iNeur🔵n

Low Level Design

Mushroom Classification

Written by	Chetan Ullagaddi
Document version	1.0
Last Revision Date	27/09/2021

Document Control

Date	Version	Author
27/09/2021	Version-1.0	Chetan Ullagaddi

Contents   Page no.
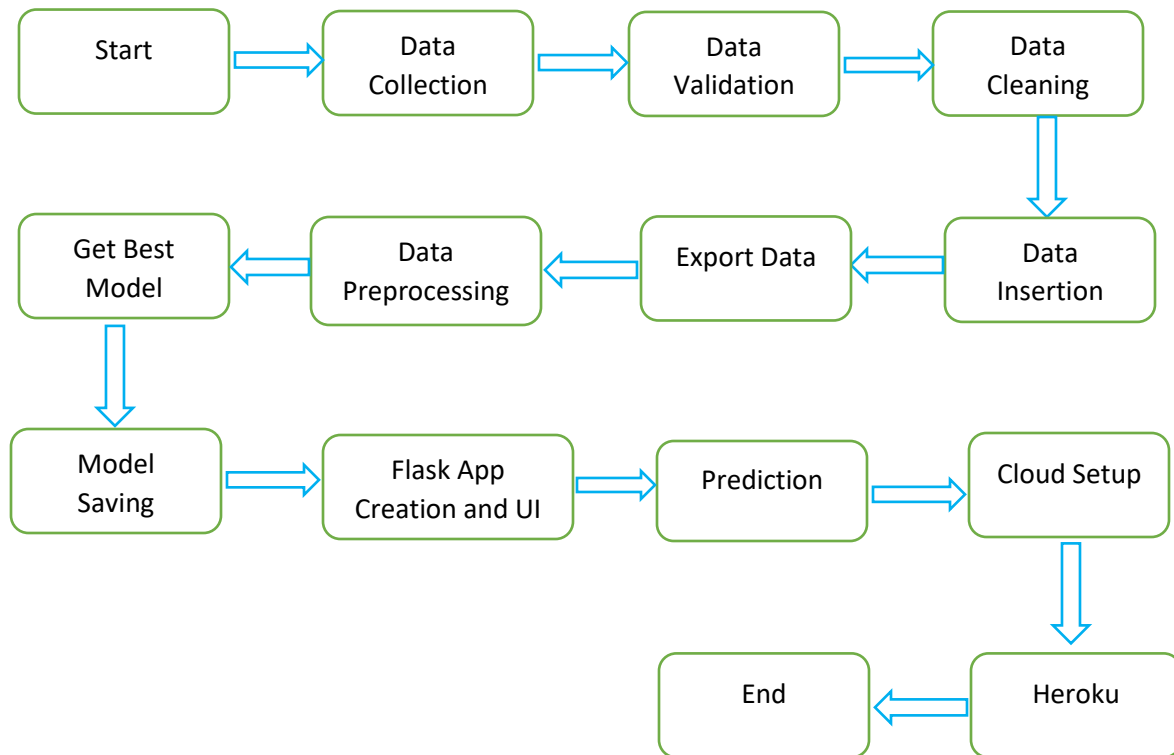
1. Introduction

1.1 What is Low Level Design Document?

The goal of LLD or a low-level design document (LLD) is to give the internal logical design of the actual program code. Low-level design is created based on the high-level design. 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.

1.2 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. Post-build, each component is specified in detail.

2 Architecture

```
Start → Data Collection → Data Validation → Data Cleaning
                                                  ↓
Get Best Model ← Data Preprocessing ← Export Data ← Data Insertion
      ↓
Model Saving → Flask App Creation and UI → Prediction → Cloud Setup
                                                             ↓
                        End ← Heroku
```

3   Architecture Description

3.1 Data Description
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.

3.2 Import Data
Data Import - The data in a stored database is imported as a CSV file to be used for Data Pre-processing and Model Training.

3.3 Data Cleaning
There are no null values in the data and all the variables are categorical, some of the observations meaningless and they are converted into meaningful observation.
Example: "?" is converted into letter "m" (Missing)

3.4 Exploratory Data Analysis
EDA is done in such way that every independent variable in the dataset is visualized by multiple bar plot with respect to dependent variable having classes of poisonous and edible mushrooms.

3.5 Data Preprocessing
Data preprocessing steps are converting categorical variables into numerical variables using label encoding method and train and test split of the data etc.

3.6 Model Building
After Data preprocessing split the data train and test (Simple Random Sampling) and implemented different Classification Machine Learning Algorithm. Random Forest model gives the better accuracy.

3.7 Model Dump
After comparing all accuracies and finding the best model for the dataset I have created a model and dumped the model in a pickle file format with the help of pickle module.

3.8 Data from User
Here the user will have to enter all the features values in correct order and have to submit it to the model with the help of UI interface. The data will be fed to the model which will predict whether the feature set represents the required traits.

3.9 Data Validation

Here Data Validation will be done, which has given by the user.
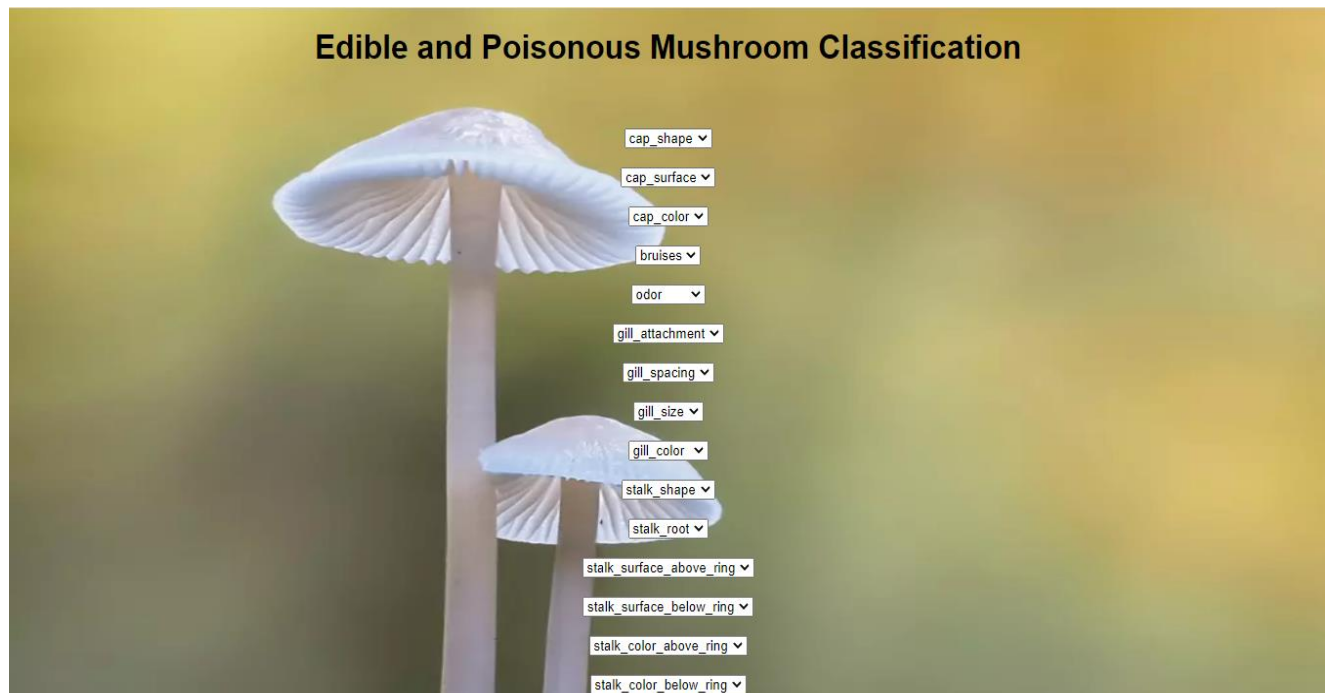
3.10 Model Call for specific input

Based on the user Input, the data will be processed in the backend with the variable format, then it will be converted into a NumPy array which will be fed to ML model. The loading of the pickle file will be done and then the model will predict whether the inputs were meeting the required traits by sending the result to our html page.
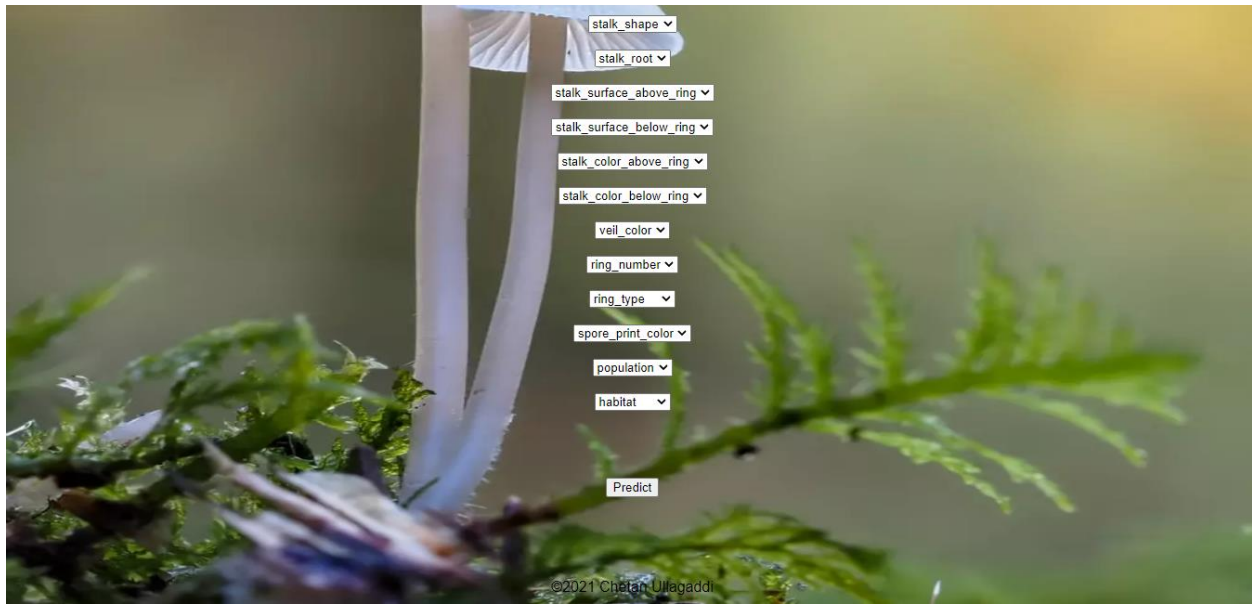
3.11 User Interface

In Frontend creation, I have made a user interactive page where users can enter their input values to application in their frontend page. a web page which is having attractive overview with CSS. This HTML user input data is transferred in variable format to the backend and same is in a decoupled format.

Input Page:

Output Page:

3.12 Deployment

The deployment of the model with the help of Heroku cloud platform.

4.Technology Stack

Front End	HTML/CSS
Back End	Flask, Pandas, NumPy, Sci-kit learn, etc
Deployment	Heroku