iNeuron

MUSHROOM CLASSIFICATION
(Machine Learning)

WIREFRAME

Project Member: Sonali Bhadra

1. Data Acquisition:

- Source data from a reliable dataset, such as the UCI Mushroom dataset.
- Import necessary libraries: pandas, numpy, matplotlib, seaborn, sklearn.

2. Data Exploration and Preprocessing:

- Load the dataset into a pandas DataFrame.
- Explore the dataset's dimensions, summary statistics, and missing values.
- Convert categorical data to numerical using LabelEncoder.
- Define feature columns (X) and target variables (Y).
- Perform Exploratory Data Analysis (EDA) to understand the distribution of features.

3. Principal Component Analysis (PCA):

- Apply PCA to reduce the dimensionality of the feature space.
- Choose an appropriate number of components

4. Model Training:

- Split the dataset into training and testing sets.
- Implement machine learning models:
  - Logistic Regression
  - K-Nearest Neighbors
  - Support Vector Classifier
  - Decision Tree Classifier
  - Random Forest Classifier
  - Gradient Boosting Classifier

5. Model Evaluation:

- Use accuracy score to evaluate model performance on the test set.
- Compare and analyze the performance of each model.

6. Model Serialization:

- Save the best-performing model (Random Forest) using joblib for future use.

7. GUI Application:

- Create a simple GUI using Tkinter for user interaction.
- Include input fields for each feature to predict the class.
- Load the saved model and make predictions based on user input.

8. Visualization:

- Plot a bar chart to display the accuracy of each model.
- Create visualizations of the dataset features.

Mushroom Classification Using Machine Learning

cap-shape:(cap shape: bell=0,conical=1,convex=5,flat=2, knobbed=3,sunken=4)

cap-surface:(fibrous=0,grooves=1,scaly=3,smooth=2)

cap-color:(brown=4,buff=0,cinnamon=1,gray=3,green=r, pink=5,purple=6,red=2,white=7,yellow=8)

bruises:(bruises=1,no=0)

odor:(almond=0,anise=3,creosote=1,fishy=8,foul=2,musty=4,none=5,pungent=6,spicy=7)

gill-attachment:(attached=0,descending=1,free=2,notched=3)

gill-spacing:(close=0,crowded=2,distant=1)

gill-size:(road=0,narrow=1)

gill-color:(black=4,brown=5,buff=0,chocolate=3,gray=2,green=8,orange=6,pink=7,purple=9,red=1,white=10,yellow=11)

stalk-shape:(enlarging=0,tapering=1)

stalk-root:(bulbous=0,club=1,cup=5,equal=2,rhizomorphs=4, rooted=3,missing=6)

stalk-surface-above-ring:(fibrous=0,scaly=3,silky=1,smooth=2)

stalk-surface-below-ring:(fibrous=0,scaly=3,silky=1,smooth=2)

stalk-color-above-ring:(brown=4,buff=0,cinnamon=1,gray=3, orange=5,pink=6,red=2,white=7,yellow=8)

stalk-color-below-ring:(brown=4,buff=0,cinnamon=1,gray=3, orange=5,pink=6,red=2,white=7,yellow=8)

veil-type:(partial=0,universal=1)

veil-color:(brown=0,orange=1,white=2,yellow=3)

ring-number:(none=0,one=1,two=2)

ring-type:(cobwebby=0,evanescent=1,flaring=2,large=3,none=4,pendant=5,sheathing=6,zone=7)

spore-print-color:(black=2,brown=3,buff=0,chocolate=1, green=5,orange=4,purple=6,white=7,yellow=8)

population:(abundant=0,clustered=1,numerous=2,scattered=3, # several=4,solitary=5)

habitat:(grasses=1,leaves=2,meadows=3,paths=4,urban=5,# waste=6,woods=0)

Predict