CLASSIFICATION OF MUSHROOM'S INTO EDIBLE OR POISONOUS USING ML

Project report submitted in partial fulfillment of the requirement for the degree of

BACHELOR OF TECHNOLOGY

IN

ELECTRONICS AND COMMUNICATION ENGINEERING

By

Nikhil Thakur (181012)
Himanshu Thakur (181022)

UNDER THE GUIDANCE OF

Dr. Alok Kumar

JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY, WAKNAGHAT May 2022

TABLE OF CONTENTS

DECLARATION

We hereby declare that the work reported in the B.Tech Project Report entitled **"Classification Of Mushroom's Into Edible Or Poisonous Using ML"** submitted at **Jaypee University of Information Technology,Waknaghat, India** is an authentic record of our work carried out under the supervision of **Dr. Alok Kumar .** We have not submitted this work elsewhere for any other degree or diploma.

Nikhil Thakur (181012) Himanshu Thakur (181022)

This is to certify that the above statement made by the candidates is correct to the best of my knowledge.

Dr. Alok Kumar
Date:

Head of the Department/Project Coordinator

ACKNOWLEDGEMENT

Knowledge, energy and time are the resources in the completion of this project but the most requisite is the proper guidance. We would like to express our deepest appreciation to all those who have been helping us throughout the project and without whom this project would have been a very difficult task. We are highly indebted to our respected project guide, **Dr. Alok Kumar** for his valuable suggestions, ideas and constant supervision as well as for providing necessary information throughout the project.

We want to express gratitude towards **Prof. Dr. Rajiv Kumar** , Head of the Department, Department of ECE, Jaypee University of Information Technology, Waknaghat for providing us this opportunity to make our project.

We would also like to thank all the panel members for their valuable suggestions in all the viva voce during evaluation from which we got to learn a lot.

No thanks can counter our indebtedness to our parents who have been with us throughout. We thank them from the core of our hearts.

LIST OF ACRONYMS AND ABBREVIATIONS

- DT- Decision Tree
- SVM- Support Vector Machine
- GA- Genetic Algorithm
- kNN- k – Nearest Neighbours
- NN- Neural Networks
- CNN- Convolutional Neural Networks

LIST OF SYMBOLS

None

LIST OF FIGURES

ABSTRACT

One of the fungal kinds' foods with the most significant nutrients on the planet is the mushroom. Mushrooms offer several health benefits.They are also helpful in research as they can destroy cancer cells. The main idea behind this major project is to determine the best approach for classification of mushrooms. Mushrooms will be divided into two categories: toxic and edible. On a dataset of mushroom photos, the suggested method will use several approaches and algorithms, such as k Nearest Neighbors algorithm, Neural Networks algorithm, Decision Tree algorithm, and Support Vector Machines algorithm on images with and without backgrounds.

The most promising approach for categorising mushroom photographs, according to the experimental findings, is kNN, which has an accuracy of 95 percent which we based on many features derived from the images of mushrooms with genuine dimensions of mushrooms, and 88 percent based on features retrieved with the help of image features solely.

CHAPTER 1
INTRODUCTION

The fungal fruiting body is referred to as a mushroom. The stem and cap of the prototype mushroom are anatomically similar, having gills on the underside. However, the word can apply to a variety of stemless gilled fungus. The spores produced by the mushroom's gills contribute to the proliferation of the fungus. For millennia, mushroom toxicity has been recognised, and it has been linked to the deaths of countless historical leaders, including Claudius, a mighty Roman Emperor.

Although many cases are deliberate ingestions, mushroom poisoning can occur when a forager misidentifies a toxic species as edible. Mushroom poisoning symptoms can range from mild gastrointestinal distress to potentially fatal indications such as liver failure, renal failure, and neurologic complications. There are up to 14 syndromes that appear differently depending on the species, poisons, and amount consumed.

One of the world's most serious health issues is mushroom poisoning. More than 5000+ mushroom species are thought to exist in the world. Only 25% to 30% of mushrooms have been identified, and only 3% of them are dangerous. In certain nations and places, mushrooms are regularly used as food, whereas in others, they are regarded with suspicion.

In Asian countries like India, Sri Lanka, wild mushroom picking is common. Mushroom gathering and eating are, however, uncommon in countries like the United Kingdom. In the United States, "mushrooming" is quite popular. In most areas, mushroom hunting expeditions are organised by groups that emphasise proper identification and safety.

Aside from intentional use of hallucinogenic mushrooms, the majority of mushroom poisonings are unintentional and result from species misidentification [1]. Suicide attempts or criminal acts with mushrooms are quite uncommon. The quantity of toxicity varies depending on the mushroom's age, season, geographical region, and cooking method [1]. While one individual may get poisoning symptoms after eating the same mushroom, another may not.

Fig 1.1 Wild Mushroom

Mushrooms are commonly used in meals for their delicious flavour and also because of their health benefits. They contain very important vitamins as well as minerals and make a fantastic complement to your diet, providing flavour to a variety of cuisines.

Crimini mushrooms are a common mushroom kind found in kitchens all over the globe. Many people are unaware that mushrooms, particularly crimini mushrooms, are a kind of fungus. They are indigenous to North America and Europe and are prized for their delicate taste and meaty texture.

1.1 Health Advantages:-

Mushrooms are a very low-calorie meal that is very high in nutrients. They've long been acknowledged as a crucial element of any meal due to their high concentration of health-promoting vitamins, minerals, and antioxidants. For example, mushrooms grown in the presence of UV light are high in Vitamin D, which is essential for bone and immunological health.

Crimini mushrooms are a very good source of zinc, an important trace element. Zinc is an essential vitamin for the immune system as well as optimum development in newborns and children.

Furthermore, researchers have discovered a number of additional compelling reasons to include mushrooms in your diet, including:

1.1.1. Reduce Blood Pressure

Mushrooms are abundant in potassium, a nutrient that has been demonstrated to help the body cope with the negative effects of salt. Potassium also helps to lower blood pressure by reducing blood vessel tightness.

1.1.2. Boost Your Immune System

The anti-inflammatory effects of mushrooms have been discovered to greatly improve immune system performance. According to studies, mushrooms stimulate macrophages in the innate immunity, which improves the immune system's ability to combat foreign bodies and makes you less susceptible to serious illnesses.

1.1.3. Loss of Weight

Long-term and short-term studies have both demonstrated that mushrooms, when combined with exercise and other lifestyle adjustments, may help people lose weight. For example, when asked to replace 20% of their beef intake with mushrooms, study participants saw improvements in their BMI and belly circumference. Mushroom antioxidants are also suggested to lower the risk of hypertension and other metabolic diseases.

CHAPTER 2
MUSHROOM TOXICITY

Poisoning from mushrooms, as well as other types of poisoning, adds to the country's high morbidity and death rates. Mushrooms are an essential part of the cuisine of several Indian ethnic groups. In recent years, the prevalence of mushroom poisoning in India has been acknowledged as a result of increased awareness and afflicted persons seeking medical help as soon as possible. The country's tropical belt, with its variety, is home to a plethora of fungal mushroom species. According to research performed in India, there are 1200 species with just 50 to 100 dangerous species. Twelve different mushroom toxins have been identified as being responsible for 14 different clinical disorders.

According to the research, 50-100 poisonous mushrooms create mycotoxins that are responsible for various clinical symptoms, which are mentioned below. The North American Mycological Association maintains a case register in the United States where cases of mushroom poisoning are recorded. Mushrooms are often consumed in India, particularly by ethnic groups.

The majority of mushroom poisoning instances result from the ingestion of wild mushrooms. However, the majority of cases go untreated and unreported, and outbreaks of mushroom poisoning are reported in the news, mainly during the monsoon season. Table 1 lists the various mushroom species, as well as their poisons and fatality rate.

According to case reports in India, the recorded cases are from tribal communities in South India, the Eastern Ghats, Northern India, and the country's north eastern region. The majority of the incidents are recorded and documented in newspapers, where the death rate of such cases is noted. The clinical symptoms of the majority of the deaths are highly indicative of Amanita phalloides poisoning.

Poisonous mushrooms contain a range of poisons with varying degrees of toxicity. Mushroom poisoning may cause anything from stomach distress to organ failure and death. Serious symptoms do not usually appear right after eating; sometimes, they do not appear until the toxin assaults the kidney or liver, which might take days or weeks.

Toxin	Toxicity
Alpha-amanitin	Deadly
Phallotoxin	Non-lethal
Orellanine	Deadly
Muscarine	Potentially deadly
Monomethylhydrazine (MMH)	Deadly
Coprine	Non-lethal
Ibotenic acid	Potentially deadly
Muscimol	Potentially deadly
Arabitol	Non-lethal
Bolesatine	Non-lethal
Ergotamine	Deadly

Figure 2.1: Toxins In Mushrooms

The most frequent symptom of mushroom poisoning is stomach discomfort. The majority of "poisonous" mushrooms include gastrointestinal irritants that induce vomiting and diarrhoea (occasionally necessitating hospitalisation), but little long-term harm. There are, however, a number of well-known mushroom toxins that have distinct, and sometimes fatal, effects:

Figure 2.2: Amanita phalloides

The fungus Amanita phalloides (shown in Fig 2.1) is responsible for the majority of deadly mushroom poisonings across the globe.

These poisonous mushrooms mimic numerous edible species regularly eaten by humans (most notably the Caesar's mushroom and the straw mushroom), increasing the potential of accidental poisoning. Amatoxins, the kind of toxin present in these mushrooms, are thermostable, meaning they don't alter when exposed to heat, hence cooking has no impact on their poisonous effects.

One of the most extensively farmed and popular mushrooms in the planet is Amanita phalloides.

Figure 2.3: Amanita phalloides

Separating edible from dangerous species takes great attention to detail; there is no one property that can be used to identify all deadly mushrooms, and there is no single trait that can be used to identify all edible mushrooms. Mycologists are those who gather mushrooms for the purpose of eating them, and mushroom hunting, or simply "mushrooming," is the act of doing so. Even edible mushrooms may cause allergic responses in those who are sensitive to them, ranging from mild asthma to severe anaphylactic shock. Small levels of hydrazines, the most prevalent of which is agaritine, may be found in even grown A. bisporus (a mycotoxin and carcinogen). When cooked, however, moderate heat destroys the hydrazine.

CHAPTER 3
IDENTIFICATION OF MUSHROOMS

The size, colour, and form of the cap and stem; whether the underside of the cap contains pores, gills, or teeth; the lack or presence of a veil; and the colour of the mushroom and its meat are all diagnostic traits used to identify mushrooms. The colour of the skin after it has been injured may also reveal information about its identity. It's vital to observe where the fungus develops; some have mycorrhizal relationships with certain tree species, while others inhabit decaying logs or live trees. The odour of several fungus is peculiar. When particular chemicals are given to the surface, meat, or spores of some mushrooms, they change colour. Ammonia, potassium hydroxide, and iron salts are examples. When iodine is added to certain spores, they change colour.

3.1 Parts that are Microscopic:-

Pileus/cap- is the mushroom's major horizontal component with lamellae/gills or tubes on the underside. It may come in a variety of forms, sizes, and colours (including overall and cross-sectional margin appearances).

Hymenium:- is the spore-bearing layer of a fungus' fruiting body that contains asci or basidia. Lamellae or tubes might be used to accomplish this. The hymenium's attachment to the pileus and stipe varies per species, with a variety of lamellae/gill and tube attachment morphologies seen in cross section. The distance between the lamellae and the gills may also vary.

The pileus has platelike features on the bottom called lamellae/gills. They are distributed radially from the stipe to the pileus's border. On the lamellae, spore-bearing basidia may be observed.

Tubes:- In place of gills, certain mushrooms (genus Boletus) feature a thick, sponge-like tube layer with holes that open to the pileus' underside.

Most mushrooms have a stalk or stipe that may be connected to the centre, off-center, or side of a pileus. The stipe may vary in form, size, and colour depending on the fungus species. The stalk may be nonexistent or extremely tiny in certain cases.

Veils:- Veils are membranes that cover and protect the juvenile mushroom or immature lamellae in certain mushrooms. Veils are divided into two categories.

The partial veil, which protects young lamellae, and the universal veil, which covers the juvenile mushroom, are both membranes. These membranes break down as the mushroom grows, occasionally leaving fragments on the stipe or pileus.

The stipe of a mushroom has a ring-like structure called an annulus. A sliver of the partial curtain may be seen here.

Volva:- It is a scaly relic of the global veil that may be found at the foot of certain mushroom stripes as a cup-like structure.

Only a few mushrooms have cortina, a silky, cobweb-like partial veil (genus Cortinarius).

3.2 Macroscopic Components:-

Hyphae:- is one of the threadlike strands that make up the fungus' mycelium. It is made up of single and double nucleus cells. The septum is a wall-like structure that divides each cell. A clamp connection is a loop-like structure present at the septum of certain fungal hyphae. The vegetative phase of a fungus is made up of mycelium, which is a mass of hyphae.

Ascus:- It is a fungus that belongs to the genus Ascus. Ascomycetes generate spores in an ascus, a tube-shaped container. An ascus typically contains 8 spores, which are distributed via a hole at the ascus' apex.

Basidia:- Basidiomycetes fungi generate spores on small appendages called sterigmata that protrude from a club-shaped basidia. Basidia usually have four spores, although there are situations when just two spores are present.

Spore:- a fungus's reproductive structure that is walled, single or multi celled, and capable of producing a new individual. Spores in fungus may take on a range of shapes and sizes.

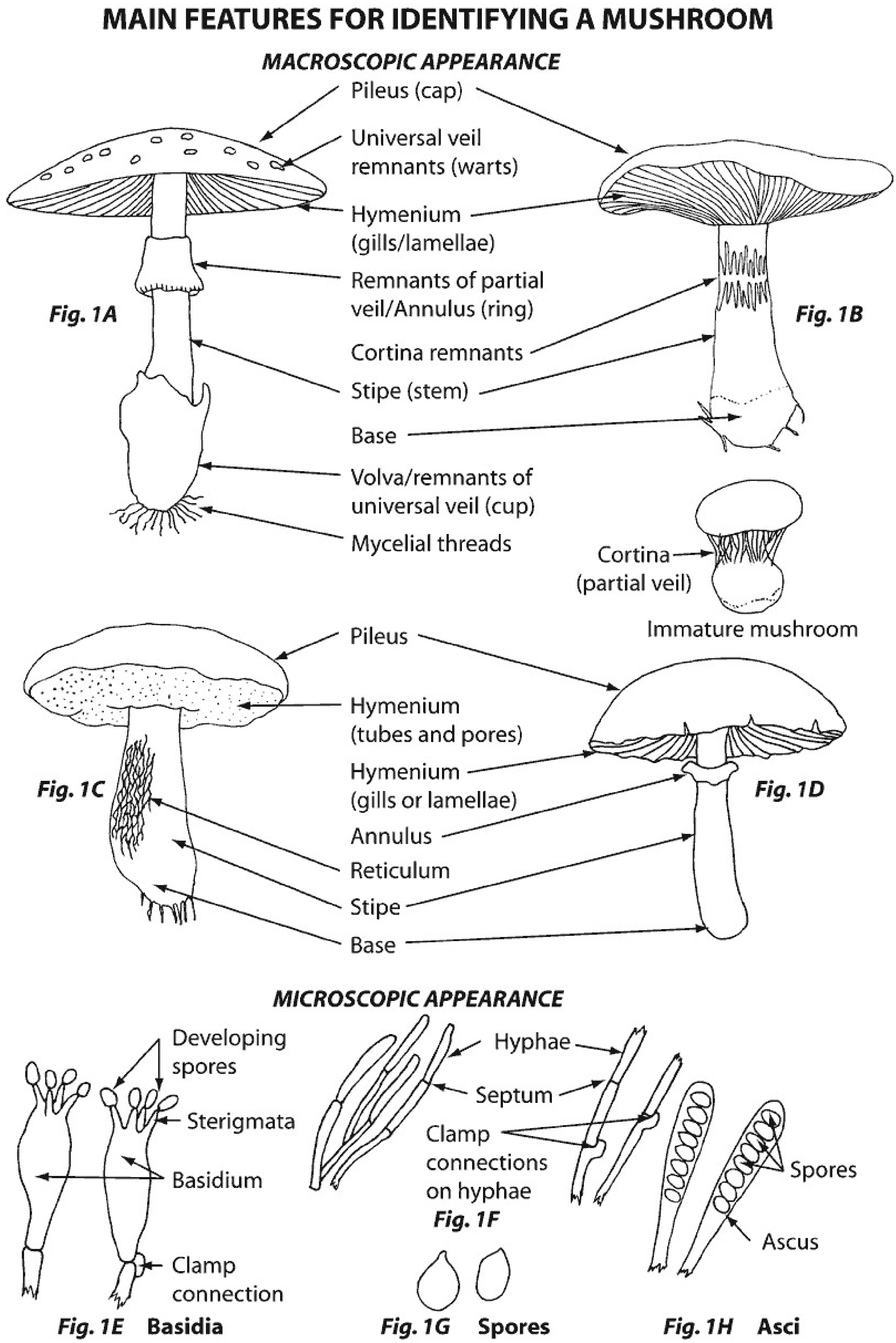


Figure 3.1: Diagram of Identifying Features of Mushroom

CHAPTER 4
LITERATURE REVIEW

There have been a variety of studies that have utilised various methodologies to classify mushrooms. [3] have designed a MDAS or Mushroom Diagnosis Assistance System, which consists of following components: a server to run the application online, a database that is preprocessed, and a client for smartphones like a app for usage on smartphones.

Naive Bayes and Classification Tree algorithms are used to identify the mushroom types. To commence, the suggested method picks out the very well known mushroom properties. Second, determine the mushroom's species. The Decision Tree algorithm beats the Naïve Bayes algorithm in terms of correctly and incorrect classified instances, as well as failure measurements.

Kumar and colleagues looked at a variety of classification algorithms used in data analysis for prediction systems in [9]. Three decision tree approaches, comprising a statistical, a multilayer perceptron, a svm classifiers, and a classification algorithm, are compared. Using four databases from distinct fields, the suggested algorithm analyses projected accuracy, failure rate, narrative coherence, categorization index, and training length..

The Training Algorithm (GA) and svms algorithms surpassed everyone else in terms of projected accuracy, according to the findings of the studies. In decisions tree-based techniques, the QUESTT technique creates trees with reduced size and depth. Finally, for their information systems, the GA-based approach is the most successful algorithm.

Babu and colleagues presented a new potential application for SVM in [10]. The suggested technique uses the Support Vector and Bayesian Network methods to classify mushrooms. The testing shown that SVM surpasses Nave Bayer's technique in terms of accuracy. Furthermore, the SVM is a versatile method which can be used to a wide range of problems.

[2] developed a categorization prediction system using Multi-Layer Perceptions for Dataset Development. In the experiment, only 8264 of the database is used for training. According to the trial findings, the best processing layer is 3, the best learning rates are 0.7, the best input layer is sigmoid, the best activation rate is 0.3, and the best epoch result is 400.

Onudu developed an improved K-mean strategy compared with the conventional k-mean technique to enhance classified database categorization while addressing the basic problem with the classic classification model in [11].

The approach proposed is based on the Eucledean distance metric. The data collection was transformed into numeric values using the specified technique. The programme then reads the data and normalises the numeric properties to prevent a large range of values. The results of the trial revealed that the proposed modified K-means approaches were quicker than the current algorithm.

Al-mejibla and Hemad [1] created a Mushroom Feature Diagnosis Assistance Software application that can be used on a smartphone or a web browser. The purpose of this programme is to guarantee that mushroom collecting is done safely. They used regression trees and naive bayes algorithms to categorise the mushrooms.

They used the very well known mushroom traits to determine the mushroom kind. This methodology includes two basic different stages: training and selecting, which are used to assign the much more active qualities in the selection procedure and determine the final result.

The experimental findings showed that decision trees beat naive bays in terms of error measurements, correctly classified samples, and misclassified samples. To review a previous mushroom data gathering, the writers of [12] employed numerous machine learning approaches including the Weka miner tool.

A voted perceptron approach, a nearest neighbour classifier, a covering algorithm to collect correct rules, an unpruned decision tree, and a nearest neighbour classifier were all used.

After testing the techniques on many groups of trees, they concluded that an unpruned tree produces the most accurate results, and they used it in a biological interface based on the interactive web to generate engaging mushroom recognition.
In [13], Chowdhory and S. Oijha discovered a technique for distinguishing many mushroom diseases using various data mining classification approaches. They employed data processing techniques such as Nave Bayes, RIGOR, and SME to analyse a dataset acquired from a mushroom farm.

They made a comparison based on a statistical method for detecting common mushroom symptoms in order to uncover mushroom sickness. When compared to other classification algorithms, they discovered that naive Bayes produces the best results.

In [14], Beniwel and Des employed data processing and mining classification algorithms such as Zero, Naive Bayes, and Bayes net to evaluate a mushroom dataset including both lethal and nonpoisonous mushrooms.

They implied the accuracy, kappa statistic, and mean absolute error to assess classification algorithms. They found out that the Bayes net, followed by Naive Bayes, delivers the lowest mean absolute error and maximum accuracy.

CHAPTER 5
METHODOLOGY

The purpose of this project is to analyse Mushroom images and classify them into 2 categories using machine learning approaches (poisonous and nonpoisonous).

5.1 Research Phase:-

In this study, we used a five-phase methodology: the first step is data collection, the second stage is preprocessing the data, the third stage is data acquisition, then ML model, and the last phase is evaluation.

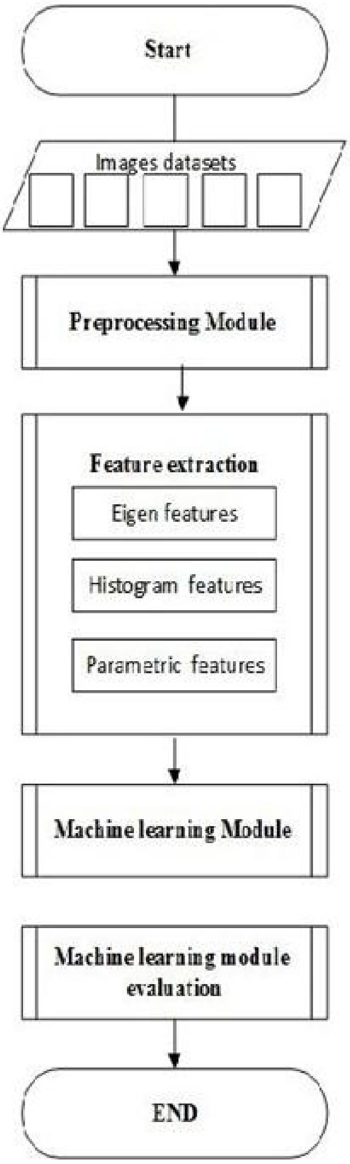


Figure 5.1: Depicts the suggested approach's study steps.

5.2 Data Set Collection:-

We began by gathering our mushroom picture collection (raw dataset) from [15], which is divided into three groups (edible, inedible, and poisonous).
Aside from photos, each mushroom type is identified by other facts, such as family, location, size, and edibility. Figure 2 shows an example of mushroom visualisations:

Figure 5.2: Mushroom Images Dataset

5.3 Highlights Extracting:-

We used Labview software to obtain all of the attributes from the original dataset's gathered pics during this stage. We initially extract the Eigen parameters after scaling each image. Second, we examine the top 100 most strong features. The dimension characteristics associated with different types in the database, such as cap girth, stem tallness, and diameter, is included into the feature matrix for the dataset. Finally, we generate the feature map, which comprises each of the measurements with Eigen properties such as cap girth, stem height, and diameter, in order to develop the ML Model.The suggested technique uses a range of ways to create the ML model, including NN, DT, SVM, and KNN. With a number of folds of 11 and an accuracy of 96 percent, the best results were achieved using KNN with cross validation. Despite the challenges of collecting high accuracy for mushrooms, we chose to rely only on photo-derived properties.

Figure 6 shows how we employ identifying edges in grayscale images mode of images to determine the size and shape of the mushroom shape inside the pictures to enhance outcomes. Next add these qualities to the Eigen properties. KNN yielded an accuracy of 84 percent, according to the trial's conclusions.

To improve the findings, we sought to extract various information from photos, such as histogram features. To compute each of height and breadth according to identify edges, we repeated the processes from the previous experiment. For the dataset, we created a new features matrix. The accuracy of histogram characteristics was found to be 87 percent in the experiment. To improve our findings, we created an algorithm that seeks to extract additional information, which we refer to as parametric features.

Local Contrast Data normalization (LCN): Influenced by cognitive neuroscience [16], it is used to differentiate features within a feature space as well as among feature maps at the same location.

Standard deviation, Sample variance, and Value of skewness are statistical meta qualities that are obtained by assessing a statistical idea, computing it for all numeric qualities, and computing the mean [17].

Entropy is calculated as the average amount of information generated by a randomized source for information [18]. It has a long record and has been the subject of various reconstructions and interpretations.

The word "mean" comes in helpful for estimating predicted losses and benefits. For instance, in the recommended approach, we applied mean to identify in the characteristics matrix, notably for computing image length and width.

Correlation: The term "correlation" relates to a reciprocal relationship or association between objects [19], and it is among the most regularly used.

Homogeneity describes the process of extracting the same set of attributes through all member nodes, and is one of the key categories of information mining.

The real or hypothetical diameter of a diameter of the mushroom stalk is termed diameter.

5.3.1 Features For Model Training:-

Figure 5.3: Mushroom gill tissue arrangement

Figure 5.4: Mushroom ring type

Figure 5.5: Mushroom stalk

Figure 5.6: Mushroom gill spacing

\

Figure 5.7: Mushroom gill attachment

Figure 5.8: Mushroom cap surface

Figure 5.9: Mushroom cap shape

5.4 Noise Reduction:-

Noise reduction is a key aspect that determines picture quality, and it works to eliminate image defects that cause difficulties. We will apply noise reduction in the suggested technique to eliminate unusable areas from original photos, such as the backdrop of photographs.

5.5 Extraction Images Without Background:-

We employed Matlab to collect Eigen characteristics for updated photographs (i.e. images missing backdrops) and reconstruct the features array. We use low contrast identification to construct a new dataset by determining the height and width of each photograph depending on the corners of the mushroom pictures. We tried the Cnn, DT, Svc, and Nearest neighbor techniques on Orange 3 software, with KNN yielding the best outcomes with an efficiency of 82%.

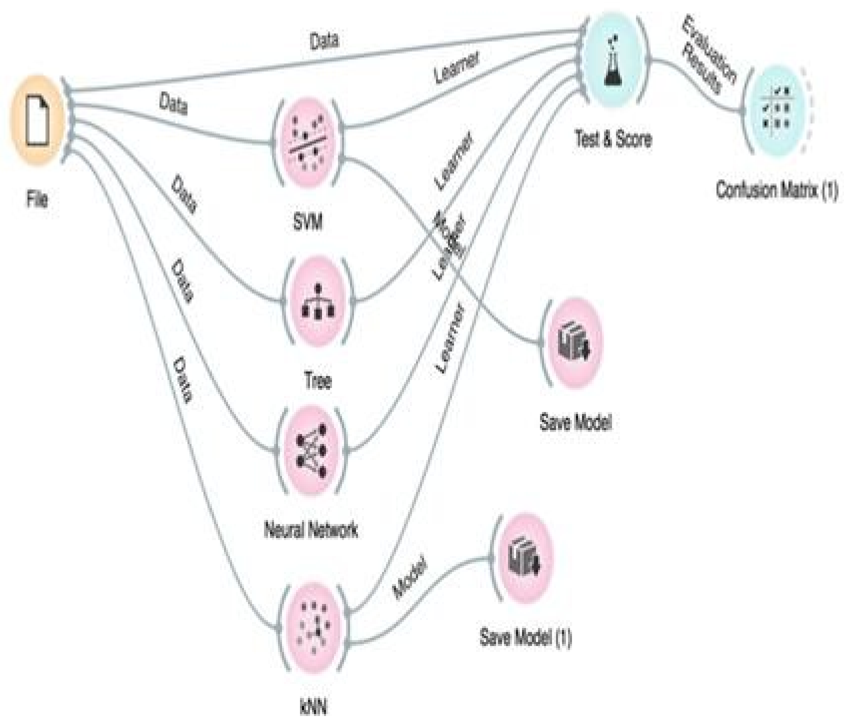


Figure 5.10: ML model made using Orange3

5.6 Model for machine learning:-

We employ Orange3 and Knime to construct a machine learning model that utilizes SVMs, Neural Network, Decision Tree, and KNNN algorithms following the feature extraction phase. We tried random sampling with a 65 percent training set size, but we were unable to get an acceptable result due to the limited number of dataset instances (350). As a result, tenfold cross validation is utilized. We examine the data in terms of accuracy, f-measure, precision, and recall after we've constructed the trained model. The confusion matrix is used to assess the percentage of events that are classed erroneously. The training model is illustrated in orange3 and Knime is displayed in blue in Figures 3 and 4.

Knime is faster than orange3 in developing a machine learning model, whereas orange3 is more user-friendly and easy to use.

5.6.1 SVM (Support Vector Machine):-

SVM is an abbreviation for supervised machine learning, which may be used for both regression and classification. Despite the difficulties associated with regression, classification is the best fit. The SVM method seeks a hyperplane inside an N-dimensional space that categorises data points unambiguously. The number of features defines the size of the hyperplane. If there are just two input characteristics, the hyperplane is merely a line. When the number of input characteristics reaches three, the hyperplane transforms into a two-dimensional plane. As the number of qualities reaches three, it becomes difficult to visualise.

5.6.2 Neural Networks (NN):-

Machine learning includes artificial neural networks (ANNs) and simulated neural networks (SNNs), which are at the core of deep learning methodologies. They get their name and structure from the human brain, and they interact with each other in the same way that real neurons do.

A node layer has a receptive field, one or more hidden layers, and an output layer in artificial neural networks (ANNs). Each node, or synthetic neuron, is linked to the others and has its own threshold and weight.

When a node's output hits a certain threshold, it is activated, and data is sent to the next tier of the network. Otherwise, no data is sent on to the network's next tier.

5.6.3 Decision Tree:-

The decision tree is the most well-known and commonly used tool for classification and prediction. It is a tree like architecture that looks like a flowchart, with each internal node representing an attribute testing, each branch representing the test's result, and each leaf in decision tree (terminating node) storing a target class.

5.6.4 kNN Algorithm:-

The K-Nearest Neighbour technique is one of the most fundamental Machine Learning algorithms, and it is based on the Supervised Learning methodology.

The K-NN approach implies that the new case/data and current cases are comparable, and it assigns the new case to the category that is most similar to the previous ones.

The K-NNN approach stores all available data and identifies data points based on their similarity to previously saved data. This means that utilising the K-NN approach, new data may be swiftly sorted into a well-defiined category.

Although the K-NNN approach may be used for both regresssion and classsification, it is more often employed for classification.

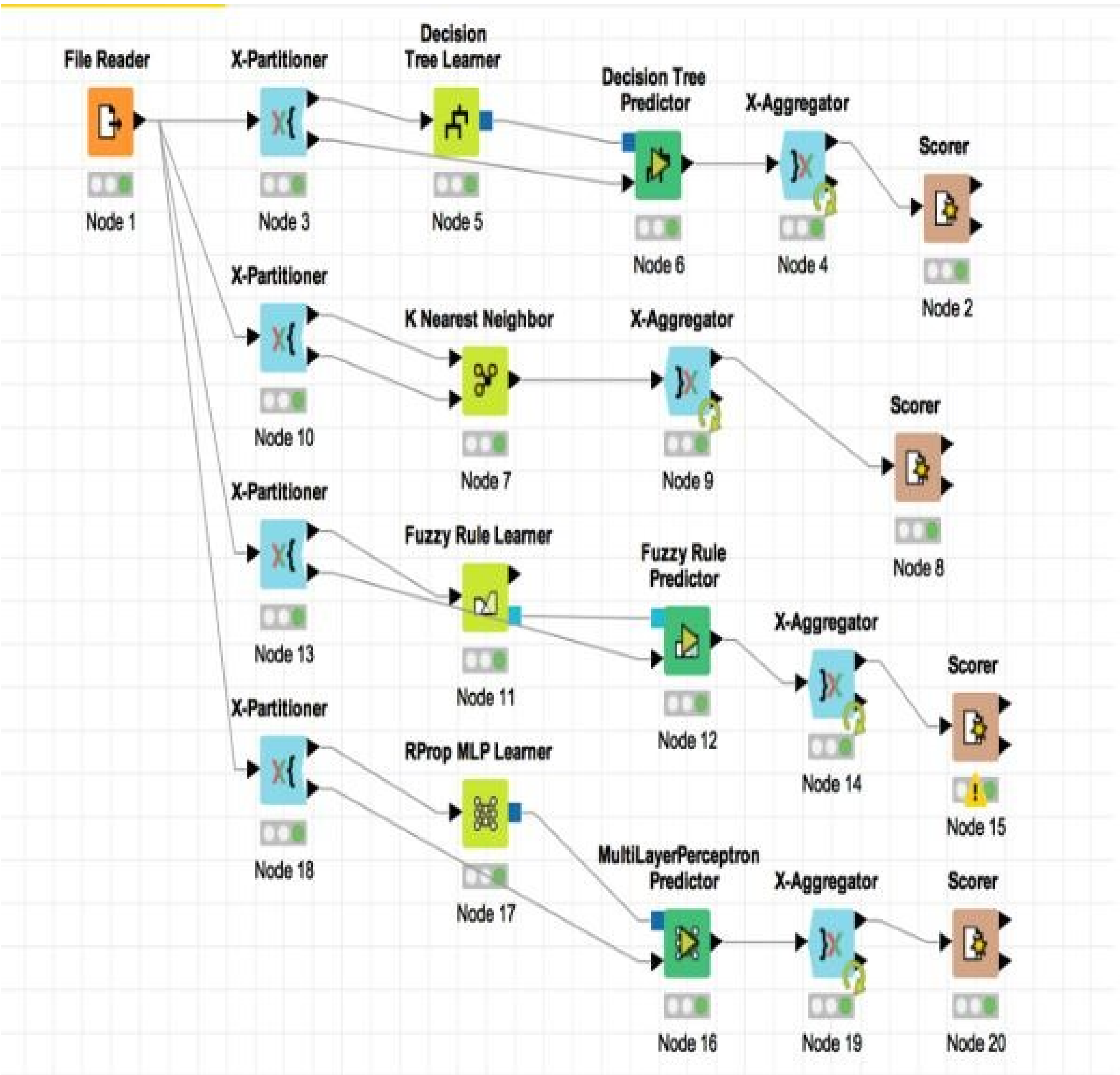


Figure 5.11: ML model using Knime

CHAPTER 6
DEPLOYMENT

We have deployed our mushroom toxicity classifier to amazon web cloud services. Now anyone can access our mushroom classifier by visiting the link https://mushroomclassifier.sa.com/

6.1 Screenshots of the web programme :-

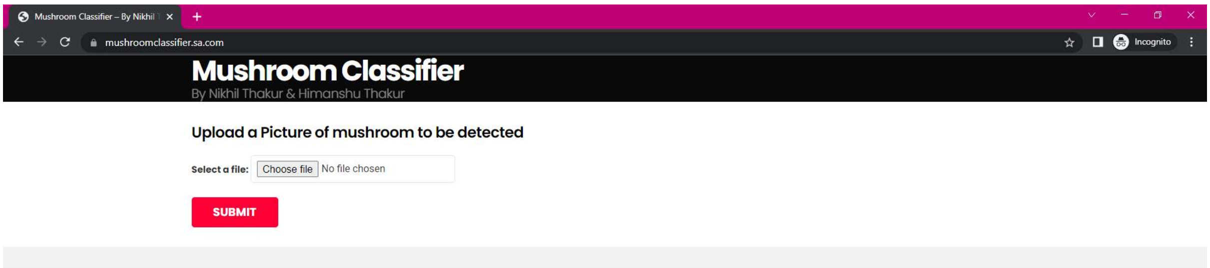


Figure 6.1: Mushroom Picture Upload Page

Figure 6.2: Mushroom Info Page 1

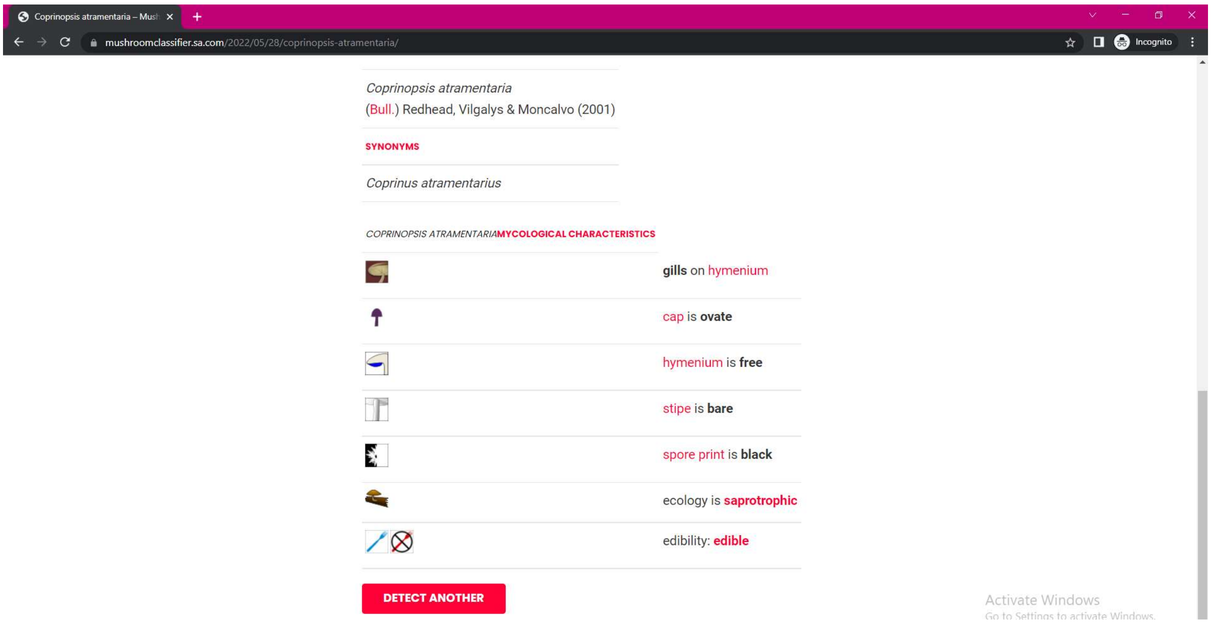


Figure 6.3: Mushroom Info Page 2

CHAPTER 7
CONCLUSION

We used the provided approach to apply each of the Support Vector Machine, Neural Network, Decision Tree, & K - nearest neighbors in various conditions, with and without backdrop, to get the best mushroom classification results.

Eigen features, histogram features, and parametric features are among the characteristics extracted from mushroom photos. In an attempt to improve the results, we eliminated the backgrounds from the images, however this approach failed. Finally, the experimental result shows that pictures without backgrounds have an advantage, especially while using the kNN Clustering techniques and Eigen features extraction with real mushroom dimensions (i.e. cup radius, stem height, and stem girth), with an accuracy of 0.89, while the final outcome when replacing real measurements with virtual dimensions (i.e. size and shape of mushroom shape inside the images) is 82 percent.
After eliminating the background from the photos, the greatest value for kNN was 0.841. We'll attempt to derive some physical dimension from mushroom photos in the future, such as cup size, stem height, colour, and texture. Also, to enhance the categorization process, we will also strive to increase the prediction data and employ additional photos.

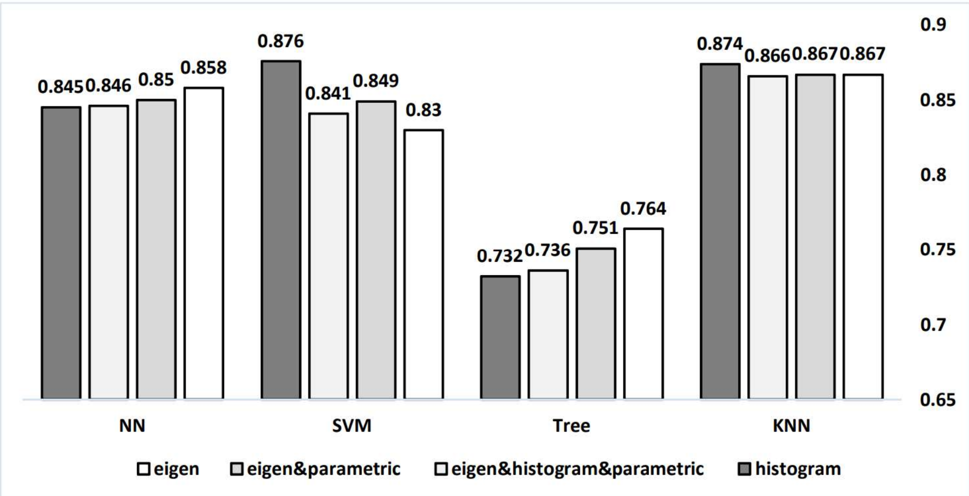


Figure 7.1: Algorithm accuracy graph

REFERENCES

[1] I. Al-Mejibli and D. Hamed Abd, "Mushroom Diagnosis Assistance System Based on Machine Learning by Using Mobile Devices Intisar Shadeed AlMejibli University of Information Technology and Communications Dhafar Hamed Abd Al-Maaref University College," vol. 9, no. 2, pp. 103–113, 2017. https://doi.org/10.29304/jqcm.2017.9.2.319

[2] M. Alameady, "Classifying Poisonous and Edible Mushrooms in the Agaricus," International Journal of Engineering Sciences & Research Technology, vol. 6, no. 1, pp. 154–164, 2017.

[3] R. LaBarge, "Distinguishing Poisonous from Edible Wild Mushrooms," 2008.

[4] I. Kononenko, "Machine learning for medical diagnosis: history, state of the art and perspective," Artificial Intelligence in medicine, vol. 23, no. 1, pp. 89–109, 2001. https://doi.org/10.1016/S0933-3657(01)00077-X [5] L. Von Ahn, B. Maurer, C. McMillen, D. Abraham, and M. Blum, "recaptcha: Human-based character recognition via web security measures," Science, vol. 321, no. 5895, pp. 1465–1468, 2008. https://doi.org/10.1126/science.1160379

[6] M. Tawarish and K. Satyanarayana, "A Review on Pricing Prediction on Stock Market by Different Techniques in the Field of Data Mining and Genetic Algorithm," International Journal of Advanced Trends in Computer Science and Engineering, vol. 3, no. 23–26, 2019. https://doi.org/10.30534/ijatcse/2019/05812019

[7] N. Bhargava and G. Sharma, "Decision Tree Analysis on J48 Algorithm for Data Mining," International Journal of Advanced Research in Decision Tree Analysis on J48 Algorithm for Data Mining, vol. 3, no. 6, pp. 1114–1119, 2013.

[8] A. Deshpande and R. Sharma, "Multilevel Ensemble Classifier using Normalized Feature based Intrusion Detection System," International Journal of Advanced Trends in Computer Science and Engineering, vol. 8, no. 3, pp. 874–878, 2019.

[9] P. Kumar, V. K. Sehgal, D. S. Chauhan, and others, “A benchmark to select data mining based classification algorithms for business intelligence and decision support systems,” arXiv preprint arXiv:1210.3139, 2012. https://doi.org/10.5121/ijdkp.2012.2503

[10] P. Babu, R. Thommandru, K. Swapna, and E. Nilima, “Development of Mushroom Expert System Based on SVM Classifier and Naive Bayes Classifier,” International Journal of Computer Science and Mobile Computing, vol. 3, no. 4, pp. 1328 –1335, 2014.

[11] F. E. Onuodu, “K-Modes Clustering Algorithm in Solving Data Mining Problems for Mushroom Dataset,” nternational Journal of Advanced Research in Computer Science and Software Engineering, vol. 5, no. 9, pp. 596–603, 2015.

[12] C. Eusebi, C. Gliga, D. John, and A. Maisonave, “Data Mining on a Mushroom Database,” Proceedings of Student-Faculty Research Day, pp. 1–9, 2008.

[13] D. Chowdhury and S. Ojha, “An Empirical Study on Mushroom Disease Diagnosis : A Data Mining Approach,” International Research Journal of Engineering and Technology(IRJET), vol. 4, no. 1, pp. 529–534, 2017.

[14] S. Beniwal and B. Das, “Mushroom Classification Using Data Mining Techniques,” International Journal of Pharma and Bio Sciences, vol. 6, no. 1, pp. 1170– 1176, 2015.

[15] “Mushroom Dataset.”, Retrevided from http://www.mushroom.world/.

[16] R. Socher, B. Huval, B. Bath, C. D. Manning, and A. Y. Ng, “Convolutional-recursive deep learning for 3d object classification,” in Advances in neural information processing systems, 2012, pp. 656–664.

[17] G. Wang, Q. Song, H. Sun, X. Zhang, B. Xu, and Y. Zhou, “A feature subset selection algorithm automatic recommendation method,” Journal of Artificial Intelligence Research, vol. 47, pp. 1–34, 2013. https://doi.org/10.1613/jair.3831

[18] F. Flores Camacho, N. Ulloa Lugo, and H. Covarrubias Martínez, “The concept of entropy, from its origins to teachers,” EDUCATION Revista Mexicana de Física E, vol. 61, no. December, pp. 69–80, 2015.

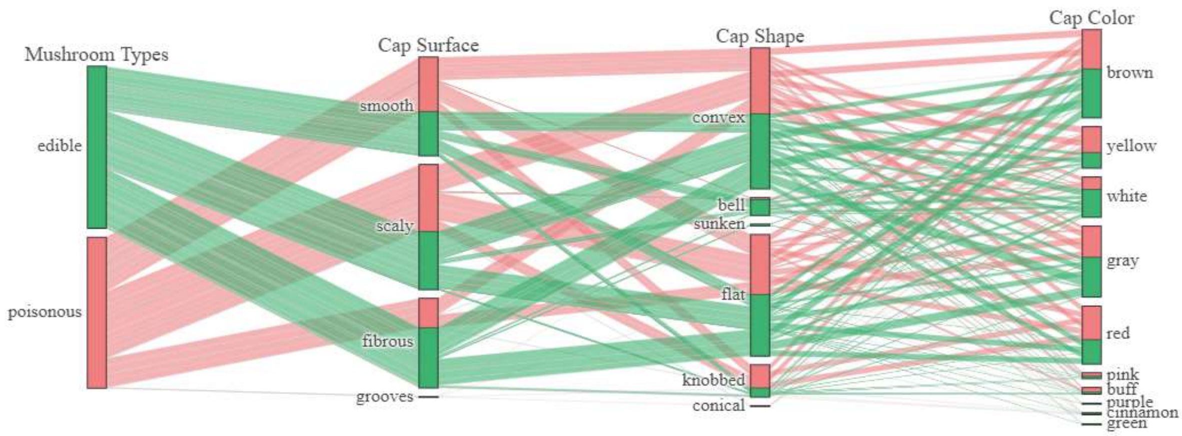
[19] R. Socher and B. Huval, “Convolutional-recursive deep learning for 3D object classification,” Advances in Neural …, no. i, pp. 1–9, 2012.

[20] C. Chang-yanab, Z. Ji-xian, and L. Zheng-jun, “Study on methods of noise reduction in a stripped image,” The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. XXXVII. Pa, no. 1, pp. 2–5, 2008.

[21] Ottom, Mohammad Ashraf & Alawad, Noor Aldeen. (2019). Classification of Mushroom Fungi Using Machine Learning Techniques. International Journal of Advanced Trends in Computer Science and Engineering. 8. 2378-2385. 10.30534/ijatcse/2019/78852019.

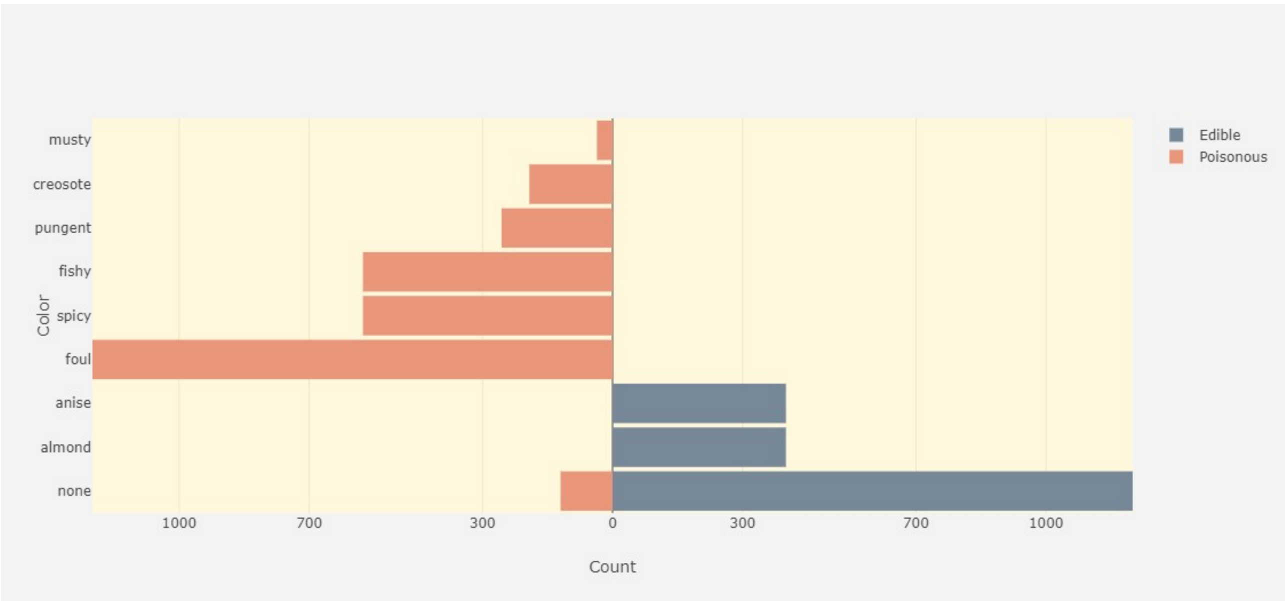
APPENDIX

A.



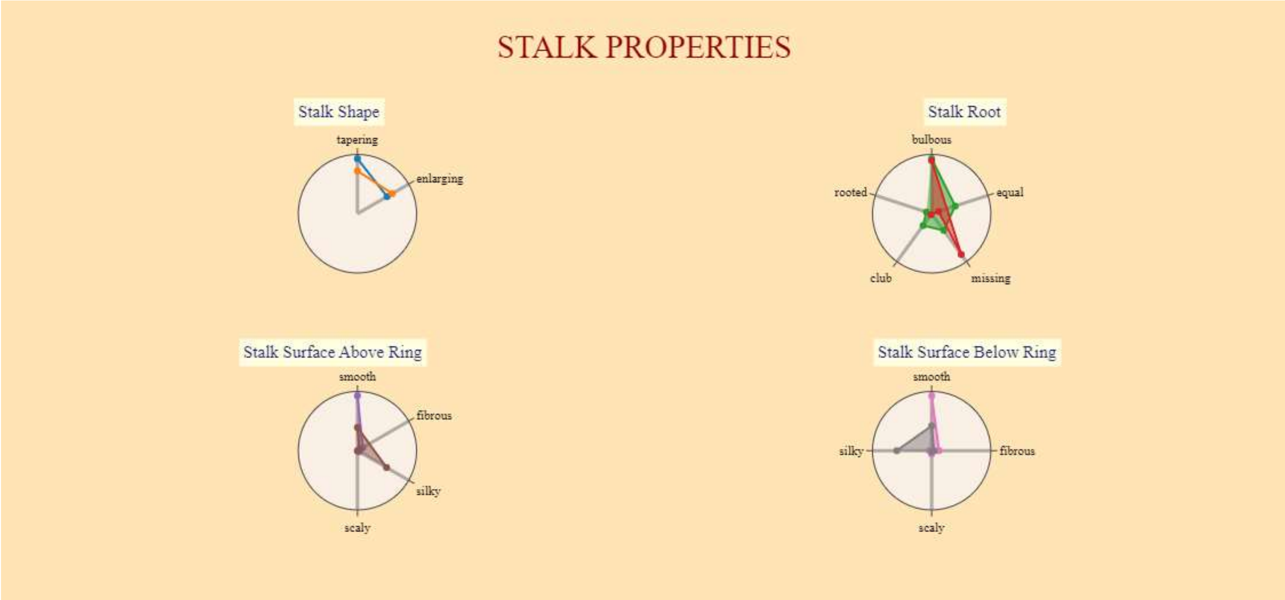
Mushroom Types
edible
poisonous
Cap Surface
smooth
scaly
fibrous
grooves
Cap Shape
convex
bell
sunken
flat
knobbed
conical
Cap Color
brown
yellow
white
gray
red
pink
buff
purple
cinnamon
green

B.

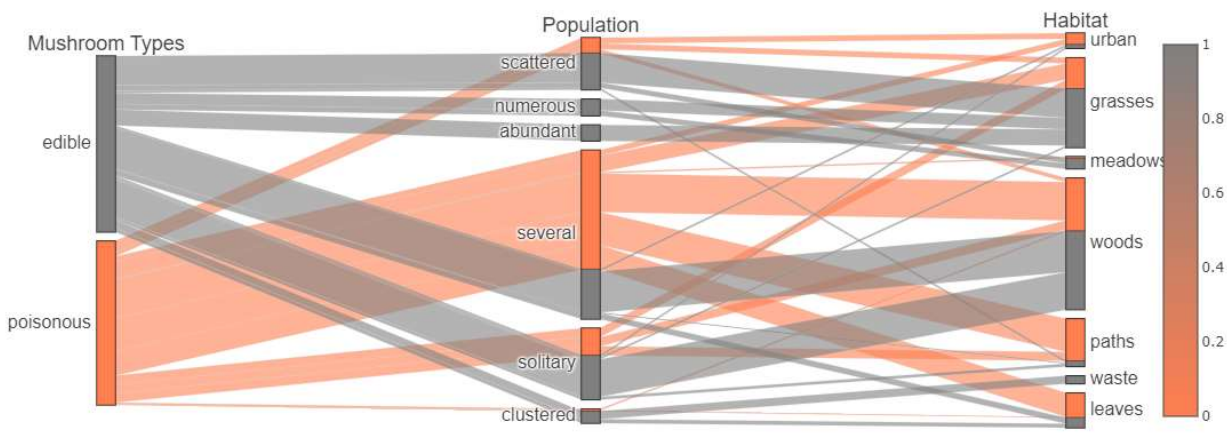


musty
creosote
pungent
fishy
spicy
foul
anise
almond
none
Color
1000
700
300
0
300
700
1000
Count
Edible
Poisonous

C.


STALK PROPERTIES
Stalk Shape
tapering
enlarging
Stalk Root
bulbous
rooted
equal
club
missing
Stalk Surface Above Ring
smooth
fibrous
silky
scaly
Stalk Surface Below Ring
smooth
silky
fibrous
scaly

D.


Mushroom Types
edible
poisonous
Population
scattered
numerous
abundant
several
solitary
clustered
Habitat
urban
grasses
meadows
woods
paths
waste
leaves
1
0.8
0.6
0.4
0.2
0

PLAGIARISM REPORT

181012_PR_28052022

ORIGINALITY REPORT

13%	10%	2%	5%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	www.warse.org Internet Source	4%
2	nceca.in Internet Source	2%
3	Submitted to Somaiya Vidyavihar Student Paper	2%
4	Submitted to Intercollege Student Paper	1%
5	Submitted to Universiti Teknologi Malaysia Student Paper	1%
6	ijircce.com Internet Source	1%
7	Submitted to Bangladesh University of Professionals Student Paper	1%
8	www.researchgate.net Internet Source	1%
9	Submitted to Coventry University Student Paper	<1%