**High-Level Design (HLD)**

**Travel Purchase Package Prediction**

|  |  |
| --- | --- |
| Written By | Aadil Hussain Teeli |
| Version | 1.0 |
| Date | 8-12-2024 |

|  |  |  |  |
| --- | --- | --- | --- |
| **Version** | **Date** | **Author** | **Comments** |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

**Document Change Control Record**

Review

|  |  |  |  |
| --- | --- | --- | --- |
| **Version** | **Date** | **Reviewer** | **Comments** |
|  |  |  |  |

**Approval**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Version** | **Review date** | **Reviewed By** | **Approved By** | **Comments** |
|  |  |  |  |  |

**Abstract**

The "Travel Purchase Package Prediction" project addresses a classification problem at the intersection of travel and data science. Focused on predicting travel package purchases, the system utilizes historical data encompassing customer interactions and purchase behaviors. The primary goal is to classify customers into distinct categories based on their likelihood of purchasing specific travel packages.

Drawing insights from past purchase records and various customer attributes, the system employs a diverse range of machine learning algorithms. These algorithms discern intricate patterns and relationships within the data, contributing to the development of a robust classification model. The model's efficacy lies in its ability to accurately categorize potential buyers, providing businesses with valuable insights to tailor marketing strategies and optimize sales efforts.

In essence, this project revolutionizes customer engagement and revenue generation in the travel industry by empowering businesses to proactively address the diverse needs and preferences of their customer base. The outcome is a sophisticated classification model that not only enhances the precision of predictions but also serves as a strategic tool for businesses seeking to elevate their marketing initiatives and maximize sales in the dynamic landscape of the travel sector.

1. Introduction

**1.1 Why these High-Level Design Documents?**

The purpose of this High-Level Design(HLD) Documents is to add necessary details to the current project description to represent a suitable for coding. This document is also intended to help detect contradictions before coding. And can be used as a reference manual for how the modules interact at a high level.

The HLD will be :

* Present all of the design aspects and define them in detail.
* Describe the user interface being implemented.
* Describe the needed Python libraries for the coding.
* Describe the performance requirements.
* Include design features and the architecture of the project.
* List and describe the non-functional attributes like:
  + Security
  + Reliability
  + Maintainability
  + Portability
  + Reusability
  + Application Compatibility
  + Resource Utilization
  + Serviceability

**1.2 Scope**

The HLD documentation presents the structure of the system, such as the database architecture, application architecture(layers), application flow (Navigation), and technology architecture, The HLD uses non-technical and mildly-technical terms which should be understandable to the administrators of the system

**1.3 Definition**

|  |  |
| --- | --- |
| TERM | Description |
| DB | Database, the cloud platform where the data will be stored. Can be considered  cloud storage. |
| ML | Machine Learning |
| AWS | Amazon Web Services |
| API or APIs | Application Programming Interface can be considered a website link from there we can extract information. |

**2. General Description**

**2.1 Product Perspective**

The "Travel Purchase Package Prediction" project is a targeted predictive tool within the travel industry, seamlessly integrating into marketing and sales processes. Through the analysis of various features and historical data, the system forecasts customer decisions on travel package selection, ensuring a dynamic approach to customer engagement and revenue optimization. This aligns with data-driven decision-making trends, enhancing adaptability and innovation for businesses by predicting whether a customer is likely to opt for a travel package.

**2.2 Problem Statement**

The aim of this project is to build a smart system that can predict how popular travel packages will be in the future. By looking at details about customers and their past choices, the system will determine if people are likely to choose a specific travel package. The challenge is to make a reliable tool that helps travel companies decide on things like advertising, managing available packages, and allocating resources wisely.

**2.3 Proposed Solution**

Our proposed solution initiates with leveraging exploratory data analysis (EDA) to uncover significant relationships between various attributes and the likelihood of customers purchasing travel packages. Using advanced machine learning algorithms, including fine-tuning through hyperparameter adjustment, we aim to construct a predictive model capable of anticipating future sales demand. Clients will engage with a user-friendly web application, inputting relevant features, which will then undergo validation, preprocessing, and backend prediction generation. The deployment of this integrated system is designed to furnish clients with actionable insights, enabling them to refine marketing strategies, allocate resources efficiently, and maximize revenue generation amidst the dynamic landscape of the travel industry.

**2.4 Data Requirements**

The necessary data for the Travel Package Purchase Prediction project is readily available within the dashboard. The dataset contains essential information pertinent to customer behavior and travel package attributes. With 4888 rows, it includes columns such as 'CustomerID', 'ProdTaken', 'Age', 'CityTier', 'DurationOfPitch', 'NumberOfPersonVisiting', 'NumberOfFollowups', 'PreferredPropertyStar', etc. oThese attributes provide valuable insights into customer demographics, preferences, and interactions with travel packages.

**2.5 Tool Used**

Python programming language and framework such as Numpy,Pandas,Scikit-learn are used to build the whole model.

   

* VS code is used as IDE.
* For visualization of the plots, Matplotlib, Seaborn are used.
* AWS is used for deployement of the model.
* Front end development is done using HTML/CSS.
* Flask is used for backend development

**2.6 Constraints**

The System should be user-friendly, the user should get all proper messages while using the web app. He/she also should get a proper error message if he/she has done something wrong On the web-app page. All the errors and results should be delivered in the easiest possible way and all the buttons are going to insert on the webpage should be labeled properly, so the user did not get confused to use the system.

**2.7 Assumptions**

The project operates under the assumption that historical data is dependable and that machine learning models are effective in forecasting future demand for travel packages. It also presupposes a level of stability in external factors that impact purchase decisions and consistency in customer behavior over an extended period

**3. Design Details**

**3.1 Process Flow**

We will be using following process flow for this project. The process will be based on modular coding i.e. use of oops concepts to build the entire project from start to end.

Start

Data gathering

Data Cleaning

Handling Missing Data

Parameter tuning

Model building

Model saving

End

Feature Generation

Deployment

Export into csv

Push to GitHub

Flask setup

Encoding Categorical Data

New feature creation

**3.2 Deployment Process**



**3.3 Error Handling**

In case of any processing errors, the system should display user-friendly error messages that are easily understandable by anyone, regardless of technical expertise. These messages should provide meaningful information to help users identify and correct their mistakes when rerunning the process. It's essential to handle all potential errors effectively, and a comprehensive error logging system must be in place to track and manage errors within the application.

**4. Performance**

The prediction of Travel Package Purchases relies on machine-learning algorithms. We will train different ML algorithms to identify the most suitable one for predicting the target. The effectiveness of our system will hinge on the quality of the data provided to these algorithms. The overall performance will be determined by the chosen model, the web application, and the deployment server. A seamless operation of our program is contingent upon the proper functioning of all these components.

**4.1 Reusability**

During the project development phase, adherence to coding guidelines is crucial for both the code and modules. The entire project code is structured in a modular fashion to ensure maintainability. Our system is designed to be flexible, capable of functioning seamlessly from any location. It should gracefully handle improper input values from users, providing clear error messages for correction. Additionally, the system is built to be reusable, accommodating various input values it has been trained on.

**4.2 Application Compatibility**

The system is constructed using various libraries and the Python programming language. Each library serves a specific function and is expected to seamlessly integrate with our dynamic system. Flask will be employed to create web APIs, while HTML/CSS will be utilized for developing the web application. It is imperative that all components of the application function seamlessly to yield results without any disruptions.

**4.3 Resource Utilization**

Our application should utilize the given resource properly and it should use a minimal amount of internet to work and call the APIs on the Web page. Our system should not use much amount of computational resources hence it will make the application slow. Our application will be deployed cloud platform and it should utilize the resource given on the cloud and work properly.

**5. Deployment**

For the deployment process, we will using AWS cloud platforms for hosting our application. The cloud platform will run the system and it will give the flexibility to use our application globally.



**6. Conclusion**

The objective of the Travel Package Purchase Prediction project is to offer valuable insights to travel companies, aiding them in refining marketing strategies, improving customer targeting, and optimizing revenue generation. Through the analysis of past customer behavior and travel package characteristics, our goal is to build a strong machine learning model capable of precisely forecasting future demand for travel packages. This predictive system is designed to enable businesses to make well-informed decisions regarding product offerings and resource allocation, while also streamlining supply chain management processes. By leveraging historical data and advanced machine learning techniques, we aim to foster growth and efficiency in the travel industry, delivering tangible value to stakeholders.

**7. Reference**

1. Google image for collection the logos and images.
2. Sketch diagram for drawing the diagrams.