# MUSHROOM CLASSIFICATION (Machine Learning)

## DETAILED PROJECT REPORT

Project Member: Sonali Bhadra

iNeuron.ai

January 2023

**1. Introduction:** The project aimed to develop a machine-learning model for the classification of mushrooms as edible or poisonous based on a comprehensive dataset. Python, with libraries like Pandas, NumPy, Matplotlib, Seaborn, and Scikit-Learn, was utilized for data exploration, preprocessing, and model development.

## 2. Dataset Overview:

The dataset consisted of 8124 instances and 23 categorical features, describing various characteristics such as cap shape, color, odor, and habitat. Before analysis, an in-depth understanding of feature values was crucial for interpreting the data correctly.

## 3. Exploratory Data Analysis (EDA):

EDA was performed to gain insights into the dataset's distribution. Visualizations were created to illustrate the categorical distributions of features, and the summary statistics provided a snapshot of the data's central tendencies.

## 4. Data Preprocessing:

To prepare the data for machine learning algorithms, categorical values were encoded into a numerical format using Label Encoding. This conversion facilitated the application of algorithms requiring numerical input.

## 5. Principal Component Analysis (PCA):

To handle the dataset's high dimensionality, PCA was employed for feature reduction. Seven principal components were selected, explaining a significant portion of the variance, and used for model training.

## 6. Model Selection and Training:

Various classifiers were trained, including Logistic Regression, K-Nearest Neighbors, Support Vector Classifier, Decision Tree, Random Forest, and Gradient Boosting. The models were evaluated based on accuracy metrics, and the Random Forest Classifier emerged as the top-performing model.

## 7. Graphical User Interface (GUI):

To enhance user interaction, a Tkinter-based GUI was implemented, allowing users to input mushroom characteristics for real-time predictions. This feature adds practicality to the model for users without programming knowledge.

## 8. Model Deployment:

The Random Forest Classifier was selected as the final model and saved for future use. The model demonstrated its effectiveness in predicting whether a mushroom is edible or poisonous.

## 9. Conclusion:

The project successfully demonstrated the application of machine learning in mushroom classification. The trained model, equipped with a user-friendly GUI, provides a valuable tool for quick and reliable mushroom assessment. Future enhancements could involve integrating additional features or exploring advanced deep-learning techniques for further accuracy improvements.

## 10. Acknowledgments:

Special thanks to the open-source community and contributors of the libraries used in this project, making advanced machine learning accessible for diverse applications.

This comprehensive project report encapsulates the journey from data exploration to model deployment, highlighting the significance of mushroom classification in ensuring food safety.