**High Level Design (HLD)**

**Optimum Product Weight Prediction**

Revision 1.0.1

Date: 17/11/2023

Description: Initial revision HLD

|  |  |  |  |
| --- | --- | --- | --- |
| **Date** | **Revision** | **Description** | **Author** |
| 10/12/2023 | 1.0.1 | Initial revision (1.0.1) | Robin Mitra |
|  |  |  |  |
|  |  |  |  |

**Document Version Control**

**Contents**

Document Version Control…………………………………………………………………………………………………………… 2

Abstract……………………………………………………………………………………………………………………………………….. 4

1 Introduction…………………………………………………………………………………………………………………….. 5

1.1 Why this High-Level Design Document?.................................................................. 5

1.2 Scope………………………………………………………………………………………………………………….. 5

1.3 Definitions………………………………………………………………………………………………………….. 5

2 General Description…………………………………………………………………………………………………………. 6

2.1 Product Perspective……………………………………………………………………………………………. 6

2.2 Problem statement…………………………………………………………………………………………….. 6

2.3 PROPOSED SOLUTION………………………………………………………………………………………… 6

2.4 FURTHER IMPROVEMENTS………………………………………………………………………………… 6

2.5 Technical Requirements……………………………………………………………………………………… 6

2.6 Data Requirements…………………………………………………………………………………………….. 6

2.7 Tools used………………………………………………………………………………………………………….. 7

2.7.1 Hardware Requirements……………………………………………………………………….. 7

2.8 Constraints………………………………………………………………………………………………………….. 7

2.9 Assumptions……………………………………………………………………………………………………….. 7

3 Design Details………………………………………………………………………………………………………………….. 8

3.1 Process Flow……………………………………………………………………………………………………….. 8

3.1.1 Model Training and Evaluation……………………………………………………………….. 8

3.1.2 Deployment Process………………………………………………………………………………. 9

3.2 Event log…………………………………………………………………………………………………………….. 9

3.3 Performance……………………………………………………………………………………………………….. 9

3.5 Reusability………………………………………………………………………………………………………….. 9

3.6 Application Compatibility……………………………………………………………………………………. 9

3.7 Resource Utilization……………………………………………………………………………………………. 9

3.8 Deployment……………………………………………………………………………………………………….. 9

4 Dashboards…………………………………………………………………………………………………………………….. 10

4.1 KPls (Key Performance Indicators)………………………………………………………………………. 10

5 Conclusion………………………………………………………………………………………………………………………. 10

**Abstract**

This study explores the optimal weight of products for shipment to warehouses, aiming to enhance supply chain efficiency. Considering factors like warehouse capacity, transportation constraints, and regional demand, the research employs predictive modeling and historical data analysis. The goal is to establish a framework for identifying the ideal weight range per shipment, balancing warehouse space utilization and transportation costs. Objectives include assessing the impact on storage capacity, analyzing transportation efficiency, and evaluating cost implications. Results can guide logistics managers in making informed decisions for efficient packaging and shipment planning, ultimately contributing to streamlined warehouse operations and cost-effectiveness in the supply chain.

**1** **Introduction**

**1.1** **Why this High-Level Design Document?**

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

The HLD will:

• Present all of the design aspects and define them in detail

• Describe the user interface being implemented

• Describe the hardware and software interfaces

• Describe the performance requirements

• Include design features and the architecture of the project

• List and describe the non-functional attributes like:

o Security

o Reliability

o Maintainability

o Portability

o Reusability

o Application compatibility

o Resource utilization

o 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 to mildly-technical terms which should be understandable to the administrators of the system.

**1.3 Definitions**

|  |  |
| --- | --- |
| Term | Description |
| Database | Collection of all the information monitored by this system |
| IDE | Integrated Development Environment |
| EDA | Exploratory Data Analysis |

**2 General Description**

**2.1 Product Perspective**

The Product Weight Calculator For Warehouses system is a machine learning-based calculation model which will help us to get the correct amount of weight to be shipped .

**2.2 Problem statement**

This exercise aims to build a model, using historical data that will determine the optimum weight of the product to be shipped each time to the warehouse.

**2.3 Proposed Solutions**

The solution proposed here is that first we get the basic details from the consumer like “Location type”, ”Zone”, “Number of times refilling”, etc. Then in second case we put these details in the Machine Learning Model. Then in last stage that is in third case it calculates the approx. weight of product to be shipped, which help them to manage the supply chain for the company.

**2.4 Further Improvements**

In this model we use sufficient amount of variables i.e. “Location type”, ”Zone”, “Number of times refilling”,”Owner Type” etc. No further improvement required as such.

**2.5 Technical Requirements**

I used python version 3.8 with some important libraries to develop a machine learning model, which accurately predicts the weight of the product to be shipped based on its details.

**2.6 Data Requirements**

Data requirement completely depend on our problem statement.

• We need past data that is balanced and must have at least 10000 or more products details.

• We require past weights products which are shipped before for each warehouse

Data dictionary as follows:

|  |  |  |
| --- | --- | --- |
| **Name** | **Data Type** | **Description** |
| **Location Type** | String | Input variable |
| **Warehouse capacity size** | String | Input variable |
| **Zone** | String | Input variable |
| **Warehouse Regional Zone** | String | Input variable |
| **Number Of Time Refill** | Integer | Input variable |
| **Transport Issue Type** | String | Input variable |
| **Competitors In Market** | Integer | Input variable |
| **Warehouse Owner Type** | String | Input variable |
| **Supply Of Electricity** | String | Input variable |
| **Distance From Hub To Warahouse** | Decimal | Input variable |
| **Number Of workers** | Integer | Input variable |
| **Storage Issue Reported** | Integer | Input variable |
| **Tempreture Regulator** | String | Input variable |
| **Type Of Government Certificate Issued** | String | Input variable |
| **Warehouse Breakdown** | Integer | Input variable |
| **Government Checking** | Integer | Input variable |
| **Product Weight** | Interger | Output variable |

**2.7 Tools and Technologies used**

**2.7.1 Hardware Requirements**

• Jupyter notebook is used for EDA and experimentation with various ML algorithms with the help of pandas, numpy, matplotlib, seaborn, scikit learn, statsmodels and xgboost

libraries.

• VsCode is an IDE used for development and deployment of the solution with logging. Used python version 3.7 and libraries include logging, pandas, numpy, scikit learn, statsmodels, Xgboost, pickle, flask and html 5.

• GitHub is used as a version control system.

• Power BI is a powerful tool for better visualization

**2.8 Constraints**

Prediction system must be user friendly, errors free and users should not be required to know about any of the workings.

**2.9 Assumptions**

Assumptions for insurance premium prediction in machine learning include accurate and representative data, relevance of selected risk factors, independence of events, stationarity, homoscedasticity, linear relationships, normality of residuals, and minimal collinearity. Regular monitoring and model adaptation are essential for dynamic real-world scenarios.

**3 Design Details**

**3.1 Process Flow**

**3.1.1 Model Training and Evaluation**

**3.1.2 Deployment Process**

**3.2 Event log**

In this project, I used the “logging” library in both the development and deployment stages, which keeps logging the events at every step into the “.log” files. One of the advantages of event logging is, it makes debugging much easier, like we can directly go to that specific line of code, having errors.

**3.3 Performance**

The ML based Insurance Premium prediction application is used for predicting the Insurance premium based on its age and other factors. So, it should be as accurate as possible, so that it will not mislead the user. Also, the model retraining is very important to keep it relevant if the new factors are added in future or to improve the performance.

**3.4 Reusability**

The code written and the components used have an ability to be reused without any problem.

**3.5 Application compatibility**

The different components or modules of this project use python version 3.8 as their interface between them. Each component has its own task to perform, and it is the job of the python version to ensure proper transfer of the information.

**3.6 Resource utilization**

In this project, any task may likely to use all the processing power available in the system, until it is accomplished.

**3.7 Deployment**

**3.8 User Interface**

Designed user interface using HTML. It looks as per the below image**.**

**4 Dashboards**

**4.1 KPls (Key Performance Indicators)**

Key performance indicators (KPIs) for insurance premium prediction models include Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Mean Percentage Error (MPE), R-squared, accuracy, precision, recall, F1 score, AUC-ROC, and feature importance. These metrics collectively assess model accuracy, efficiency, and ability to predict premiums accurately.

**5 Conclusion**

**Optimum product weight calculator is used to predict the optimum weight to be shipped from a warehouse so that the company will minimize the loss for during supply of products in future.**