overview entails utilizing machine learning techniques to analyze and predict online shoppers' intentions on e-commerce platforms. This includes gathering and preprocessing relevant data like user interactions, purchase history, and demographics. To extract meaningful patterns, feature engineering is used, followed by the construction and training of machine learning models. These models are used to forecast a variety of intentions, such as purchase likelihood, product interests, or user segmentation. By understanding and predicting online shoppers' behavior, the ultimate goal is to improve user experience, personalize marketing efforts, and improve sales forecasting.

* 1. **Purpose**

The purpose of this project is to harness machine learning to deeply understand online shoppers' behaviors, predict their intentions, and leverage these insights to enhance user experiences, optimize marketing strategies, and drive improved sales outcomes on e-commerce platforms.

1. **Literature Survey**
   1. **Existing Problem**

**Data Quality and Quantity:** Limited or poor-quality data might hinder model accuracy. Incomplete or biased data could lead to skewed predictions.

**Overfitting and Generalization:** Models might overfit to the training data, impacting their ability to generalize to new shopper behaviors.

**Privacy and Ethics:** Balancing data utilization with user privacy concerns and ethical considerations is crucial.

**Changing Shopper Behavior:** Rapid changes in consumer behavior can make trained models outdated, requiring continuous retraining.

**Model Deployment:** Integrating ML models into existing e-commerce systems while maintaining performance and stability can pose technical challenges.

* 1. **References**
  2. **Problem Statement Definition**

Title: Predictive Analysis of Online Shopper Behavior Using Machine Learning

**Objective:**

The objective of this project is to employ machine learning techniques to comprehend and predict online shoppers' intentions on an e-commerce platform. The primary goal is to enhance user experience, optimize marketing strategies, and improve sales outcomes by accurately predicting and understanding shopper behaviors.

**Scope:**

This project aims to analyze diverse datasets encompassing user interactions, browsing history, purchase patterns, demographic information, and other relevant variables. The focus will be on predicting purchase likelihood, identifying product interests, and segmenting users based on behavior patterns.

**Target Audience:**

The project targets regular users of the e-commerce platform, aiming to understand and predict behaviors of both new and returning customers across various demographics and purchase histories.

**Data Focus:**

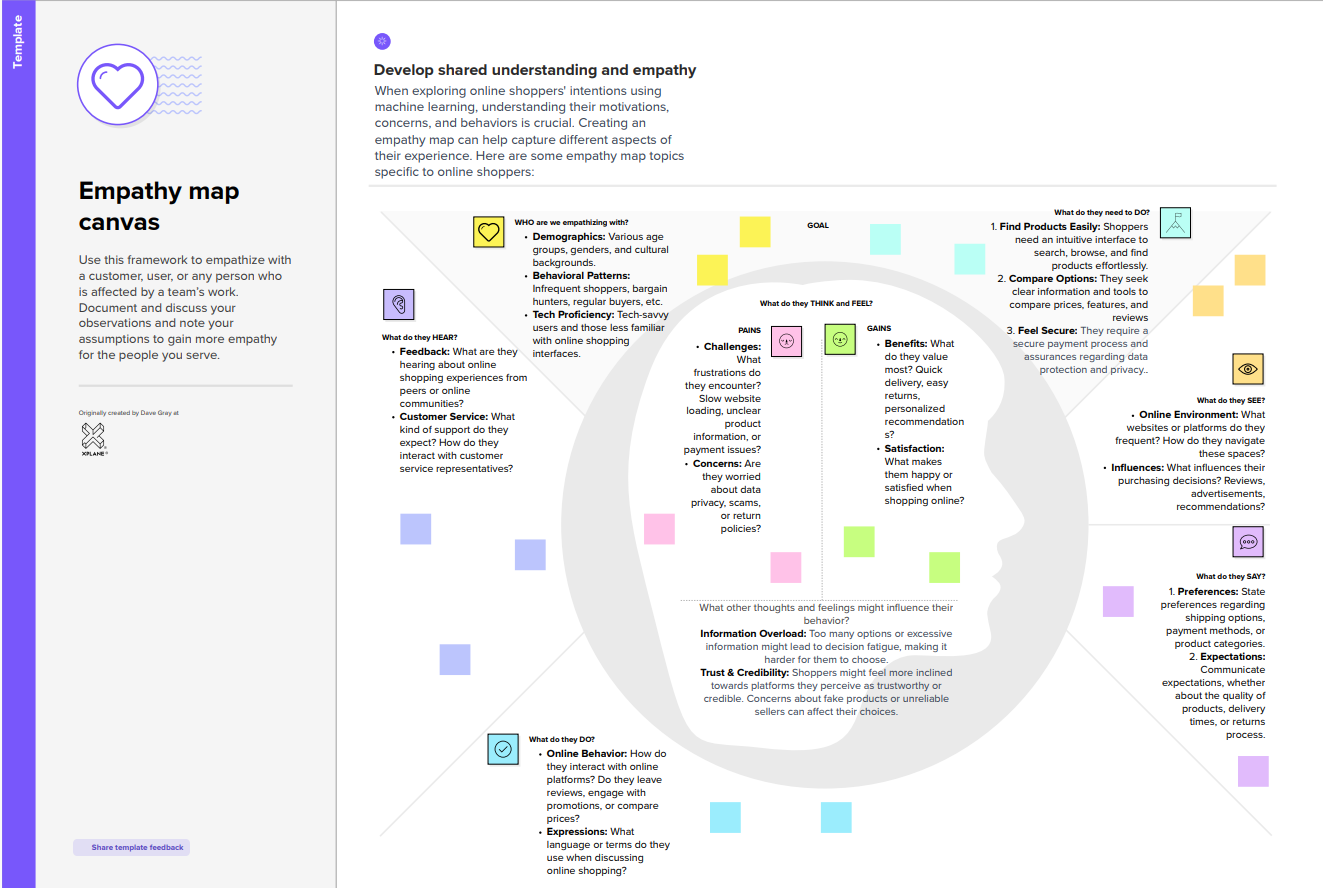
Utilizing comprehensive datasets containing user interactions, product views, cart additions, purchase records, demographics, and additional behavioral attributes to develop predictive models.

**Outcome Measurement:**

The success of this project will be assessed by the accuracy of predictive models in determining purchase intentions, the effectiveness of personalized recommendations, improvements in conversion rates, and enhancements in user engagement metrics.

1. **IDEATION & PROPOSED SOLUTION**
   1. **Empathy Map Canvas**

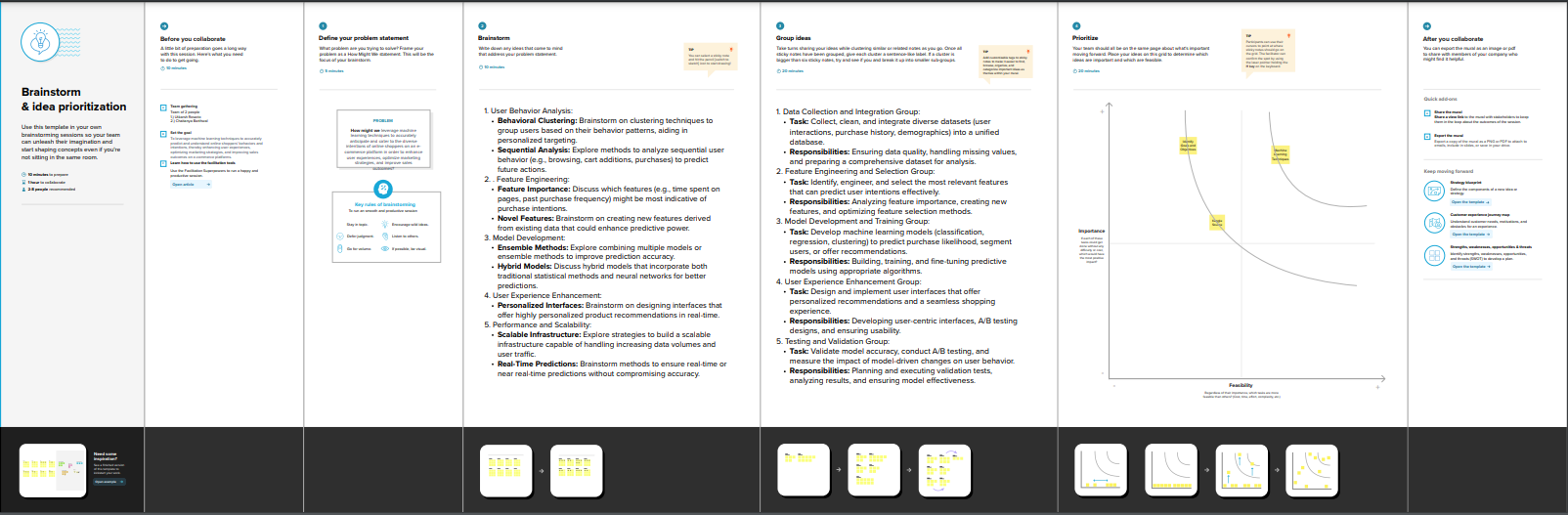
An empathy map is a simple, easy-to-digest visual that captures knowledge about a user’s behaviors and attitudes. It is a useful tool to help teams better understand their users. Creating an effective solution requires understanding the true problem and the person who is experiencing it. The exercise of creating the map helps participants consider things from the user’s perspective along with his or her goals and challenges.



* 1. **Ideation & Brainstorming**

Brainstorming provides a free and open environment that encourages everyone within a team to participate in the creative thinking process that leads to problem-solving. Prioritizing volume over value, out-of-the-box ideas are welcome and built upon, and all participants are

encouraged to collaborate, helping each other develop a rich amount of creative solutions. Use this template in your own brainstorming sessions so your team can unleash their imagination and start shaping concepts even if you're not sitting in the same room.

****

1. **Requirement Analysis**
   1. **Functional requirement**
2. **Models of Prediction:**

* **Purchase Likelihood:** Create a model to predict the likelihood of a user making a purchase based on their behavior and history.
* **Product Recommendations:** Create algorithms to generate personalized product recommendations for users.
* **User Segmentation:** Create models to segment users based on their behavior patterns for targeted marketing.

1. **Data Collection and Preprocessing:**

* **Data Integration:** Gather and integrate disparate data sources (user interactions, purchase history, demographics) into a unified dataset.
* **Data Cleaning:** Cleanse and preprocess the data to handle missing values, outliers, and ensure data consistency.

1. **Integration and Deployment:**

* **System Integration:** Integrate machine learning models into the backend of the e-commerce platform to make real-time predictions.
* **Scalability:** Ensure the system's scalability to handle increasing user data and requests.
  1. **Non-Functional Requirement**

1. **Performance:**

* **Speed:** Models should provide predictions in real-time or near real-time to avoid delays in user interactions.
* **Scalability:** The system should be able to handle increasing data volumes and user traffic without compromising performance.

1. **Accuracy and Reliability:**

* **Model Accuracy:** To ensure effective personalization and targeted marketing, models should predict user intentions with high accuracy.
* **Reliability:** During peak usage periods, the system should maintain consistent performance and reliability.

1. **Security and Privacy:**

* **Data Security:** Implement measures to ensure user data privacy and protection against potential security threats or breaches.
* **Ethical Use:** Follow ethical guidelines for collecting and utilizing user data.

1. **Project Design**
   1. **Data Flow Diagrams & User Stories;**

A Data Flow Diagram (DFD) is a traditional visual representation of the information flows within a system. A neat and clear DFD can depict the right amount of the system requirement graphically. It shows how data enters and leaves the system, what changes the information, and where data is stored.

**5.2 Solution Architecture**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **User Type** | **Functional**  **Requirement**  **(Epic)** | **User Story Number** | **User Story / Task** | **Acceptance criteria** | **Priority** | **Release** |
| Customer  (Mobile user) | Registration | USN-1 | As a user, I can say I am getting accurate result. | I can access my dashboard | High | Sprint-1 |
|  |  | USN-2 | As a user, I am reliable and it safes my life every time. | I can access my dashboard | High | Sprint-1 |
|  |  | USN-3 |  | I can access my dashboard | Low | Sprint-2 |
|  |  | USN-4 | As a user, I can register for the application through Gmail. | I can access my dashboard | Medium | Sprint-1 |
|  | Login | USN-5 | As a user, I can log into the application by entering email & password | I can access my dashboard | High | Sprint-1 |
|  | Dashboard |  |  |  |  |  |
| Customer (Web user) |  |  |  |  |  |  |
| Customer Care Executive |  |  |  |  |  |  |
| Administrator |  |  |  |  |  |  |

1. **Project Planning and Scheduling**
   1. **Technical Architecture**
   2. **Sprint Planning & Estimation**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Sprint** | **Functional**  **Requirement (Epic)** | **User Story**  **Number** | **User Story / Task** | **Story Points** | **Priority** | **Team**  **Members** |
| Sprint-1 | Registration | USN-1 | As a user, I can say I am getting accurate result. | 2 | High | Utkarsh Rosario |
| Sprint-2 |  | USN-2 | I am reliable how it safes my life everytime. | 1 | High | Utkarsh Rosario |
| Sprint-3 |  | USN-3 | Now we can easily supply water which we don’t have any idea whether it is safe or not | 2 | Low | Chaitanya Barthwal |
| Sprint-4 | Login | USN-5 | Safe to drink water now | 1 | High | Chaitanya Barthwal |

* 1. **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 | 20 | 21 Days | 1 November 2023 | 21 October 2023 | 20 | 29 October 2023 |
| Sprint-2 | 20 | 6 Days | 21 October 2023 | 29 November 2023 | 20 | 2 November 2023 |
| Sprint-3 | 20 | 6 Days | 1 November 2023 | 9 November 2023 | 20 | 10 November 2023 |
| Sprint-4 | 20 | 6 Days | 11 November 2023 | 20 November 2023 | 20 | 22 November 2023 |

1. **Performance Testing**
   1. **Performance Metrics:**

**A black rectangular object with a white stripe

Description automatically generated**

1. **Results**
   1. **Output Screenshots:**

**A screenshot of a computer

Description automatically generatedA screenshot of a computer

Description automatically generated**

1. **Advantages and Disadvantages**

**Advantages:**

* **Personalized Recommendations:** By utilizing machine learning algorithms to examine user behavior, the shopping experience can be improved with personalized product recommendations.
* **Improved User Engagement:** Understanding consumer intents enables personalized interactions, which increase user engagement and loyalty.
* **Optimized Marketing Strategies:** By more precisely targeting customers according to their preferences, predictive models aid in the optimization of marketing efforts.
* **Improved Sales Conversion:** By presenting appropriate products at the correct time, retailers can increase sales by accurately forecasting buyer intents.
* **Behavioral Insights:** ML models can provide useful insights into user behavior, allowing firms to make more educated decisions.

**Disadvantages:**

* **Data Completeness and Bias:** Prediction accuracy can be affected by incomplete or biased data, which is a major dependency of machine learning models.
* **Interpretation of Complex Models:** It may be difficult to interpret advanced machine learning models, which makes it difficult to communicate predictions to users.
* **Privacy and Ethical considerations:** Using user data for prediction poses privacy and ethical considerations, necessitating careful treatment of sensitive information.
* **Overfitting of models:** Models may perform well on training data but struggle to generalize to new or unknown data, resulting in overfitting.
* **Continuous Model Maintenance:** As shopper behavior evolves, models require continuous monitoring and adaption, which can be time-consuming.

1. **Conclusion**

Online Shoppers Intentions by utilized machine learning to comprehend online shoppers' behavior, enhancing user experiences and optimizing marketing strategies. Key insights obtained from predictive models led to improved sales outcomes and personalized interactions, highlighting the potential of ML in understanding and catering to online shoppers' intentions effectively. Recommendations for further enhancements and future research were also outlined, showcasing the project's impact and paving the way for continued advancements in this domain.

1. **Future Scope**
2. **Appendix**

**Source Code:**