High-Level Design (HLD)

**BACK-ORDER PREDICTION**

**Technologies** Machine Learning

**Domain** E-commerce

Document Version Control

|  |  |  |  |
| --- | --- | --- | --- |
| Date Issued | Version | Description | Author |
| 13/04/2023 | 1 | Initial HLD - 1 .0 | SAMAR SINGH |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

**Contents**

Documents Version Control…………………………………………………………………………2

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

1. Introduction ………………………………………………………………………………………….5
   1. Why this High-Level Design Document? ………………………………………….5
   2. Scope ………………………………………………………………………………………………5
   3. Definition …………………………………………………………………………………….….5
2. General Description ………………………………………………………………………………6
   1. Product Perspective ………………………………………………………………………..6
   2. Problem Statement …………………………………………………………………………6
   3. Proposed Solution …………………………………………………………………………..6
   4. Further Improvements ……………………………………………………………………6
   5. Tools Used ………………………………………………………………………………………6
3. Design Details………………………………………………………………………………………..8
   1. Process Flow ……………………………………………………………………………………8
   2. DataBase …………………………………………………………………………………………8
   3. Deployment …………………………………………………………………………………….8

Abstract

This paper presents a study on backorder prediction using a real-world dataset from a manufacturing company. The dataset includes various features related to customer orders, inventory levels, and production capacity, as well as a binary target variable indicating the presence or absence of backorders. The proposed approach involves the use of machine learning algorithms for building a predictive model based on historical data. The model is trained and evaluated using various performance metrics, and its potential for proactive inventory management and production planning is demonstrated. The study contributes to the field of operations management by demonstrating the potential of data-driven approaches in improving supply chain efficiency and customer satisfaction.

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

### 1.3 Definitions

|  |  |
| --- | --- |
| **TERM** | **DESCRIPTION** |
| Database | Collection of the Information |
| IDE | Integrated Development Environment |
| HLD | High Level Document |

# 2 General Description

## 2.1 Product Perspective

The Back-Order Prediction Analysis is a Machine Learning model designed to predict backorder. This analysis tries to minimize the effects by predicting the products which are likely to go on a back order before its occurrence, so that businesses can take necessary actions at an earlier point in time, thus improving overall efficiency.

### 2.2 Problem statement

Backorders are unavoidable, but by anticipating which things will be backordered, planning can be streamlined at several levels, preventing unexpected strain on production, logistics, and transportation. ERP systems generate a lot of data (mainly structured) and also contain a lot of historical data; if this data can be properly utilized, a predictive model to forecast backorders and plan accordingly can be constructed. Based on past data from inventories, supply chain, and sales, classify the products as going into backorder (Yes or No)

### 2.3 Proposed Solution

Using all the standard techniques used in the life cycle of a Data Science project starting from Data Exploration, Data Cleaning, Feature Engineering, Model Selection, Model Building and Model Testing, and Back-Order Prediction analysis model will be developed, evaluated, and deployed to provide valuable insights.

2.4 FURTHER IMPROVEMENTS

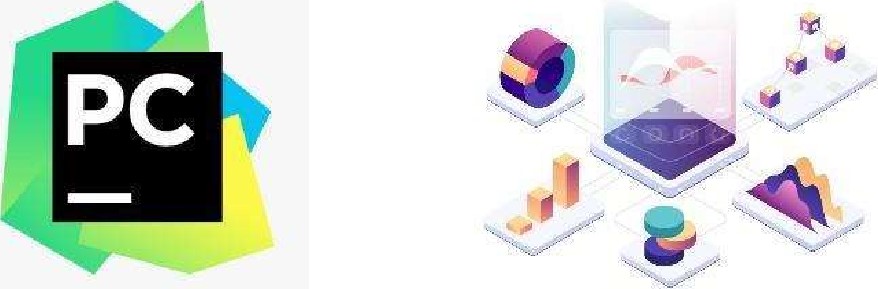
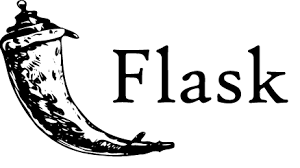
The Back-Order analysis model is a powerful tool that provides businesses with valuable insights. This model is designed to deliver the most accurate and effective Back-Order Prediction possible, enabling businesses to make informed decisions about marketing, product development, and customer engagement.

This can be further improved by training more data in the model.

#### 2.7 Tools used

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

* PyCharm is used as IDE.
* For visualization of the plots, Matplotlib, Seaborn and Plotly are used.
* AWS is used for deployment of the model.
* DataStax as a DataBase to store the Data.
* Front end development is done using HTML/CSS

 GitHub is used as version control system

#### 3. Design Details

#### 3.1Process Flow

**Data**

**Pre-processing / Feature Selection**

**Model Deployment**

**Model Training**

**Model Selection**

**Model Tuner**

**Data Ingestion**

**3**.**2 Database:**

DataStax is used as Database. It is NoSql Database known as Apache Cassandra Database.



**3.3 Deployment:**

For Deployment Aws is used. After storing the Code in Github Repository with the help of Elastic BeanStalk and Codepipeline Project in Deployed.

