

Mushroom Classification Using Machine Learning Models

Jana Ali Redaini, Wejdan Mostafa Banafa, Samar Mishal Alamri, and Nada Osamah Oraijah

Instructor: Dr. Eaman Alharbi

Department of Artificial Intelligence, King Abdulaziz University, Saudi Arabia

Abstract

This project presents a machine learning approach for classifying mushrooms as edible or poisonous using the UCI Mushroom Dataset, which contains 8124 samples described by 22 categorical attributes. After applying one hot encoding to prepare the data, two classification models were developed and evaluated: a Decision Tree classifier and a Random Forest classifier. The Decision Tree achieved 99.75% accuracy, while the Random Forest achieved 99.51%, demonstrating that both models can effectively identify toxicity based on physical characteristics such as odor and spore print color. A robustness experiment was also conducted by removing the two most influential features, and the Random Forest maintained strong performance with 98.46% accuracy. These results indicate that machine learning provides a reliable solution for mushroom toxicity prediction, with the Random Forest offering stronger generalization and resilience to feature reduction.

Index Terms

Machine Learning, Mushroom Classification, Decision Tree, Random Forest.

I. INTRODUCTION

In the fields of mycology and food safety, the difference between edible and dangerous mushrooms is essential. It is critical to accurately determine the edibility of mushrooms to avoid any possible health risks related to consuming harmful species. Also, the gathering of wild mushrooms is a popular hobby practiced by enthusiasts worldwide. However, it is an activity filled with risk. Many toxic mushroom species look very similar to their edible counterparts, and misidentification can lead to severe illness or even death. This high risk "classification" problem has traditionally been solved by human experts, but this knowledge is specialized and not always readily available.

This project suggests that machine learning (ML) offers a reliable and scalable solution to this challenge. By training a model on a dataset of known mushroom characteristics, an

ML algorithm can learn the complex patterns and subtle connections that separate an edible species from a poisonous one. The goal is to build a model that can, with high accuracy, classify a mushroom as "edible" or "poisonous" based on a set of descriptive features.

This project will use the UCI Mushroom Dataset, which contains over 8,000 samples described by 22 physical features, like odor and habitat. The project will utilize the Decision Tree and Random Forest algorithms, and a comparative analysis will be conducted between them.

II. RELATED WORK

Machine learning techniques have been widely applied to mushroom classification, particularly for distinguishing edible from poisonous species. Researchers have explored various approaches, including image-based analysis, structured categorical data, and ensemble methods, to enhance classification accuracy, interpretability, and computational efficiency. In this section, we provide a chronological overview of notable studies in this area, highlighting the datasets, methodologies, and key findings that have shaped current approaches.

One of the earlier efforts in this domain was conducted by Vanitha et al. [1], who presented "Classification of Mushrooms to Detect their Edibility Based on Key Attributes," focusing on structured categorical data from the UCI Mushroom dataset. The authors applied two attribute-selection techniques (Wrapper and Filter) in the WEKA framework to identify the most influential features, finding that *odor* and *spore_print_color* were the most significant predictors. A decision-tree classifier was trained on the selected attributes, and its performance was evaluated using precision, recall, and F1-score. This study demonstrated the importance of attribute selection for interpretability and classification performance, though it did not perform extensive comparisons across multiple classifier types or ensemble methods.

Building upon this foundation, Kanchi Tank [2] investigated dimensionality reduction and supervised learning in "A Comparative Study on Mushroom Classification using Supervised Machine Learning Algorithms," utilizing the UCI Mushroom dataset together with Principal Component Analysis (PCA). The study evaluated Logistic Regression, Decision Tree, K-Nearest Neighbors (KNN), Support Vector Machine(SVM), Naïve Bayes, and Random Forest models, comparing their performance using ROC curves and discussing how PCA enhances classifier efficiency and interpretability. The authors concluded that ensemble and tree-based approaches frequently deliver strong performance on the structured UCI dataset, confirming the value of combining feature-selection and classifier optimization.

Further advancements were introduced by Tarawneh et al. [3], who proposed a hybrid ensemble model in "Mushroom Classification Using Machine-Learning Techniques." Using the UCI dataset, they trained several individual classifiers, which are Naïve Bayes, Decision Tree, SVM, and Artificial Neural Network (ANN). And combined the top-performing models

into a fused decision system. The ensemble model achieved approximately 94% accuracy, outperforming the individual classifiers, and highlighted the robustness benefits of integrating multiple classifiers. The study underscored the value of ensemble approaches for improving predictive performance, while still relying on structured categorical data rather than images or other complex features.

Building on this line of work, the study by Samantara et al. [4] “Mushroom Quality Prediction using Machine Learning Classification”, focuses on automating the identification of edible and poisonous mushrooms using machine learning approaches utilizing the UCI Mushroom dataset. Before classification, the dataset underwent basic preprocessing in WEKA: the authors handled missing values, ensured nominal data were binary-encoded, and verified that the class distribution was nearly balanced (51.28% edible and 48.2% poisonous). Two machine learning classifiers were compared, first the Naïve Bayes, which assumes independent feature contribution, and second the Sequential Minimal Optimization (SMO). Both models were evaluated using 10-fold cross-validation, and performance was assessed using accuracy, kappa statistic, mean absolute error (MAE), root mean squared error (RMSE), relative absolute error (RAE), and root relative squared error (RRSE). The results showed that SMO achieved perfect classification with 100% accuracy, zero error rates, and $k = 1.0$, while Naïve Bayes reached 95.82% accuracy with a kappa value of 0.9162. The confusion matrix revealed that Naïve Bayes occasionally misclassified mushrooms due to correlated attributes like odor and spore color, whereas SMO accurately separated both classes without overlap. The authors concluded that SMO (SVM) is more robust for categorical datasets with complex boundaries, demonstrating its effectiveness in mushroom quality prediction.

Similarly, Shehab et al. [5] presented “The Classification of Mushroom Using Machine Learning.” Used the Kaggle mushroom dataset, originally derived from The Audubon Society Field Guide to North American Mushrooms. The researchers followed a three-phase methodology involving data preprocessing, algorithm application, and performance evaluation. They compared four machine learning algorithms: Random Forest, Decision Tree, Logistic Regression, and Multilayer Perceptron (MLP) to assess their classification performance. The findings revealed that the Random Forest model achieved the highest accuracy of 98.7%, surpassing the Decision Tree (98.0%), MLP (98.54%), and Logistic Regression (94.38%). Furthermore, the evaluation based on precision, recall, specificity, and F1-score validated the robustness of the Random Forest approach.

A more specialized perspective was later offered by Ujir et al. [6], who introduced “Automating Mushroom Culture Classification: A Machine Learning Approach,” focusing on image-based prediction of growth phases and contamination during *Pleurotus ostreatus* cultivation. Working with 1400 images and evaluating both HSV histogram-based descriptors and CNN-derived features, the study found that CNN architectures (MNet/MConNet) substantially outperformed handcrafted color-based features, achieving average accuracies near 92.15% for growth prediction and 97.81% for contamination detection. Their results demonstrate

how convolutional neural networks can support cultivation-monitoring tasks complementary to toxicity and species classification research.

Ortiz-Letechipia et al. [7] presented a comprehensive feature-selection study titled “Classification and selection of the main features for the identification of toxicity in *Agaricus* and *Lepiota* with machine learning algorithms.” Recognizing that some mushroom species are beneficial while others are toxic, the study analyzed organoleptic, ecological, and morphological characteristics to identify the most significant features for fungal toxicity prediction. Using genetic algorithms (GALGO) and LASSO regression, the authors found that odor, spore print color, and habitat were the most important features according to GALGO, while odor, gill size, stalk shape, and twelve additional features were significant in the LASSO selection. Quantitative importance scores highlighted odor at 99.99%, gill size at 73.7%, and stalk shape at 39.9%, with other features below 18%. The selected feature sets were used to train logistic regression, KNN, and XGBoost classifiers, evaluated using sensitivity, specificity, and accuracy metrics. Models based on LASSO-selected features achieved a maximum AUC of 0.99 with XGBoost and KNN, while GALGO selections reached a maximum AUC of 0.98. The study illustrates that careful feature selection combined with robust classifiers can significantly improve the reliable identification of toxic mushrooms, which has direct implications for preventing mushroom-related health issues.

Most recently, in 2025, Singh et al. [8] focused on improving deep-learning performance in “Optimized DenseNet Architectures for Precise Classification of Edible and Poisonous Mushrooms.” The study addresses the challenge that subtle differences between edible and toxic species often lead to misclassifications, and traditional ML models struggle with feature extraction. To overcome overfitting and computational constraints, the authors proposed modifications to DenseNet-121, including freezing upper layers to preserve lower-level features, implementing dropout and weight decay regularization, and automating hyperparameter tuning via KerasTuner. Results showed that the modified DenseNet-121 achieved 0.97 for accuracy, precision, recall, and F1-score, significantly outperforming the standard DenseNet-121, which reached 0.90–0.91 across these metrics. These improvements increased reliability in distinguishing subtle species differences, reducing misclassification risks.

III. DATASET

This report employs the *Mushroom Classification* dataset sourced from the UCI Machine Learning Repository, accessed via Kaggle. [9] The dataset covers 8124 mushroom instances, each characterized by 22 categorical features describing physical attributes such as cap shape, cap color, gill size, odor, and habitat. The target variable labels each mushroom as either edible or poisonous. All input attributes are nominal and therefore require encoding prior to analysis.

IV. METHODOLOGY

A. Problem Definition

The objective of this report is to develop predictive models capable of deciding whether a given mushroom is edible or poisonous based on its observable characteristics. The task is formulated as a binary classification problem, since the target variable consists of two discrete categories.

B. Data Pre-processing

To ensure that the dataset is fit for classifier training, the following pre-processing method will be applied:

- **Encoding Categorical Attributes:** Because the dataset contains only non-numeric features, all attributes will be transformed using one-hot encoding, converting each categorical value into a set of binary indicator variables, to prevent the introduction of unintended ordinal relationships between categories.
- **Feature Removal (for Random Forest – Drop):** The two most influential feature groups, *odor* and *spore-print-color*, were removed from the dataset by dropping all corresponding one-hot encoded columns. This allows the evaluation of how model performance changes when the strongest predictors are intentionally excluded.

C. Model Selection

Two classification models were selected for this report: a Decision Tree classifier and a Random Forest classifier.

- **Decision Tree:** The Decision Tree model was chosen because of its capability to capture non-linear relationships between variables, and because its structure is inherently interpretable through rule-based splitting.
- **Random Forest (Full):** The Random Forest classifier was chosen because of its ability to improve predictive accuracy and reduce overfitting by aggregating the results of multiple decision trees.
- **Random Forest (Drop):** A second Random Forest model was included to evaluate the sensitivity of the classifier to key predictors.

D. Model Training and Evaluation

Both models will be trained on the preprocessed training set. for the Decision Tree classifier, node-splitting will be based on entropy or Gini impurity. for the Random Forest classifier,

an ensemble of decision trees will be trained on random subsets of the data and features, which is expected to enhance predictive accuracy and reduce overfitting.

Model performance will be evaluated on the held-out test set using standard classification metrics, including accuracy, precision, recall, F1-score, and the confusion matrix. These metrics will provide insight not only into overall predictive performance, but also into the models' ability to correctly identify poisonous mushrooms, a critical consideration given the practical safety implications.

V. RESULT AND DISCUSSION

A. Decision Tree

The Decision Tree classifier effectively separates edible and poisonous mushrooms by learning simple, interpretable rules from the feature set. As shown in Table I, the model achieves an overall accuracy of 0.9975, with nearly perfect precision, recall, and F1-scores for both classes. The confusion matrix in Fig. 1 confirms this performance, showing only 4 edible samples misclassified as poisonous and none of the poisonous samples predicted as edible. The feature importance plot in Fig. 2 further shows that the model relies heavily on a small set of highly discriminative features, particularly *odor* and *spore_print_color*, which explains its strong accuracy and interpretability.

TABLE I: Classification Report for Decision Tree

Class	Precision	Recall	F1-score
Edible (False)	1.00	1.00	1.00
Poisonous (True)	0.99	1.00	1.00
Accuracy	0.9975		
Macro Avg	1.00	1.00	1.00
Weighted Avg	1.00	1.00	1.00

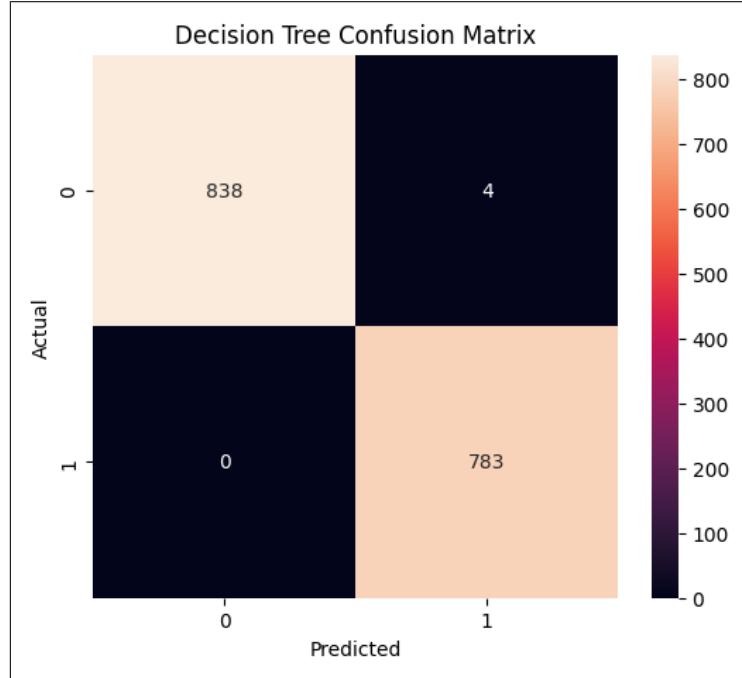


Fig. 1: Confusion matrix for Decision Tree

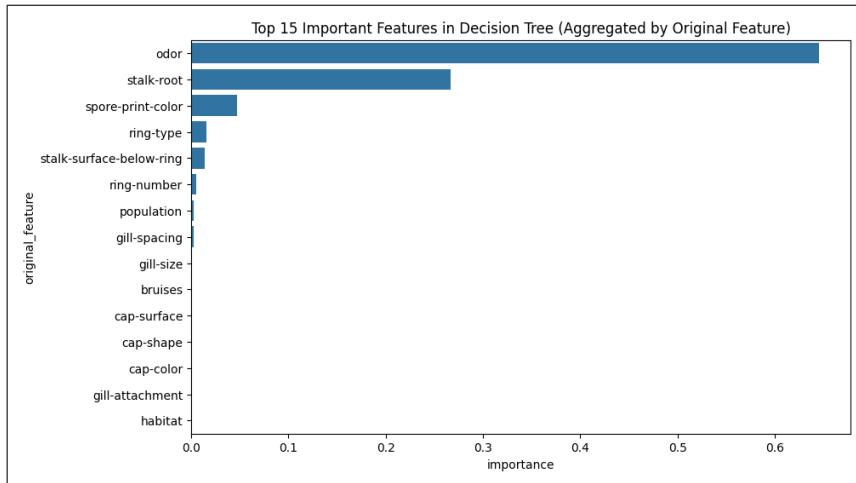


Fig. 2: Feature importance for the Decision Tree model

Finally, the Decision Tree demonstrates that even a single, well-constructed model can capture the key patterns in the dataset with high accuracy, making it both effective and easy to interpret.

B. Random Forest

The Random Forest classifier demonstrated excellent performance when trained using the complete set of mushroom attributes (RF-Full). As shown in Table II, the model achieved

an overall accuracy of 0.9951, with macro-averaged precision, recall, and F1-score all approximately 1.00. The confusion matrix in Fig. 3 illustrates near-perfect classification behavior, with almost no poisonous mushrooms misclassified as edible, a critical requirement in toxicity detection tasks. The 5-fold cross-validation results further support this stability, yielding a mean accuracy of 0.9954, indicating that the model generalizes consistently across different training partitions. Feature importance analysis in Fig. 4 highlights *odor* and *spore_print_color* as the most influential predictors, which aligns with biological domain knowledge and explains the exceptionally high accuracy of the RF-Full configuration.

TABLE II: Classification Report for Random Forest (Full Features)

Class	Precision	Recall	F1-score
Edible (False)	1.00	1.00	1.00
Poisonous (True)	1.00	0.99	1.00
Accuracy			0.9951
Macro Avg	1.00	1.00	1.00
Weighted Avg	1.00	1.00	1.00

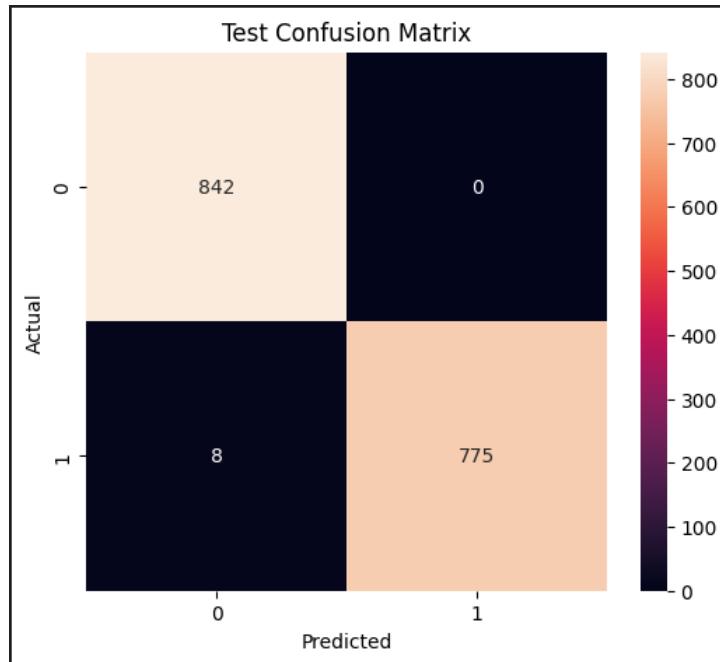


Fig. 3: Confusion matrix for Random Forest using all features (RF-Full)

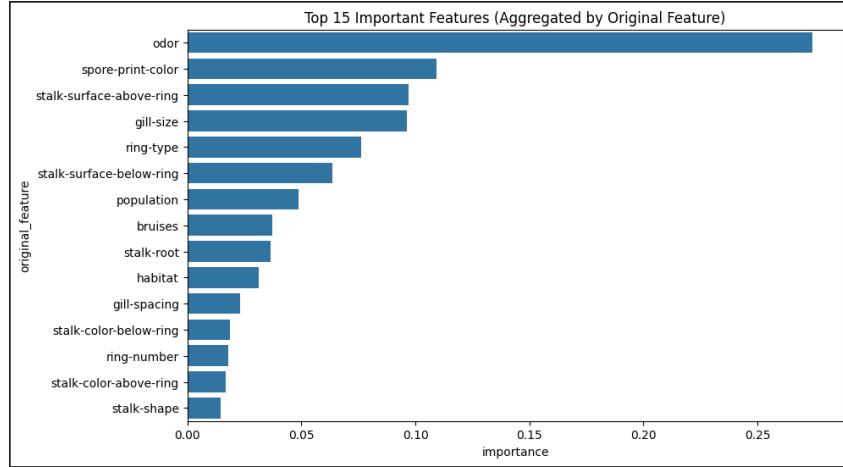


Fig. 4: Feature importance for the Random Forest model (RF-Full)

To further evaluate the robustness of the model, a second experiment was conducted in which the two highest ranking features *odor* and *spore_print_color* were removed (RF-Drop). Despite removing the dominant attributes, the Random Forest still maintained a strong performance, achieving 0.9846 accuracy, with a macro-averaged precision of 0.99, recall of 0.98, and F1-score of 0.98 (Table III). Although a small increase in false negatives was observed, as indicated in the confusion matrix in Fig. 5, the classifier continued to differentiate edible and poisonous mushrooms with high reliability. Importantly, the 5-fold cross-validation accuracy remained stable at 0.9802, confirming that the reduced-feature model retains strong generalization capability. This resilience demonstrates that Random Forest does not heavily depend on a single feature and can still learn meaningful decision boundaries even after removing top predictors.

TABLE III: Classification Report for Random Forest (After Dropping Top Features)

Class	Precision	Recall	F1-score
Edible (False)	0.98	1.00	0.99
Poisonous (True)	0.99	0.97	0.98
Accuracy	0.9846		
Macro Avg	0.99	0.98	0.98
Weighted Avg	0.99	0.98	0.98

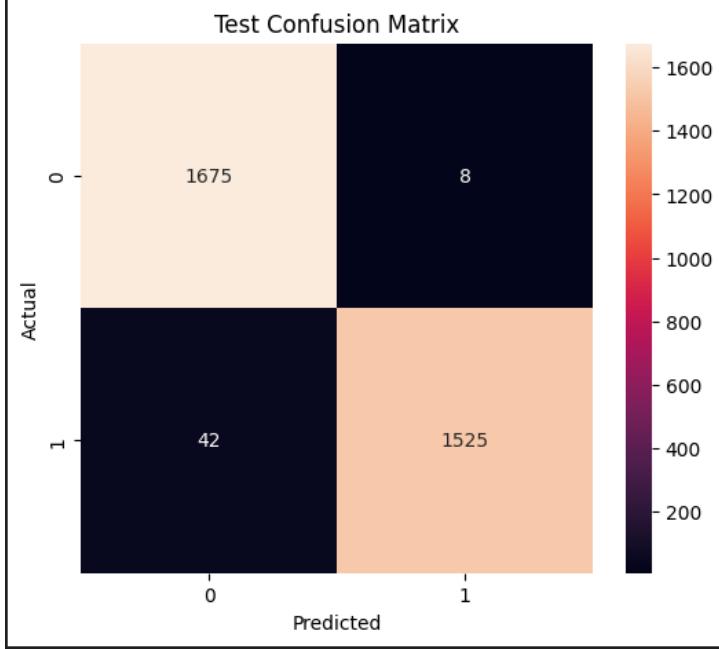


Fig. 5: Confusion matrix for Random Forest after dropping top features (RF-Drop)

Overall, the comparison between RF-Full and RF-Drop reveals that while the inclusion of odor and spore-print color significantly enhances predictive accuracy, the Random Forest model continues to perform at a high level even without them. This indicates strong robustness and feature redundancy within the dataset an important property when dealing with real-world scenarios where certain attributes may be noisy, missing, or unavailable during inference.

C. Performance Comparison of Models

Comparing the Decision Tree and Random Forest classifiers highlights both similarities and differences in their predictive behavior. As shown in Table IV, both models achieve exceptionally high accuracy, reflecting the strong signal present in the mushroom dataset. However, the Random Forest, being an ensemble of multiple decision trees, provides more stable and robust predictions, which is evident from its slightly lower variance in cross-validation scores and its more balanced feature importance distribution.

While the single Decision Tree relies heavily on a few dominant attributes, such as *odor* and *spore_print_color*, the Random Forest aggregates patterns across many trees. This allows it to incorporate additional cues like *gill-size* and *stalk-surface* more effectively, reducing the risk of overfitting. The results of the RF-Drop experiment further illustrate this point: even after removing the top two features, the Random Forest exhibits only a modest decline in accuracy, demonstrating its resilience and ability to generalize from a broader set of attributes.

All things considered, the comparison in Table IV shows that while the Decision Tree is highly interpretable and performs extremely well, the Random Forest offers stronger generalization

and robustness, making it a more reliable choice for real-world toxicity detection where missing or noisy data may occur.

TABLE IV: Comprehensive Performance Comparison: Decision Tree vs. Random Forest (Full & Drop)

Metric	Decision Tree	RF (Full Features)	RF (Drop Features)
Accuracy	99.75%	99.51%	98.46%
Most Important Feature	Odor	Odor	Distributed
Interpretability	High (White Box)		Low (Black Box)
Robustness	High		Very High

VI. CONCLUSION AND FUTURE WORK

The results demonstrate that both the Decision Tree and Random Forest are highly effective for distinguishing between edible and poisonous mushrooms, achieving near-perfect classification on this dataset with 99.75% accuracy for Decision Tree, 99.51% for Random Forest (Full), and 98.46% for Random Forest (Drop). Among the two, the Random Forest stands out as the more reliable model, offering improved robustness and better handling of feature redundancy while still maintaining interpretability through feature importance analysis. This robustness was further supported by the feature-drop experiment, where the Random Forest maintained high accuracy even after removing the most influential attributes, indicating strong generalization capability. Future extensions of this work could explore incorporating additional data sources, such as environmental factors or image-based mushroom characteristics, to evaluate whether the models generalize beyond the tabular dataset used here. Testing the classifiers on real-world or heterogeneous datasets would also help assess performance under practical conditions. Furthermore, experimenting with other ensemble methods, such as Gradient Boosting or XGBoost, could provide insight into whether even more nuanced decision boundaries lead to further improvements in predictive accuracy and reliability.

GITHUB REPOSITORY

To access the project code on GitHub, click [here](#).

REFERENCES

- [1] V. Vanitha, “Classification of mushrooms to detect their edibility based on key attributes,” *Bioscience Biotechnology Research Communications*, vol. 13, pp. 37–41, 12 2020.
- [2] K. Tank, “A comparative study on mushroom classification using supervised machine learning algorithms,” *International Journal of Trend in Scientific Research and Development (IJTSRD)*, vol. 5, no. 5, pp. 716–723, August 2021. [Online]. Available: <https://www.ijtsrd.com/papers/ijtsrd42441.pdf>
- [3] O. Tarawneh, M. Tarawneh, Y. Sharab, and M. Altarawneh, “Mushroom classification using machine-learning techniques,” 07 2022.

- [4] A. Samantara, T. Nayak, and B. Parida, "Mushroom quality prediction using machine learning classification," *International Journal of Computer Applications*, vol. 185, no. 9, pp. 14–23, 2023. [Online]. Available: <https://www.ijcaonline.org/archives/volume185/number9/samantara-2023-ijca-922746.pdf>
- [5] R. Sallam, E. Shehab, and S. Shehab, "The classification of mushroom using ml," *Kafrelsheikh Journal of Information Sciences*, vol. 4, pp. 1–7, 11 2023.
- [6] H. Ujir, I. Hipiny, M. H. Bolhassan, K. N. F. K. Azir, and S. Ali, "Automating mushroom culture classification: A machine learning approach," *International Journal of Advanced Computer Science and Applications*, vol. 15, no. 4, 2024. [Online]. Available: <http://dx.doi.org/10.14569/IJACSA.2024.0150454>
- [7] J. S. Ortiz-Letechipia, C. E. Galvan-Tejada, J. I. Galván-Tejada, M. A. Soto-Murillo, E. Acosta-Cruz, H. Gamboa-Rosales, J. M. Celaya Padilla, and H. Luna-García, "Classification and selection of the main features for the identification of toxicity in agaricus and lepiota with machine learning algorithms," *PeerJ*, vol. 12, p. e16501, 2024. [Online]. Available: <https://doi.org/10.7717/peerj.16501>
- [8] J. P. Singh, D. Ghosh, J. Singh, A. Bhattacharjee, and M. K. Gourisaria, "Optimized densenet architectures for precise classification of edible and poisonous mushrooms," *International Journal of Computational Intelligence Systems*, vol. 18, no. 1, p. 143, June 2025. [Online]. Available: <https://doi.org/10.1007/s44196-025-00871-y>
- [9] U. M. L. Repository, "Mushroom classification," <https://www.kaggle.com/datasets/uciml/mushroom-classification>, 2016.