High Level Design (HLD)

Shipping Price Prediction

Revision Number: 2.0

Last date of revision: 04/03/2023

Document Version Control

|  |  |  |  |
| --- | --- | --- | --- |
| Date Issued | Version | Description | Author |
| 04/03/2023 | 1 | Initial HLD | Lovely Patra |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

Contents

Document Version Control...

Abstract.........................................................................

1. Introduction.5

1 .1 Why this High-Level Design Document? ..... . . . . . . . . . .5

1 .2 Scope. . . . . .5

1.3 Definitions ... .5

1. 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.6

2.6 Data Requirements .

2.7 Tools used........................................................................

 2.7. 1 Hardware Requirements........................................................................

2.7.2 ROS(Robotic Operating System) .. .9

2.8 Constraints

2.9 Assumptions................................................................... .. ..

1. Design Details ..... ..... ..... ................................ . . . . . . 1 0
   1. Process Flow....................................................................................................... 1 0
      1. Model Training and Evaluation . . . . . . 1 0
      2. Deployment Process 1 1
   2. Event log 1 1
   3. Error Handling . . . . . . . 1 1
   4. Performance........................................................................................................12
   5. Reusability................................................. . . . . . 1 2
   6. Application Compatibility. . . . . . 1 2
   7. Resource Utilization. . . . . . 1 2
   8. Deployment ..... ..... ........................ .. . . . . . . . 1 2
2. Dashboards. . . . . . 1 3
   1. KPls (Key Performance Indicators).............................................. . . . . . 1 3
3. Conclusion .. . . . 1 4

Abstract

Not having an exact cost standard can present a problem for setting the shipping costs on a freight brokerage platform. Transport brokers who use their high market position to charge excessive commissions can also make it difficult to set rates. In addition, due to the absence of a quantified fare policy, fares are undervalued relative to the labor input. Therefore, vehicle owners are working for less pay than their efforts. This study derives the main variables that influence the setting of the shipping costs and presents the recommended shipping cost given by a price prediction model using machine learning methods. The cost prediction model was built using five algorithms: KNN Regressor, Decision Tree Regressor , XGB Regressor, Random Forest Regressor, and AdaBoost Regressor. R-squared was used as the performance evaluation index. In view of the results of this study, Random Forest Regressor was chosen as the model with the greatest explanatory power and the fastest processing. Furthermore, the range of the predicted shipping costs was determined considering realistic usage patterns. The confidence interval was used as the method of calculation for the range of the predicted shipping costs, and, for this purpose, the dataset was classified using the K-fold cross-validation method. This paper could be used to set the shipping costs on freight brokerage platforms and to improve utilization rates.

# 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:

 Security

 Reliability o Maintainability

 Portability

 Reusability

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

# 2 General Description

### 2.1 Problem statement

### The market for supply chain analytics is expected to develop at a CAGR of 17.3 percent

### from 2019 to 2024, more than doubling in size. This data demonstrates how supply

### chain organizations are understanding the advantages of being able to predict what will

### happen in the future with a decent degree of certainty. Supply chain leaders may use

### this data to address supply chain difficulties, cut costs, and enhance service levels all at

### the same time.

### The main goal is to predict the supply chain shipping pricing based on the available

### factors in the dataset.

### 2.3 PROPOSED SOLUTION

The solution proposed here is a machine learning-based shipping cost prediction method for a domestic freight transportation environment using data from a freight brokerage platform. It also shows that predictive models can set the shipping costs appropriately, and it compares the predictive power to present the best predictive model.

##### 2.6 Data Requirements

Used transportation-related data for 6 months from the freight brokerage platform. To identify the major factors, new factors were added, and various preprocessing methods were applied. Correlational analysis and a step selection method were used to derive the major factors. After that, a fare prediction model was developed using the derived factors with a machine learning algorithm. The machine learning algorithms used were KNN Regressor, Decision Tree Regressor , XGB Regressor, Random Forest Regressor, and AdaBoost Regressor. Random Forest Regressor is a model that reduces the learning time compared to other model.

#### 2.7 Tools used

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

* For visualization of the plots, Matplotlib, Seaborn and Plotly are used.
* AWS is used for deployment of the model.
* MySQL/MongoDB is used to retrieve, insert, delete, and update the database.
* GitHub is used as version control system.

##### 2.8 Constraints

The shipment pricing prediction model must be user friendly, as automated as possible and users should not be required to know any of the workings.

##### 2.9 Assumptions

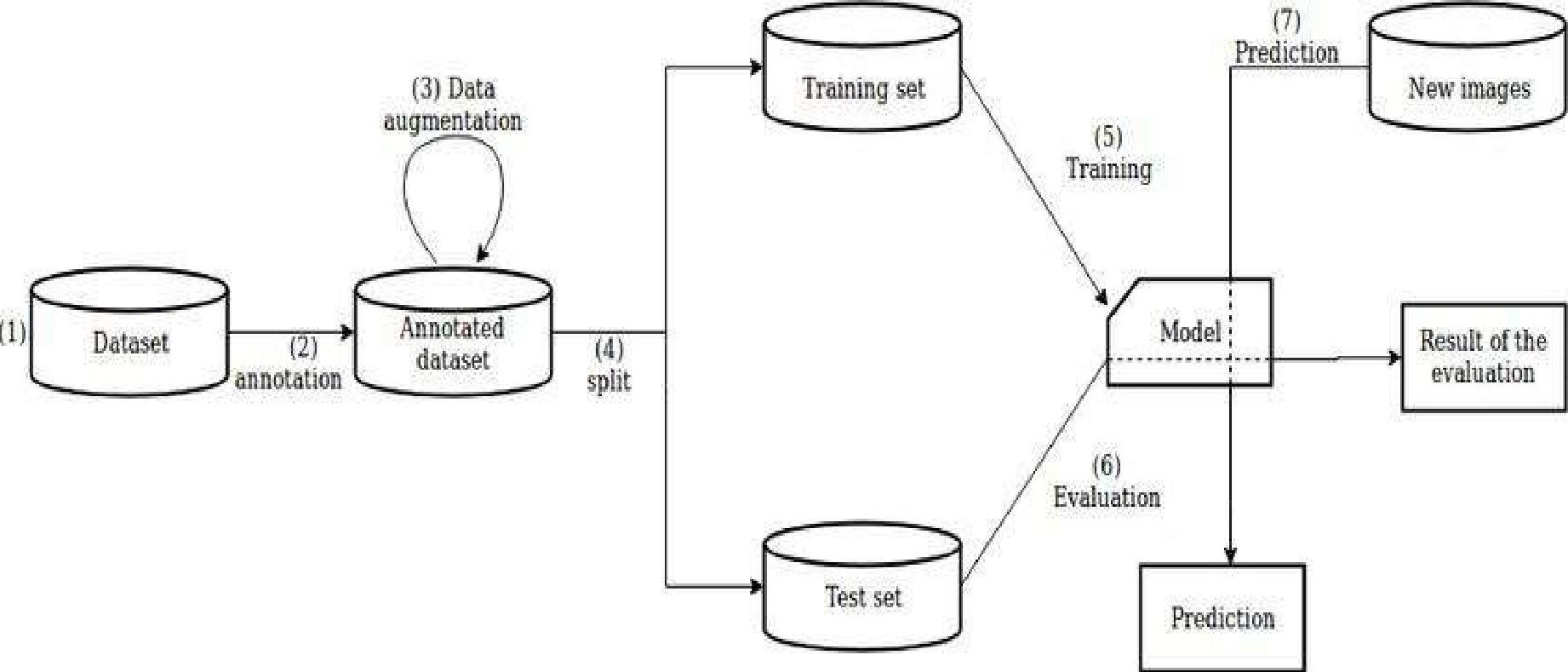
The main objective of the project is to implement the use cases as previously mentioned (2.2 Problem Statement) for new dataset. Machine Learning based price detection model is used for detecting the above-mentioned use cases based on the input data. It is also assumed that all aspects of this project have the ability to work together in the way designer is expecting.

1. Design Details

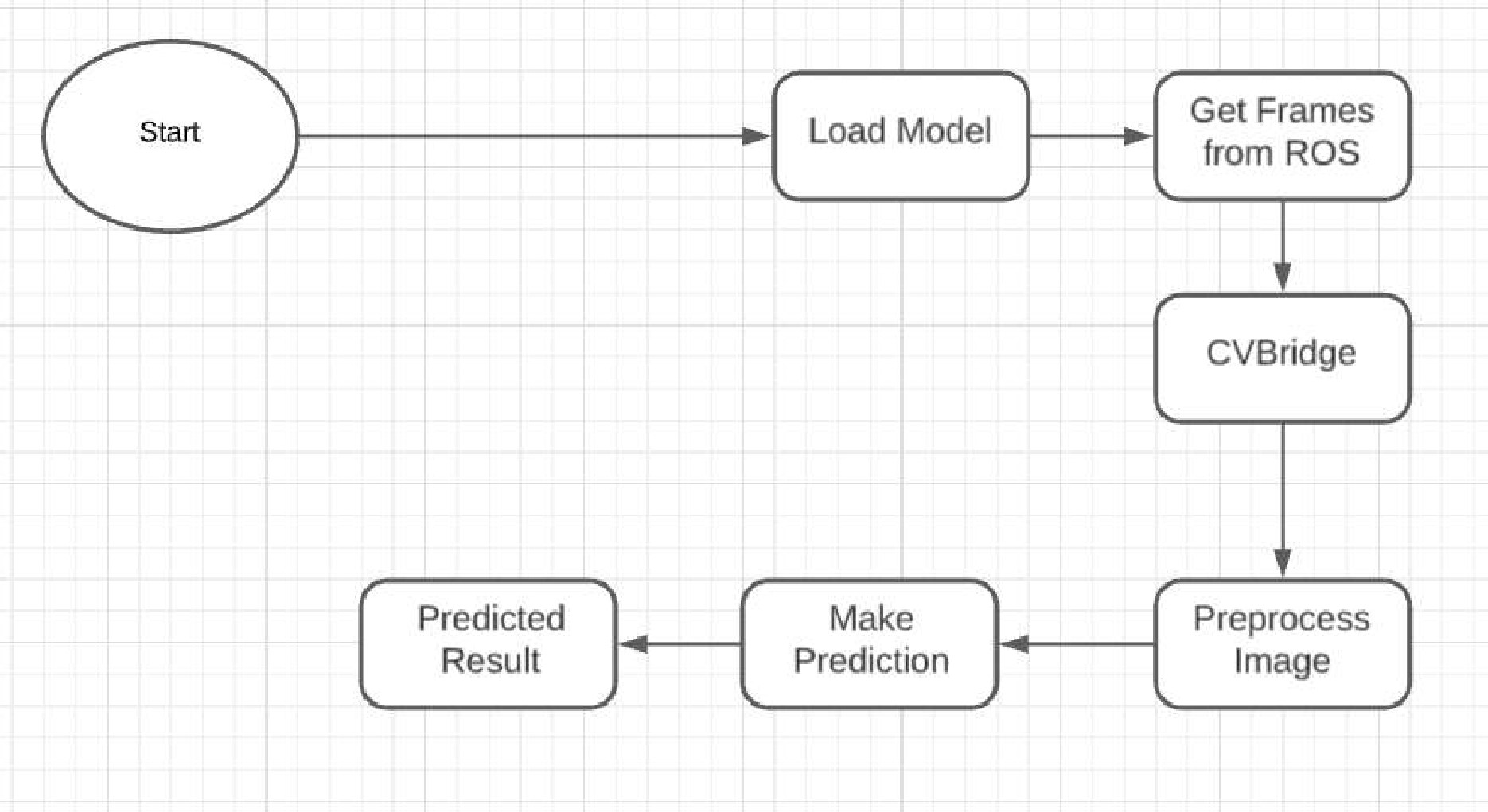
3.1 Process Flow

For identifying the different types of anomalies, we will use a deep learning base model. Below is the process flow diagram is as shown below.

##### 3.1.1 Model Training and Evaluation



##### 3.1.2 Deployment Process



#### 3.2 Event log

The system should log every event so that the user will know what process is running internally.

Initial Step-By-Step Description:

1. The System identifies at what step logging required
2. The System should be able to log each and every system flow.
3. Developer can choose logging method. You can choose database logging/ File logging as well.
4. System should not hang even after using so many loggings. Logging just because we can easily debug issues so logging is mandatory to do.

#### 3.3 Error Handling

Should errors be encountered, an explanation will be displayed as to what went wrong? An error will be defined as anything that falls outside the normal and intended usage.

#### 4.1 Reusability

The code written and the components used should have the ability to be reused with no problems.

#### 4.2 Application Compatibility

The different components for this project will be using Python as an interface between them. Each component will have its own task to perform, and it is the job of the Python to ensure proper transfer of information.

#### 4.3 Resource Utilization

When any task is performed, it will likely use all the processing power available until that function is finished

4.4 Deployment

Microsoft

Azurewebservicesruamazon

Google Cloud

Conclusion

To solve the problem of fare setting on a freight transportation brokerage platform, where there is no standardized shipping cost, the main factors that affect the shipping cost setting were derived in this study, and a price prediction model was built using machine learning. Factors that affect the shipping cost were selected using correlational analysis and the stepwise method from a total of 73 factors, including factors that were obtained from the freight brokerage process and environmental factors such as precipitation. The results of the analysis showed that the KNN Regressor, Decision Tree Regressor , XGB Regressor, Random Forest Regressor, and AdaBoost Regressor models, which are machine learning models, performed better than the linear regression, which is a traditional analysis method. The model that showed the best predictive power among the models used was the Random Forest Regressor model.

References

1. https://www.researchgate.net/publication/366972748\_Prediction\_of\_Shipping\_Cost\_on\_Freight\_Brokerage\_Platform\_Using\_Machine\_Learning