iNeuron

LLD (LOW LEVEL DESIGN)

MUSHROOM CLASSIFICATION

AUTHOR : RAVINDRA V

Index　Page no

Problem Statement :

To build a model to predict whether a mushroom is edible or poisonous , by looking at 23 species of gilled mushrooms we have to identify whether a mushroom is edible or poisonous
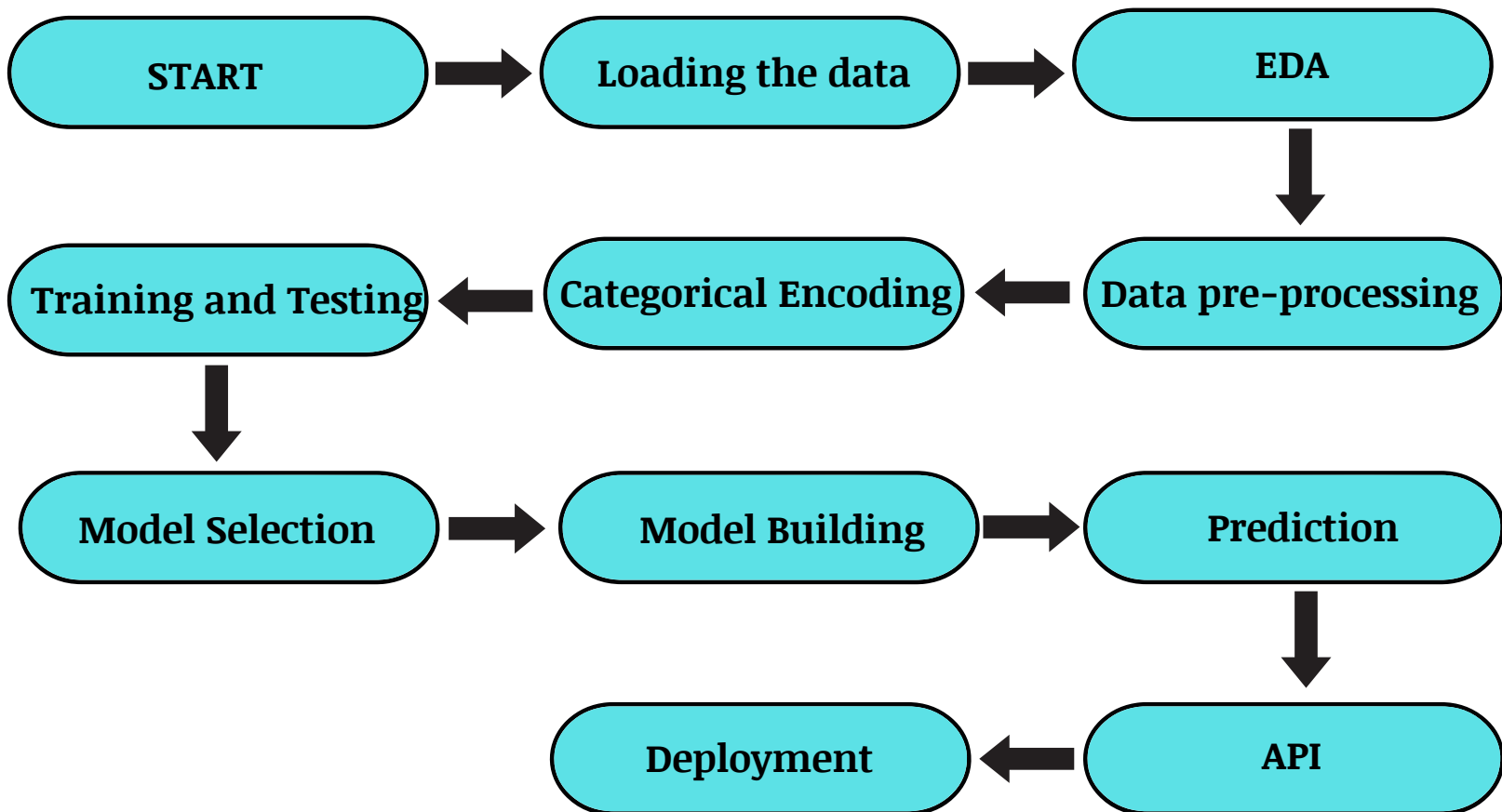
Introduction :

Low-level design is a detailed description of every module of software. It describes every module in detail by incorporating the logic behind every component in the system. It delves deep into every specification of every system, providing a micro-level design.

The goal of LLD (Low level design document) is to give the internal logical design of the actual program code for mushroom classification. LLD describes the diagrams with the methods and relation between features and level columns. It helps the user to understand the code and document

Scope :

Low-Level design(LLd) is a process that follows a step-by-step refinement process. This process can be used for designing data structures, required software architecture, source code and performance algorithms, It describes every module in detail by incorporating the logic behind every component in the system.

Architecture :

```
START  →  Loading the data  →  EDA
                                 ↓
Training and Testing  ←  Categorical Encoding  ←  Data pre-processing
        ↓
Model Selection  →  Model Building  →  Prediction
                                          ↓
Deployment  ←  API
```

ARCHITECTURE DESCRIPTION

DATA :

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. It consist of 8123 rows and 23 columns,the dataset is exported from kaggle to local device

Loading the data :

Data export from database - the data is stored locally which is exported as csv file kaggle ,this data will be used in further data training and model building

EDA :

we are using EDA (Exploratory data analysis) to analyze and investigate data sets and summarize their main characteristics, often employing data visualization methods. this method will help us to get all the insights of mushroom classification

Data pre-processing :

Data pre-processing the one of the import step in ML, data preprocessing include cleaning, instance selection, normalization, Label encoding one hot encoding, transformation, feature extraction and selection, etc. The product of data preprocessing is the final training set.

Categorical Encoding :

Categorical encoding is the process where we convert our categorical data into numerical data , the mushroom classification dataset consist of categorical data by the help of Label encoding we converted all categorical features into numerical values.

Training and Testing :

This is the process where we divide some of the data into training and some of the data into testing, in this case we took 67% of data for training and 33% of data for testing with the help of train_test_split

Model Selection :

Model selection refers to the process of choosing the model that best generalizes. Model selection is a process that can be applied both across different types of models (e.g. logistic regression, SVM, KNN, Random forest, DecisionTree, etc.)

Model Building :

After working and checking with every model , we can conclude with the model which is providing us the best accuracy and the best accuracy model is saved in pickle file format

Prediction :

After finding the best model we will make the predictions with the test dataset if the model gives the best results than we can conclude with the results.

API :

By using Streamlit API application we can test and run the project locally as well as in cloud platforms

Deployment :

After running the project successfully in locally as well as in cloud platforms (eg. Heroku) we can share the project via link and the users can easily access the link as it is shared publicly

Conclusion :

This is a web based application, I have used streamlit for the user interface. For making predictions whether a mushroom is edible or poisonous the user have to enter all the features of mushroom as input