* **Functional requirements, specifying what the final product will do**

We have created a website that helps the farmer to determine the characteristics of the mushrooms he has collected on his farm. To do this, he has to answer 21 questions that determine the properties of the mushrooms, so that the result then shows him whether these mushrooms are edible or poisonous.

* **Data requirements, specifying what data will be processed**

This data set includes descriptions of hypothetical samples corresponding to 23 species of gilled mushrooms in the Agaricus and Lapita Family . 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. The Guide clearly states that there is no simple rule for determining the edibility of a mushroom; no rule like ``leaflets three, let it be'' for Poisonous Oak and Ivy.

صورة تحتوي على نص, رقم, لقطة شاشة, الخط

تم إنشاء الوصف تلقائياً

While preparing the data set, features contain 2480 of Nans in stalk-root , we removed bad column.

* Original X shape: (8124, 22)
* X shape after bad column removal: (8124, 21)صورة تحتوي على نص, لقطة شاشة, برمجيات, أيقونة الحاسوب

  تم إنشاء الوصف تلقائياً

صورة تحتوي على نص, لقطة شاشة, برمجيات, أيقونة الحاسوب

تم إنشاء الوصف تلقائياً

Fix unbalanced data and Check for outliers

* Otliers found: 26
* X shape after otliers removal: (8098, 21)
* y shape otliers removal: (8098, 1)

Original data contains 4200 eatable and 3898 poisonous mushrooms. We added more poisonous with SMOTE.

* X shape after SMOTE: (8400, 21)
* y shape after SMOTE: (8400, 1)

Check variance and apply dimension reduction, we can lower the dimension to 19 without losing 0.001% of variance.

X shape after dimension reduction: (8400, 19)

* **Design requirements, specifying design structure, modules, algorithms, packages to be used**

|  |  |  |
| --- | --- | --- |
| Libraries and packages | | |
| pandas | | |
| ucimlrepo | fetch\_ucirepo | |
| sklearn | Preprocessing | LabelEncoder |
| MinMaxScaler |
| neighbors | LocalOutlierFactor |
| decomposition | PCA |
| model\_selection | train\_test\_split |
| sklearn.model\_selection | train\_test\_split | |
| keras.layers | Dense | |
| Dropout | |
| imblearn.over\_sampling | SMOTE | |
| tensorflow | | |
| tensorflow.keras.models | Sequential | |
| matplotlib.pyplot | | |

* **Steps**

1. Import Libraries and packages
2. Load the data
3. Prepare the data
4. Split the data
5. Prepare the model
6. Learn from the data
7. Plot

* **User requirements, specifying user types, user functionalities (functional requirements breakdown in terms of user types), look and feel**
* The website is aimed at a user (the farmer) who has to answer 21 multiple-choice questions in which he determines the shape and characteristics of the mushrooms he collected.

**Choose the shape and characteristics of the mushroom below**

Q1: Cap shape

bell  conical  convex  flat  knobbed  sunken

Q2: Cap surface

fibrous  grooves  scaly  smooth

Q3: Cap color

brown  buff  cinnamon  gray  green  pink  purple  red  white  yellow

Q4: Bruises

bruises  no

Q5: Odor

almond  anise  creosote  fishy  foul  musty  none  pungent  spicy

Q6: Gill attachment

Attached  descending  free  notched

Q7: Gill spacing

close  crowded  distant

Q8: Gill size

broad  narrow

Q9: Gill color

black brown buff  chocolate  gray  green orange pink purple  red  white  yellow

Q10: Stalk shape

enlarging  tapering

Q11: Stalk surface above ring

fibrous  scaly  silky  smooth

Q12: Stalk surface below ring

fibrous  scaly  silky  smooth

Q13: Stalk color above ring

brown  buff  cinnamon  gray  orange  pink  red  white  yellow

Q14: Stalk color below ring

brown  buff  cinnamon  gray  orange  pink  red  white  yellow

Q15: Veil type

partial  universal

Q16: Veil color

brown  orange  white  yellow

Q17: Ring number

none  one  two

Q18: Ring type

cobwebby  evanescent  flaring  large  none  pendant  sheathing  zone

Q19: Spore print color

black  brown  buff  chocolate  green  orange  purple  white  yellow

Q20: Population

abundant  clustered  numerous  scattered  several  solitary

Q21: Habitat

grasses  leaves  meadows  paths  urban  waste  woods

* After completing the 21 single-choice questions of the system, the result shows whether the mushroom is edible or poisonous.
* Warning message to the user: Our model can predict poisonous mushrooms based on the provided description with 99.9% accuracy. But not 100%. So, be careful with all the mushrooms you are going to eat.
* **Any other non-functional requirements such as performance requirements, security requirements**
* The website will be design a user-friendly interface with clear navigation, labels, and instructions for easy mushroom identification .
* The website works well across different browsers, devices, and operating systems.
* Ability to handle increased load or expand functionality without major architectural changes.
* Making it easy for developers to update, maintain, and enhance the website over time.