

NAME: Adeyemi Abodunrin , Data:<https://archive.ics.uci.edu/dataset/73/mushroom>,collab:
Report on classification model.ipynb

Report on Classification Models: SVM, Decision Tree, and Random Forest for Mushroom Classification

Introduction: The aim of this study is to explore and evaluate the performance of three different machine learning algorithms, which are; Support Vector Machine (SVM), Decision Tree, and Random Forest, in classifying mushrooms as either edible or poisonous based on certain features. Each model's hyperparameters were optimized using grid search, and their performances were compared and contrasted to identify the most effective classifier for this specific task.

Features Used for ClassificationThe classification task was performed using the following features extracted from the mushroom dataset:

- Cap-shape
- Cap-surface
- Cap-color
- Bruises
- Odor
- Gill-attachment
- Gill-spacing
- Gill-size
- Gill-color
- Stalk-shape
- Stalk-root
- Ring-number
- Stalk-surface-above-ring
- Stalk-surface-below-ring
- Stalk-color-above-ring
- Stalk-color-below-ring
- Veil-type
- Veil-color
- Habitat
- Population
- Spore-print-color
- Ring-type

Topic 1 - Support Vector Machine (SVM)

Support Vector Machine (SVM) is a powerful and versatile machine learning algorithm used for both classification and regression tasks. In classification, SVM aims to find the hyperplane that **best** separates the data points of different classes in the feature space. The algorithm works by maximizing the margin between the closest points of different classes, known as support vectors.

Result

Best Hyperparameters:

- C: 10
- Gamma: scale
- Kernel: linear

4128	8
0	3916

predicted poisonous as edible 80 times

Confusion Matrix:SVM

Accuracy: 99.02%

Classification Report:D-tree:precision , F1 and Recall : 1.00

Classification Report:SVM

	precision	Recall	F1-score
0	1.00	0.98	0.99
1	0.98	1.00	0.99

Decision Tree Results

Best Hyperparameters:

- Criterion: gini
- Max Depth: None
- Min Samples Leaf: 2
- Min Samples Split: 2

Confusion matrix: D-TREE

4208	0
0	3916

Accuracy: 100.00%

Comparison and Conclusion

All three models performed exceptionally well with 100% accuracy in classifying mushrooms as edible or poisonous. The Decision Tree and Random Forest models achieved a slightly better performance compared to the SVM model.

Summary of Key Features:The key features used for classification were highly informative and discriminative for distinguishing between edible and poisonous mushrooms.

Recommendations

Given the comparable performance of the Decision Tree and Random Forest models with 100% accuracy, either of these models can be

Random Forest Results

Best Hyperparameters:

- Bootstrap: False
- Max Depth: None
- Min Samples Leaf: 1
- Min Samples Split: 2
- N Estimators: 200

Confusion Matrix: RANDOM FOREST

4208	0
0	3916

Classification Report:Random forest

precision , F1 and Recall : 1.00

recommended for practical application in mushroom classification. The SVM model, although slightly less accurate, can also be considered for its robust and reliable performance.

References

1. Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20(3), 273-297.
2. Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5-32.
3. Quinlan, J. R. (1986). Induction of decision trees. *Machine learning*, 1(1), 81-106.

