

# Mushroom classification

## Document version control

Date issued	version	Description	author
23/8/2021	1	Initial HLD	Nishanth kumar S
13/9/2021	2	Updated Design	Nishanth kumar S

## Contents

<b>Document version control</b>	<b>1</b>
<b>Contents</b>	<b>2</b>
<b>Abstract</b>	<b>3</b>
<b>Introduction</b>	<b>3</b>
Why this High-Level Design Document:	3
Scope:	4
<b>General Description</b>	<b>4</b>
Product perspective:	4
Problem Statement:	5
Proposed Solution:	5
Technical requirements:	5
Data Requirements:	5
Tools used:	5
7.Hardware requirements:	6
8. Streamlit:	6
9.Assumptions:	6
<b>Design Details</b>	<b>7</b>
Process Workflow	7
Event log	8
Error Handling	9
<b>Performance</b>	<b>9</b>
1.Reusability:	9
2. Application Compatibility:	9
3 Resource Utilization	9
4. Deployment:	9
<b>Conclusion</b>	<b>10</b>

## Abstract

The Audubon Society Field Guide to North American Mushrooms contains descriptions of hypothetical samples corresponding to 23 species of gilled mushrooms in the Agaricus and Lepiota Family Mushroom (1981). Each species is labelled as either definitely edible, definitely poisonous, or maybe edible but not recommended. This last category was merged with the toxic category. The Guide asserts unequivocally that there is no simple rule for judging a mushroom's edibility, such as "leaflets three, leave it be" for Poisonous Oak and Ivy.

The main goal is to predict which mushroom is poisonous & which is edible.

## Introduction

### 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

## 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.

## General Description

### 1. **Product perspective:**

The mushroom classifier is a machine learning model which will help us to find a mushroom is edible or poisonous based on the given information

## 2. Problem Statement:

The Audubon Society Field Guide to North American Mushrooms contains descriptions of hypothetical samples corresponding to 23 species of gilled mushrooms in the Agaricus and Lepiota Family Mushroom (1981). Each species is labelled as either definitely edible, definitely poisonous, or maybe edible but not recommended. This last category was merged with the toxic category. The Guide asserts unequivocally that there is no simple rule for judging a mushroom's edibility, such as "leaflets three, leave it be" for Poisonous Oak and Ivy. The main goal is to predict which mushroom is poisonous & which is edible.

## 3. Proposed Solution:

We used both decision trees and random forest classification with feature selection and got a 100% accuracy in testing phase and also cross verified if there is data leakage . In the implementation phase we used random forest model

## 4. Technical requirements:

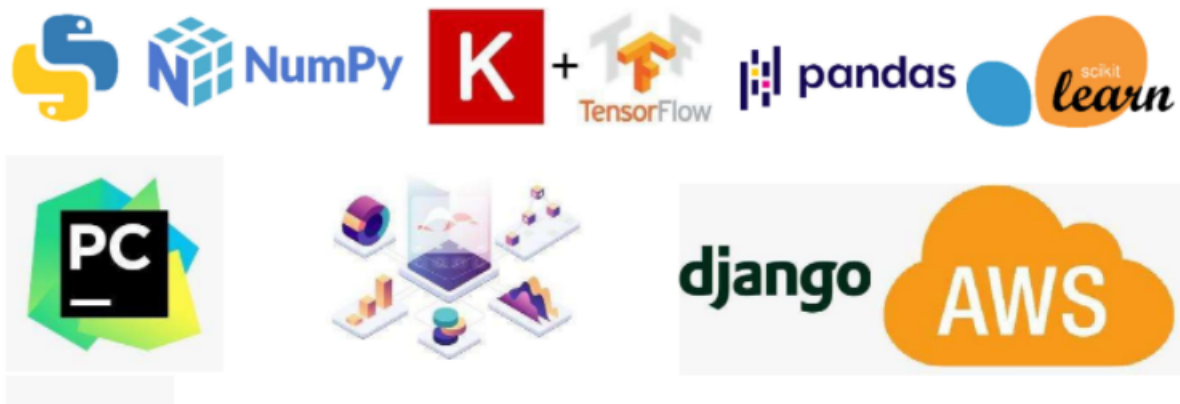
Operating system: **Linux- Ubuntu 16.04 to 17.10**, or Windows 7 to 10, with 2GB RAM (4GB preferable) 2. You have to install Python 3.6 and related packages

## 5. Data Requirements:

'cap-shape', 'bruises', 'odor', 'gill-spacing', 'gill-size','gill-color', 'stalk-root', 'stalk-surface-above-ring','stalk-surface-below-ring', 'stalk-color-above-ring','stalk-color-below-ring', 'ring-type', 'spore-print-color','population', 'habitat' values are required

## 6. Tools used:

Python programs and framework such as Numpy,Pandas,Scikit-Learn,Streamlit ,Pycharm is used to build the whole model



- PyCharm is used as IDE.
- For visualization of the plots, Matplotlib, Seaborn and Plotly are used.
- Python Streamlit is used for backend development.
- GitHub is used as version control system

## 7. Hardware requirements:

Operating system: **Linux- Ubuntu 16.04 to 17.10**, or Windows 7 to 10, with 2GB RAM (4GB preferable) 2. You have to install Python 3.6 and related packages

## 8. Streamlit:

Streamlit is an open-source Python library that makes it easy to create and share beautiful, custom web apps for machine learning and data science. In just a few minutes you can build and deploy powerful data apps

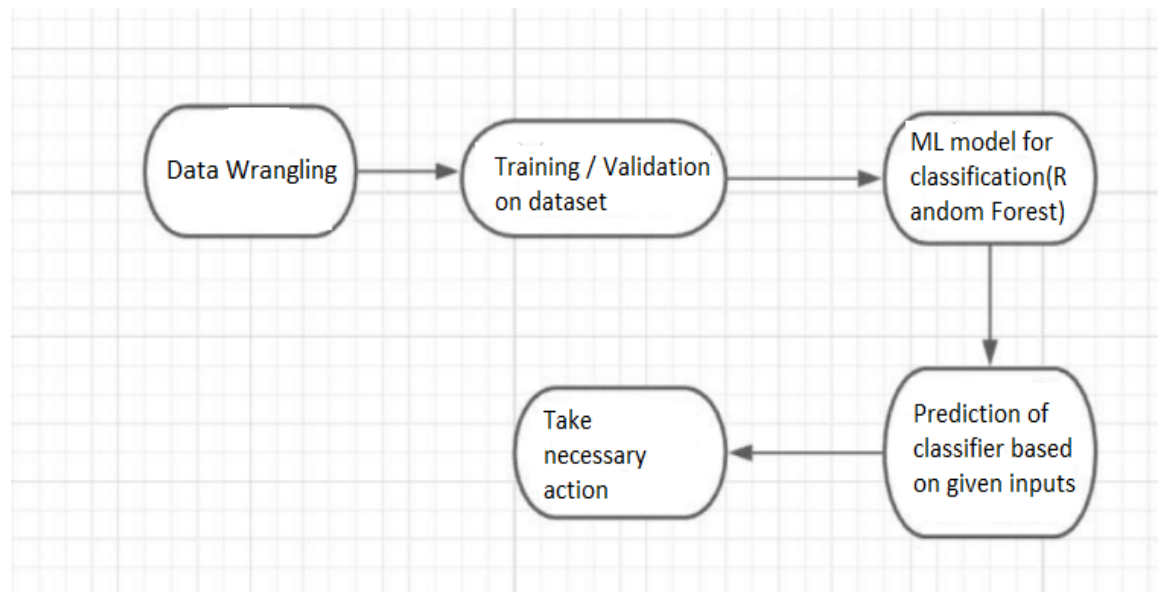
## 9. Assumptions:

The user is having all the 15 feature values {'cap-shape', 'bruises', 'odor', 'gill-spacing', 'gill-size', 'gill-color', 'stalk-root', 'stalk-surface-above-ring', 'stalk-surface-below-ring', 'stalk-color-above-ring', 'stalk-color-below-ring', 'ring-type', 'spore-print-color', 'population', 'habitat' values are required}

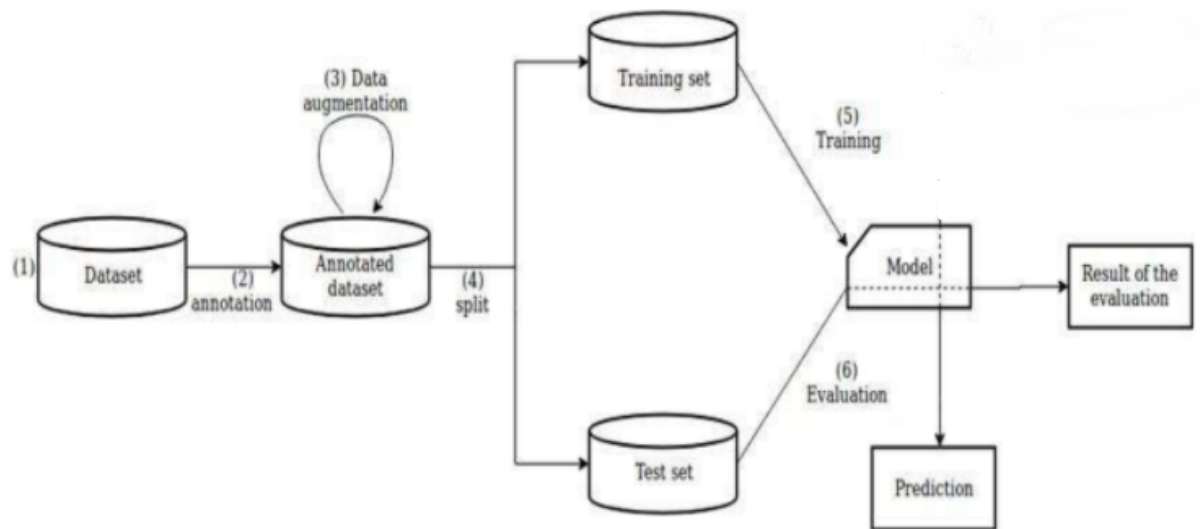
## Design Details

### 1. Process Workflow

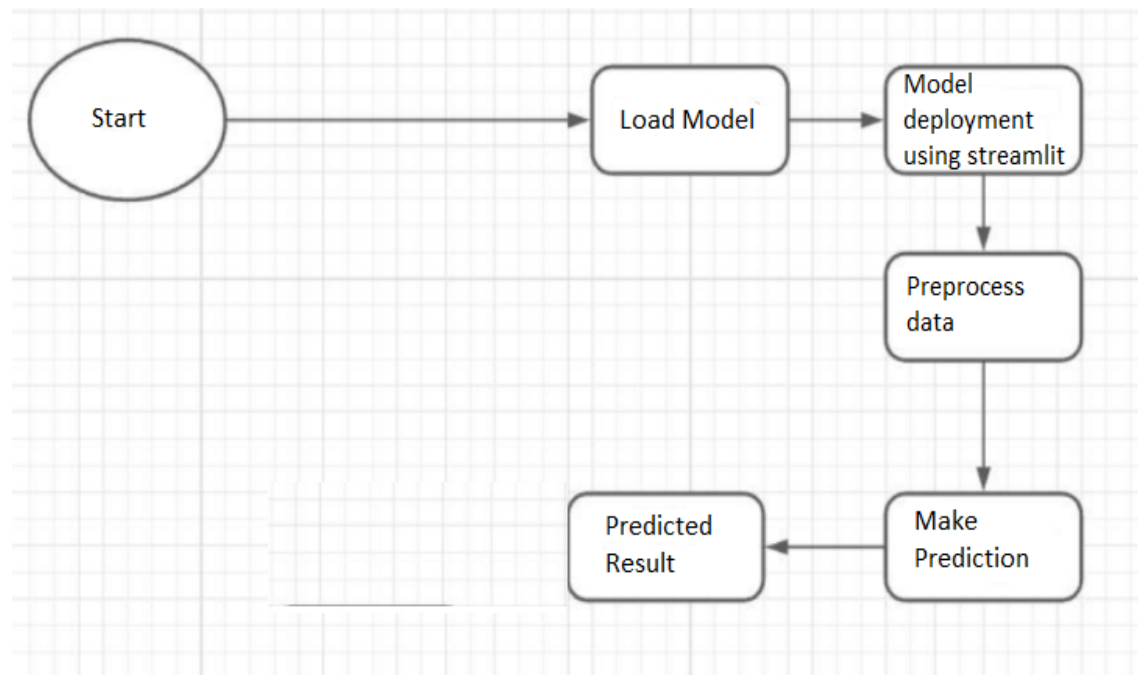
- Proposed Methodology



- Model Training & evaluation



- Deployment Process



## 2. Event log

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed diam nonummy nibh euismod tincidunt ut laoreet dolore magna aliquam erat volutpat.



### 3. Error Handling

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed diam nonummy nibh euismod tincidunt ut laoreet dolore magna aliquam erat volutpat.

## Performance

### 1.Reusability:

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

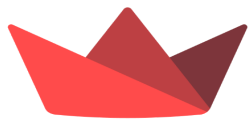
### 2. Application Compatibility:

- 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,

### 3 Resource Utilization

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

### 4. Deployment:



# Streamlit

## Conclusion

- The Designed Mushroom Classifier will Predict a mushroom is edible or not based on various anomalies data used to train our algorithm, so we can identify the poisonous mushroom and can take necessary action to stop them immediately, so we can minimize the death rate which is caused by eating poisonous mushroom and also for better health conditions ( Mushrooms are a rich, low calorie source of fiber, protein, and antioxidants. They may also mitigate the risk of developing serious health conditions, such as Alzheimer's, heart disease, cancer, and diabetes. They're also great sources of: Selenium. )