



Mushroom Classification

Abstract:

When compared to all other vegetables, mushrooms are a significant fungus that provides a decent supply of vitamin B and a significant quantity of protein. It aids in the prevention of cancer, aids in weight loss, and strengthens human immunity. On the other side, certain mushrooms are poisonous and might be risky. Consequently, identifying edible from toxic mushrooms is a crucial endeavor.

Project Scope:

- Data parsing (data preparation)
- Creating a machine learning model to predict whether or not a mushroom is edible based on its other properties.

Design Steps/Methodology:

1. Data Preparation:
 - Retrieve the dataset. (Loading the dataset into the jupyter notebook)
 - Understand the data.
 - Preprocess the data by dealing with missing values, encoding categorical variables, and carrying out the necessary transformations.
2. Feature Selection and Engineering:

- Determine the characteristics that can be used to predict the edibility of mushrooms.
- If required, create additional features based on domain expertise or data insights.
- 3. Model Selection and Training:
 - Split the dataset into training and testing sets to evaluate the performance of model.
 - Choosing the machine learning algorithm for binary classification. Ex: - Logistic Regression, Decision trees, Random Forest classifiers.
 - Train the selected model on training data and tune hyperparameters using techniques like cross-validation or grid search.
- 4. Model Evaluation:
 - Evaluate performance of trained model using evaluation metrics like accuracy, precision, recall and F1 score.
 - Examine any class discrepancies and contemplate the utilization of methods such as over- or under-sampling to rectify the discrepancies.

Result and Outcomes:

The methods that we used here are: -

- Support vector machine = 100% accuracy
- Decision Tree Classifier = 97.9% accuracy
- Logistic Regression = 95% accuracy

Benefits:

Classifying mushrooms has various advantages, particularly when it comes to identifying edible mushrooms from harmful ones.

The key benefits include: -

1. Food Safety
2. Culinary Applications
3. Medicinal Uses
4. Biodiversity Conservation
5. Scientific Research
6. Recreational and Educational Purposes

Accurate identification plays a crucial role in maximizing the benefits and minimizing the risks associated with mushrooms.

Lessons Learnt:

1. Understanding the problem: - It took some time for me to grasp the major goal of this specific endeavor. After understanding the objectives, needs, and restrictions, I was able to readily categorize the information.
2. Data preprocessing: - I learnt that this is a very critical stage in handling any data. Understanding the variables and statistical quality of the provided data, as well as processing the data from a

csv file into a legible manner. I spent most of my time cleaning the data, removing null values, formatting the values, restoring missing values, and encoding the data to binary values.

3. Model selection: - There are several techniques for encoding data. I employed three approaches in picking the machine learning algorithm: logistic regression, decision tree classifier, and support vector machine (SVM).
4. Model Interpretation and Error analysis: - As this portion was entirely new to me, splitting the data into training and testing was a time-consuming work for me. I examined how it should be done and the underlying method behind it. While coding, I did encounter several difficulties and make mistakes that I was unable to correct. then worked on those with assistance from Stack Overflow and my colleagues.
5. Continuous learning: - I was overjoyed to be given this project because I am interested in AI and machine learning. As this was my first project, I was eager to learn about every aspect of it, and I read through the majority of the documentation and several forums.

Conclusion:

The accuracy of the procedures utilized is greater than 95%, implying that what we have is edible. Most of the classifiers perform well as the dataset is a clean set.