High Level Design (HLD)

Mushroom Classifications

By Rohan Kadam & Vijay Kamble

Document Control

Date	Version	Description	Author
12/09/2023	1.0	Abstract, Introductions, Scope	Rohan Kadam
14/09/2023	1.2	Method	Vijay Kamble
17/09/2023	1.3	Tool , Deployment	Rohan Kadam
19/09/2023	1.4	Report	Vijay Kamble

Contents Document Version

Abstract

Mushrooms are used in the medical field to cure diseases like anemia, increase body immunity, diabetes, and cancer treatment. Some of the mushroom varieties are edible, but others are highly poisonous. The identification of mushrooms whether edible or poisonous is a difficult process because of the large number of mushrooms have similar characteristics. The principle of this paper is to classify the mushrooms by using Machine learning classification algorithms through a data mining tool. For the best classification, five classification algorithms were compared.

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 before 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.

2 General Description 2.1 Product Perspective & Problem Statement

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 denitely edible, denitely 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 "leaets three, leave it be" for Poisonous Oak and Ivy. The main goal is to
predict which mushroom is poisonous & which is edible.
Approach: The classical machine learning tasks like Data Exploration, Data Cleaning, Feature Engineering, Model Building and Model Testing.
Try out different machine learning algorithms that’s best t for the above case.
Results: You have to build a solution that should able to predict which mushroom is poisonous & which is edible.

Tools


Google Colaboratory
NumPy
seaborn
pandas
jupyter
python
matplotlib

2. Deployment

Deployment on Google collbs python IDE

Google Colaboratory
Colab is a hosted Jupyter Notebook service that requires no setup to use and provides free access to computing resources, including GPUs and TPUs. Colab is especially well suited to machine learning, data science, and education.

Applying Mushroom classifications

Sr.no	Classifications	Result
1	#Random Forest Classification	100%
2	#Decision tree classification	100%
3	#naive bayes classfications	91.88%
4	#SVM Classification	100%
5	#K-Nearest neighbour	100%
6	#Logistic Regressions	94%

Publish Datasets and Reports

From the confusion matrix, we saw that our train and test data is balanced.
Most of classfication methods hit 100% accuracy with this dataset.

In conclusion, the application of machine learning in mushroom classification has demonstrated its remarkable potential in automating and enhancing the accuracy of identifying mushroom species. Through the utilization of advanced algorithms and vast datasets, we have witnessed the development of robust models capable of distinguishing between edible and toxic mushrooms with a high degree of confidence. As technology continues to advance and more research is conducted, we can anticipate even greater strides in the accuracy and efficiency of mushroom classification using machine learning. This, in turn, will contribute to safer mushroom foraging practices, greater understanding of fungal biodiversity, and the preservation of ecosystems.