1. Choose a project from an internship portal and try to write a HLD and LLD based on the sample given in your portal for a respective project .

**HLD – Mushroom Classification**

Mushroom Classification

## High Level Design Document

Revision Number: 0.0

Last date of revision: 26/06/2023

High Level Design Document 1

[Mushroom Classification](#_TOC_250000)

Contents

[Abstract 4](#_bookmark0)

1. [Introduction 5](#_bookmark1)
   1. [Why this High Level Design Document 5](#_bookmark2)
   2. [Scope 5](#_bookmark3)
2. [General Description 6](#_bookmark4)
   1. [Product Perspective 6](#_bookmark5)
   2. [Problem Statement 6](#_bookmark6)
   3. [Proposed Solution 6](#_bookmark7)
   4. [Technical Requirements 6](#_bookmark8)
   5. [Data Requirements 6](#_bookmark9)
   6. [Tools Used 6](#_bookmark10)
      1. [Hardware Requirements 7](#_bookmark11)
   7. [Constraints 7](#_bookmark12)
3. [Design Details 8](#_bookmark13)
   1. [Process Flow 8](#_bookmark14)
      1. [Model Training and Evaluation 8](#_bookmark15)
      2. [Deployment Process 9](#_bookmark16)
4. [Performance 10](#_bookmark17)
   1. [Reusability 10](#_bookmark18)
   2. [Application Compatibility 10](#_bookmark19)
   3. [Resource Utilization 10](#_bookmark20)
   4. [Deployment 10](#_bookmark21)

High Level Design Document 2

Mushroom Classification

# Document Version Control

|  |  |  |  |
| --- | --- | --- | --- |
| **Date Issued** | **Version** | **Description** | **Author** |
| 26/06/2023 | 1.0 | Initial HLD – V1.0 | Md Imran |
|  |  |  |  |

High Level Design Document 3

Mushroom Classification

# Abstract

Mushrooms are widely known for their great taste and amazing health benefits. Packed with a ton of essential vitamins and minerals, they make for an excellent addition to your diet, adding flavor to many different recipes.

Mushrooms are a low-calorie food that packs a nutritional punch. Loaded with many health-boosting

vitamins, minerals, and antioxidants, they’ve long been recognized as an important part of any diet.

In a lot of countries, mushrooms are abundant and people collect wild species for their unique taste and health benefits. Mushroom hunting(also known as shrooming) is a free activity for people who like exploring but not all of them are edible. Some species are poisonous as they contain different types of naturally occurring toxicants, and it can be risky even for people skilled with identifying them.

Hence, we try to identify which mushrooms are edible by creating a classification model and using the dataset(Mushroom Classification) that was originally contributed to the UCI Machine Learning repository

High Level Design Document 4

Mushroom Classification

# Introduction

### 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

### Scope

The HLD documentation presents the structure of the system, such as 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.

High Level Design Document 5

Mushroom Classification

# General Description

### Product Perspective

This application is useful to classify whether the mushroom is edible or poisonous, which will help to prevent humans from dangerous

### Problem Statement

To create a machine learning model to classify if the mushroom is edible or not.

### Proposed Solution

A machine learning model is to be implemented to perform the mushroom classification. It will train the model and get the input from UI it is required to predict the category of the mushroom, which is supposed to be classified. The input will be processed and given to the model to predict the output, then the output will be displayed to the web page, where they started the request.

### Technical Requirements

This application is based on python language, so python should be installed and it should be properly configured with the PATH.

* The application should have the access to use the python environment
* Necessary packages listed in the requirements.txt should be made available for this application

### Data Requirements

This application uses very less data when compared to other applications which are in the market now. It only carries 22 categorical data, which only requires very less data.

### Tools Used

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

* + VSCode is used as Code Editor
  + Google colab notebook is used for data analysis
  + For visualization of the plots Matplotlib and Seaborn were used.
  + Python Flask library is used for backend development.

High Level Design Document 6

Mushroom Classification

* + GitHub is used as a version control system.
  + AWS is used for deployment of the model.



* + 1. Hardware Requirements

PC or Laptop or any remote server which can run the application with necessary software which are mentioned in the technical requirements.

### Constraints

This classification system must be user friendly, and the user should not be required to know any of the workings of ML models.

High Level Design Document 7

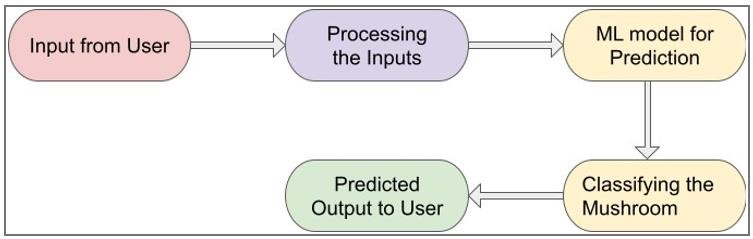
Mushroom Classification

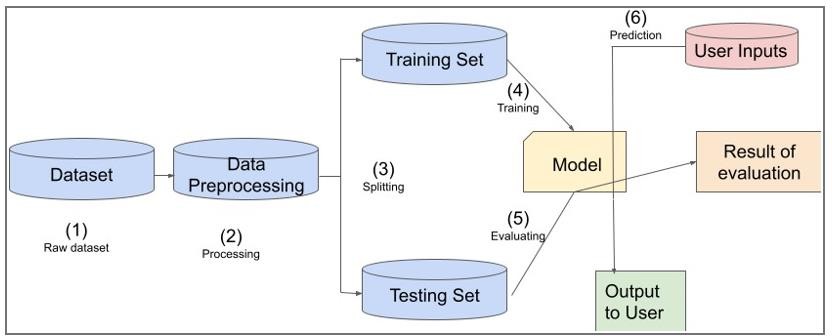
# Design Details

### Process Flow

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

Proposed methodology

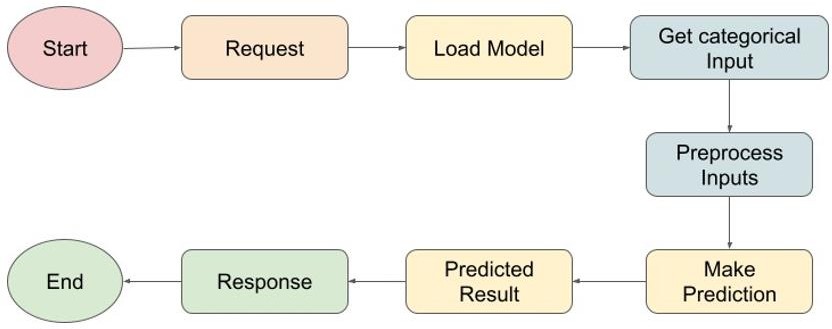


* + 1. Model Training and Evaluation

High Level Design Document 8

Mushroom Classification

* + 1. Deployment Process



### Event log

The system should identify where logging is required. It should maintain logs for different steps of code for reference during program debugging. Logging is to be maintained in files.

System should not hang even after using too many loggings.

### Error Handling

Exceptions (Anything that happens outside the normal and intended usage) should be handled for every part of code. They should be recorded in log file for reference.

High Level Design Document 9

Mushroom Classification

### Performance

This classifier is used for classifying whether the mushroom is edible or poisonous when providing the required inputs. The model should be performing well with good accuracy and recall.

### Reusability

The code written to check different models based on training is in a modular form like Object. So, it can be reused wherever and whenever needed.

### Application Compatibility

All the requirements are listed in the requirements.txt for packages. And the software required for running the packages is Python 3.X >= Python 3.6

### Resource Utilization

When a request is made, then it will utilize only the main thread. So, it will utilize less resources.

### Deployment

This project is to be deployed on AWS using Code pipeline.



High Level Design Document 10

Mushroom Classification

Conclusion

The model should be able to identify if a mushroom is edible or not based on its features. The model should be able to accept input from users and data should be validated.

Also, the model should have minimum ‘Recall’ as classifying poisonous mushrooms as edible

can be dangerous.

This will be helpful for users to identify which mushrooms should they consume.

**LLD – Mushroom Classification**

Mushroom Classification

Mushroom Classification

## Low Level Design Document

Revision Number: 0.0

Last date of revision: 26/06/2023

Low Level Design Document 1

[Mushroom Classification](#_TOC_250000)

Contents

1. [Introduction 4](#_bookmark0)
   1. [What is Low Level Design Document 4](#_bookmark1)
   2. [Scope 4](#_bookmark2)
2. [Architecture 5](#_bookmark3)
3. [Architecture Description 6](#_bookmark4)
   1. [Data Description 6](#_bookmark5)
   2. [Data Ingestion 6](#_bookmark6)
   3. [Data Transformation 6](#_bookmark7)
   4. [Model Evaluation 6](#_bookmark8)
   5. [Deployment 7](#_bookmark9)
   6. [Data from User 7](#_bookmark10)
   7. [Prediction 8](#_bookmark11)

[Unit Test Cases 9](#_bookmark12)

Low Level Design Document 2

Mushroom Classification

# Document Version Control

|  |  |  |  |
| --- | --- | --- | --- |
| **Date Issued** | **Version** | **Description** | **Author** |
| 19/03/2023 | 1.0 | Initial HLD – V1.0 | Md Imran |
|  |  |  |  |

Low Level Design Document 3

Mushroom Classification

# Introduction

### What is Low Level Design Document

The goal of LLD or a low-level design document (LLDD) is to give the internal logical design of the actual program code for the Mushroom Classifier. LLD describes the class diagrams with the methods and relations between classes and program specs. It describes the modules so that the programmer can directly code the program from the document.

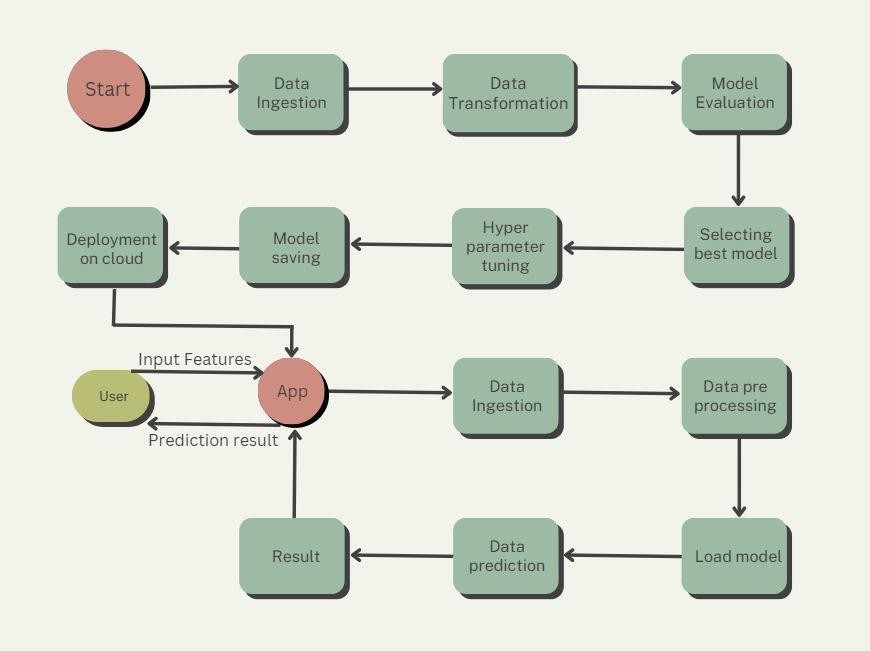
### Scope

Low-level design (LLD) is a component-level design process that follows a step-by-step refinement process. This process can be used for designing data structures, required software architecture, source code and ultimately, performance algorithms. Overall, the data organization may be defined during requirement analysis and then refined during data design work

Low Level Design Document 4

Mushroom Classification

# Architecture



Low Level Design Document 5

Mushroom Classification

# Architecture Description

### Data Description

Data is taken from the Kaggle Competition - Mushroom Classification.

This dataset includes descriptions of hypothetical samples corresponding to 23 species of gilled mushrooms in the Agaricus and Lepiota Family Mushroom drawn from the Audubon Society Field Guide to North American Mushrooms (1981). Each species is identified as definitely edible, definitely poisonous, or of unknown edibility and not recommended. This latter class was combined with the poisonous one.

### Data Ingestion

Data is downloaded in csv format from Kaggle and is split in training and testing set and again stored in different csv.

### Data Transformation

Data type of columns – All features of this data is in object format. They will be converted to numerical format through one hot encoding.

Null values in columns – This data does not contain null values but there are some missing values which are to be filled through imputation.

Name of columns - The name of the columns should be as per the columns name of the dataset. Target column will be separated for training and testing data and preprocessing will be done on it. Preprocessing object will be stored in pickle format for further use.

### Model Evaluation

These models will be tested in the data using

1. XGBoostClassifier
2. GradientBoostClassifier
3. Random Forest

Best model will be selected according to accuracy and will be used train and predict data.

Low Level Design Document 6

Mushroom Classification

### Deployment

Model will be deployed on AWS server.

### Data from User

Inputs will be collected from users using dropdown functionality in the Web Application Interface, which contains the hypothetical data.

Input features are,

* Cap-Shape
* Cap-Surface
* Cap-color
* Bruises
* Odor
* Gill-Attachment
* Gill-Spacing
* Gill-Size
* Gill-Color
* Stalk-Shape
* Stalk-root
* Stalk-Surface-Above-Ring
* Stalk-Surface-Below-Ring
* Stalk-Color-Above-Ring
* Stalk-Color-Below-Ring
* Veil-Type
* Veil-Color
* Ring-Number
* Ring-Type
* Spore-Print-Color
* Population
* Habit

Low Level Design Document 7

Mushroom Classification

### Prediction

Data will be entered by user through a HTML/CSS front end application and Flask at back end.

Data will then be pre-processed through transformation pickle and it will then be used to make predictions using pickle file for model.

The prediction results will then be showed to user on user interface.

Low Level Design Document 8

Mushroom Classification

# Unit Test Cases

|  |  |  |
| --- | --- | --- |
| **Test Case Description** | **Expectation** | **Outcome** |
| Verify whether the Application URL is accessible to the user | 1. Application URL should be defined | Application URL is accessible to the user. |
| Verify whether the Application loads completely for the user when the URL is accessed | 1. Application URL is accessible 2. Application is deployed | The Application is correctly deployed. |
| Verify whether user is able to see input fields on Web Page | 1. Application is accessible 2. Application is deployed | User is able to see input fields on Web Page. |
| Verify whether user is able to edit all input fields | 1. Application is accessible 2. Application is deployed | User is able to edit all input fields. |
| Verify whether user gets Submit button to submit the inputs | 1. Application is accessible 2. Application is deployed | User gets Submit button to submit the inputs. |
| Verify whether user is presented with recommended results on clicking submit | 1. Application is accessible 2. Application is deployed | User is presented with recommended results on clicking submit button. |
| Verify whether the recommended results are in accordance to the selections user made | 1. Application is accessible 2. Application is deployed | The recommended results are in accordance to the selections user made. |
| Verify whether Predicted Output modify as per the user inputs for the different mushroom data | 1. Application is accessible 2. Application is deployed | Predicted Output modifies as per the user inputs for the different mushroom data |