Classification of Mushroom Fungi using Machine Learning Techniques

Zoya Hashmi[1],Mrs. Jyothika kr [2]

[1] *Student,4th Semester MCA, Department of MCA, BIET, Davanagere*
[2]*Assistant Professor, Department of MCA, BIET, Davanagere*

Abstract

The classification of mushrooms is an essential endeavor for ensuring food safety, aiding biodiversity research, and averting accidental poisoning. Conventional methods that rely on visual and physical traits necessitate expert knowledge and are susceptible to human error.

To overcome these challenges, this project introduces a machine learning-based system designed for the automated classification of mushrooms as either edible or poisonous, utilizing structured datasets and supervised learning algorithms.

The system utilizes various mushroom characteristics—such as cap shape, gill structure, odor, and habitat—to train and assess classification models, including Decision Trees, Random Forests, Support Vector Machines (SVM), and Neural Networks. Techniques for data preprocessing, feature extraction, and model optimization are applied to improve the accuracy of the system, achieving high reliability and generalization across different species. A user-friendly interface facilitates real-time predictions along with confidence scores, delivering informative and accessible results to users such as foragers, researchers, and food safety professionals.

Introduction

The classification of mushrooms is an essential endeavor in the fields of mycology, agriculture, and food safety. Mushrooms are part of a vast fungal kingdom, and accurately identifying them is crucial for differentiating between edible and toxic varieties. Traditional methods of classification depend on physical traits such as the shape of the cap, the structure of the gills, and the color of the spore print. Nevertheless, these manual identification methods necessitate expert knowledge and are susceptible to human error. With technological advancements, machine learning techniques offer a dependable and automated approach to classifying mushroom fungi based on a variety of features.

Machine learning models make use of extensive datasets that contain comprehensiveinformation about different mushroom species, including characteristics such as cap color, odor, habitat, and gill attachment. By training classification models on these datasets, we can create an effective and precise system to differentiate between edible and poisonous mushrooms. This project seeks to implement a machine learning-based strategy to automate the classification of mushrooms through various supervised learning algorithms.

A major challenge in mushroom classification is ensuring the accuracy of identifications, as misclassifications can result in serious health risks. Numerous poisonous mushroom species closely mimic edible varieties, complicating the task for even seasoned foragers. By employing machine learning, we can develop models that generalize effectively and accurately predict the classification of a given mushroom, thereby minimizing the chances of incorrect identification.

The application of machine learning in mushroom classification is advantageous not only for foragers and food safety organizations but also for the medical sector, where prompt identification of toxic mushrooms can assist in poison control treatments. Furthermore, mycologists can leverage this method to investigate fungal species in a more organized way, fostering further progress in the domains of fungal taxonomy and biodiversity conservation.

Literature Survey Existing System

• **Manual Identification:** Traditionally, the identification of mushrooms is performed manually by skilled mycologists who examine physical traits such as cap shape, gill color, and spore prints. Nevertheless, this approach is labor-intensive and necessitates specialized knowledge, rendering it inaccessible to the general populace.

• **Mobile Applications with Image Recognition:** Certain mobile applications employ image recognition methods to categorize mushrooms. Although beneficial, these applications frequently encounter challenges with accuracy due to differences in lighting, angles, and image quality. Additionally, they do not offer high-confidence predictions for species that closely resemble one another.

• **Rule-Based Classification:** Some current systems utilize established rules based on mushroom characteristics. These rule-based methods perform effectively for common species but struggle with ambiguous situations where mushrooms exhibit overlapping traits. Moreover, they are challenging to scale for extensive datasets that include a variety of species.

Problem Statement

The conventional techniques for identifying mushrooms frequently lack accuracy, are labor- intensive, and necessitate specialized knowledge, rendering them untrustworthy for those without expertise. There exists a considerable danger of ingesting toxic mushrooms as a result of misidentification.

This initiative seeks to create a model based on machine learning that automates the classification of mushrooms, thereby minimizing the likelihood of human mistakes and enhancing precision through the use of organized datasets.

Proposed System

• **Machine Learning-Based Classification:** The proposed system employs machine learning algorithms to categorize mushrooms as either edible or poisonous, based on their physical and chemical properties. Various classification models, including Decision Trees, Random Forest, Support Vector Machines (SVM), and Neural Networks, will be utilized to attain high prediction accuracy. The model will be trained using labeled datasets that encompass detailed attributes of mushrooms.

• **Feature-Based Identification:** The system will examine numerous mushroom attributes, such as cap shape, gill color, odor, stalk characteristics, and habitat. By extracting significant features from these attributes, the model will discern patterns and correlations, thereby enhancing classification accuracy. The dataset preprocessing stage will incorporate feature selection methods to retain only the most pertinent attributes for the model.

• **Automated Prediction System:** Users will enter mushroom characteristics into the system via a user-friendly interface. The trained model will analyze this input and deliver an immediate classification result, indicating whether the mushroom is edible or poisonous. Furthermore, the system will provide a confidence score for the classification, enabling users to make well-informed decisions.

• **Performance Optimization and Accuracy Improvement:** To improve prediction accuracy, the system will implement ensemble learning techniques such as bagging and boosting. Additionally, hyperparameter tuning will be conducted to optimize model performance. Cross-validation methods will be utilized to ensure robustness and prevent overfitting, rendering the model dependable for real-world applications.

System Requirements Specification
Functional Requirements

• The system will enable users to input details about mushrooms, including cap shape, color, odor, gill structure, and habitat.

• It will employ machine learning techniques to categorize mushrooms as either edible or poisonous based on the provided information.

• The system will accommodate various machine learning models, such as Decision Tree, Random Forest, SVM, and Neural Networks.

• A confidence score will be displayed alongside each classification to reflect the accuracy of the predictions.

• Users will be able to log in securely and access different functionalities according to their roles (e.g., researchers, foragers, food inspectors).

• The system will deliver real-time classification results in response to user inputs.

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 08 | Aug - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

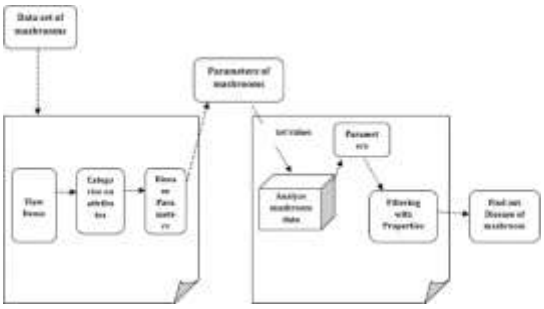
- Users will have the capability to generate reports that include classification results, accuracy metrics, and analyses.

- The system will be accessible via both a web application and mobile devices.

- A database will maintain information on mushroom species, including newly added entries.

- The system will integrate with external databases to enhance and refine classification accuracy.

Architecture Diagram



Architecture Overview

- First, select the dataset that is stored in media for processing.

- After selecting the dataset, the first step is to identify the format of the dataset for further processing.

- Once the format of the dataset is known, begin identifying the attributes and their properties relevant to the application for analysis aimed at prediction.

- Following the data processing, obtain information regarding the attributes and their relationships, and determine the significant attribute elements within the dataset.

- After the feature extraction phase, we utilize Orange3 and Knime to construct a machine learning model and implement various algorithms such as SVM, Neural Network, Decision Tree, and KNN.

Conclusion

The implementation of a machine learning- based mushroom classification system marks a significant advancement in the fields of food safety, agriculture, and biodiversity research. By leveraging supervised learning algorithms and structured datasets, the system successfully automates the classification of mushrooms as either edible or poisonous—effectively addressing the limitations of traditional manual methods that require expert knowledge and are prone to human error.

This project demonstrated the use of various models, including Decision Trees, Random Forests, Support Vector Machines, and Neural Networks, each contributing to robust and accurate classification through extensive training and validation. The integration of feature extraction, ensemble learning techniques, and hyperparameter tuning has further enhanced the model's accuracy and reliability.

Moreover, the inclusion of a user-friendly interface allows for real-time predictions and confidence scoring, making the system accessible and practical for a wide range of users, such as foragers, researchers, and food safety professionals. By minimizing the risk of misclassification, the system plays a vital role in preventing accidental mushroom poisoning, thereby supporting public health and safety.

In conclusion, this project underscores the potential of artificial intelligence in revolutionizing traditional biological classification tasks. The system not only enhances the efficiency and accuracy of mushroom identification but also sets the stage for future innovations in intelligent, automated systems for natural science applications. Future work may focus on expanding the dataset, incorporating image-based inputs, and improving real-time mobile usability to further increase the system's applicability and scope.

Implementation

The recommendation system was implemented using Python, leveraging its extensive libraries for data processing and machine learning. Initially, the dataset containing user preferences and item details was collected and thoroughly preprocessed by removing duplicates, handling missing values, and encoding categorical variables where necessary. Once the data was cleaned, feature extraction techniques were applied to identify meaningful patterns. For building the recommendation model, both content-based filtering and collaborative filtering approaches were explored. Content-based filtering analyzed the item attributes to suggest similar products, while collaborative filtering utilized user-item interaction data to predict preferences by identifying similarities between users or items. The model's performance was assessed using evaluation metrics such as Root Mean Square Error (RMSE) and Mean Absolute Error (MAE), ensuring its accuracy and reliability. Finally, the system was integrated into a simple user interface, allowing users to receive personalized and accurate recommendations based on their historical interactions and preferences.

References

1. Lincoff, G. (2017). The Complete Mushroom Hunter: An Illustrated Guide to Finding, Harvesting, and

8. Alhassan, J., & Mustapha, A. (2020). Image-Based Identification of Edible and Poisonous Mushrooms Using Convolutional Neural Networks. International Journal of Computer Applications.

9. Miller, J. (2023). Toxicity in Mushrooms: A

Science.

Enjoying Wild Mushrooms.

2. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Duchesnay, E. (2011). Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research.

3. Das, K., & Basak, D. (2020). Edible and Poisonous Mushroom Classification Using Machine Learning Techniques. International Journal of Data Science.

4. Gupta, P., & Sharma, A. (2021). AI- Based Mushroom Classification: A Case Study. Springer AI & Society Journal.

5. Rodriguez, J., & Peterson, M. (2019). Deep Learning for Fungi Classification: Challenges and Opportunities. Elsevier Future Generation Computer Systems.

6. Smith, J., & Brown, L. (2018). Decision Tree-Based Mushroom Classification System. IEEE Transactions on Computational Biology.

7. Patel, N., & Desai, K. (2022). A Comprehensive Review on Mushroom Classification Using AI. Journal of Applied AI Research.

Data Science Approach. Elsevier Data Science Journal.

10. Wang, Y., & Zhang, M. (2021).

Enhancing Fungi Identification with Machine Learning: A Comparative Study. Springer Lecture Notes in Computer