IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 9, Issue 3, March 2021

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 7.488

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 3, March 2021 ||

| DOI: 10.15680/IJIRCCE.2021. 0903106|

Mushroom Quality Classification using Machine Learning

Mr.M.Vivek Kumar M.E., Mohanapriya.S, Nirmala.P, Nishanthini.M, Petchi Kavipriya.R

Assistant Professor, Department of ECE, Sri Eshwar College of Engineering, Coimbatore, Tamil Nadu, India

B.E Student, Department of ECE, Sri Eshwar College of Engineering, Coimbatore, Tamil Nadu, India

B.E Student, Department of ECE, Sri Eshwar College of Engineering, Coimbatore, Tamil Nadu, India

B.E Student, Department of ECE, Sri Eshwar College of Engineering, Coimbatore, Tamil Nadu, India

B.E Student, Department of ECE, Sri Eshwar College of Engineering, Coimbatore, Tamil Nadu, India

ABSTRACT: Mushroom is one of the most powerful nutrients in the plant in fungal species. The medicinal benefits of mushrooms include cancer-cell killing. This research aims at finding the best technology to classify mushrooms and classifies mushrooms into two groups, poisonous and non-poisonous. The method proposes implementing several techniques and algorithms, including the Neural Network (NN), the Support Vector Machines (SVM), the Decision Tree, and k Nearest Neighbours, on mushroom sets, with source and background images included in the dataset. Experimental results showed that kNN, with 94% precision, is the best technique to distinguish images from images with actual mushroom varieties, and 87% based on features taken from images.

KEYWORDS: Machine Learning, Mushroom Classification, Supervised Learning.

I. INTRODUCTION

Currently, structures that interpret broad and complicated data for better decisions are questioned. This thesis aims to establish a new method for classifying mushroom images using various techniques of Machine Learning based on various features (ML). Classification processes strive to anticipate category class marking or goal value[1], i.e. feed- to-forward neural artificial network (ANN), and to map data to pre-defined classes or groups[2]. We used the training data collection of mushroom images in the suggested solution to divide it into poisonous and non-poisonous. When we apply the characters of mushrooms to various techniques of machine learning, our methodology is designed to classify and forecast mushrooms for the class (groups).A mushroom is one of the most powerful nutrients in the plant in terms of fungal forms. The benefits of mushrooms include killing cancer cells, viruses and improving the human immune system. The mushroom now refers to the robot operating in the food industry. This technique has been used to restrict color functions. The mushroom machine recently used specific features that maximize mushroom selection. This method relies on assessing and evaluating characteristics such that the classification based on the well- known characteristics is better[3].To enable systems based on history, knowledge, instances, and data, Machine Language (ML) has made considerable advancements in many ways, under the umbrella of artificial intelligence [4]. Machine learning involves the study of artificial intelligence computer learning and theory of pattern recognition. Machine learning also shines light on how strategies can be built that can learn and predict available data. For instance, applications such as network intruder's detection or e-mail filtering, OCR, and PC vision [5].We will use various algorithms and techniques for the machine learning of mushrooms in this analysis. Some of them are described below.Neural network: is a hierarchical matrix system that is applied in various applications, such as data and pattern classification, new case prediction, or example prediction and pattern recognition applications. NN simulates the capacity of people to think and learn from the biological cells of man [5][6].Decision Tree: is one of the most common machine learning classification methods for decision support systems. This decision aims to distinguish objects (instances). [7][8].kNN is a machine learning algorithm classified. The kNN had a low number of training parameters, which did not have high computing sophistication and adequate performance [7].

II. RELATED WORK

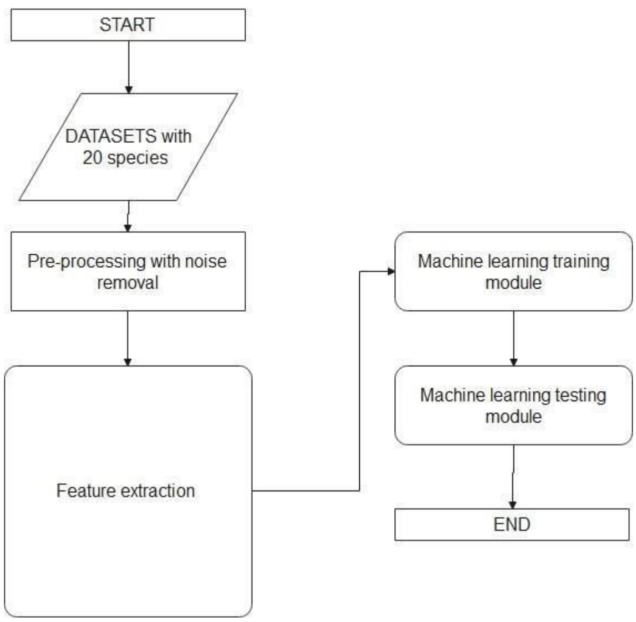
Several scientific studies are using various methods used to classify mushrooms. [3] suggested to include three components of the Mushroom Diagnostic Assistance System (MDAS) which include the web application (server), consolidated database, or the customer smartphone software (client), which are used on mobile telephones. For

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |
|| Volume 9, Issue 3, March 2021 ||
| DOI: 10.15680/IJIRCCE.2021. 0903106|

determining mushroom varieties, the Bays of Naive and the Decision Tree classifiers are used. The suggested scheme selects the most recognized properties of the mushroom. Secondly, enter the form of mushrooms. The test results reveal that inaccurate and wrong categorized instances and error calculations Decision Tree is superior to the Naïve Bays classifier. In [9], Kumar and others compared various typing strategies used in decision-making data mining. The algorithms represented by a statistic, artificial neural network, a support vector machine, and a clustered algorithm are comparable across three decision trees, respectively. Four databases from several disciplines are used to measure the predictive precision, error rate understandability, classification index, and training time of the proposed solution.
The experimental findings have shown that the predictive accuracy parameter compares GM and help algorithms of vector machines with others. QUEST Algorithm produces trees with smaller width and depth in decision tree-based algorithms. The best algorithm that can be used for decision assist systems is, in short, the GA-based algorithm. A smartphone and web software called Mushroom Diagnostic Assistance Framework can be used by Al-mejibli A smartphone and web software called Mushroom Diagnostic Assistance Framework can be used by Al-mejibli and Hamad in [1] built to provide protection when harvesting mushrooms. They used the tree of judgment and naïve bays for grouping the varieties of mushrooms. They were determined by the best-known mushroom characteristics. The key phases of this model are preparation and selection, which will delegate more active functions in the selection process and find the final decision. The experimental findings revealed that the decision tree is higher than naive bays based on calculations by mistake, samples correctly categorised and samples wrongly graded. The authors [12] have been using various methods for data mining and weka- mining software to evaluate a previous mushroom data collection. They used the closest neighbour classifier, an algorithm to compile proper rules, an untapped decision book and an algorithm of voted interpretation. They came from the use of techniques by stockholders on various categories, to have the most reliable outcome in an untapped tree and then use dynamic mushroom recognition with human-machine application based on the Internet.Chowdhury and S. Ojha in [13] described many mushroom diseases through various methods of data mining. They used real data collected by data mining like Naïve Bayes, RIDOR, and SMO algorithms from a mushroom farm. They conducted a statistical comparison to detect common mushroom symptoms to detect mushroom diseases. They also reached the naive Bayes produces the best results as compared with others. Data Mining methods like Zero, naïve Bayes and Bayes network were employed by Beniwa and das in [14] to examine mushroom dataset containing different types of mushrooms which are toxic or not toxic. They assessed grading methods with precise use of kappa statistics and an outright mistake. They also considered Bayes net to be the lowest mean absolute mistake and the greatest precision.

METHODOLOGY:The purpose of this analysis is to recognise Mushroom pictures and to classify them through machine learning technologies into two groups (poisonous and non- poisonous).
RESEARCH PHASES:The first step involves the collection of datasets and the second phasing of the dataset, the third phase comprises of the extraction, followed by the master training and then an assessment phase. Our approach consists of five stages in this study. The analysis phases of the suggested solution are shown in Figure 1.

Figure 1: Research phases



| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 3, March 2021 ||

| DOI: 10.15680/IJIRCCE.2021. 0903106|

III. COLLECTING DATASET

In the first stage, we extracted our dataset from [15], where the collected data consists of three categories, a collection of mushroom photographs (edible, inedible, and poisonous).Apart from pictures: families, location, measurements and edibility each type of mushroom is identified by different details. An example of mushroom photos can be found in Figure 2 below:

Figure 2: Dataset sample

FEATURES EXTRACTING: During this process, we used Matlab to remove all the features from the raw data collection of collected images. First, after resizing each image, we extract the Eigen characteristics. Second, we are the best in the top 100. Include details with function matrix for the dataset, i.e. the cap width, stub tall and width, which is available for each data type of dataset. Finally, we build the ML (Machine Learning) model to include the function matrix, which contains each of Eigen's dimensions with cap width, stencil high and diameter.The ML model was implemented using various techniques in the suggested solution, such as: neural network (NN), decisions tree (DT), support vector machine (SVM), and kNN. KNN with cross validation was better achieved with 10 folds in number and precision 94 percent. Although we opted to use derived features from photographs only because of the difficulties in achieving the actual measurements of mushrooms.

width and height for mushroom shapes can be found in photographs, using grey image scale detection of edges as seen in figure 6, to try to improve the results.In addition to the Eigen functions, add these dimensions. The findings of the experiments showed that KNN had 86 percent accuracy.We attempted to extract different features from images to enhance the results like histogram features. We applied the same Steps for each of height and width calculation by detection edges in the previous experiment. We also developed new data set features matrix. The test results indicate that 87% of histograms have been accurate. We are creating an algorithm in order to refine our results so that we can draw on more features, called this category, which are: Local comparison normalisation (LCN): used to contrast characteristics both within a function map and across functional maps in the same spatial area, where it is based on neuroscience computing [16].Skewness, default and curtosis: the terms in the measured statistic meta-function are considered. Calculate this with all numeric properties and take the mean by considering a mathematical concept[17].Entropy: a dynamic history phenomenon that is subjected to a wide array of reconstructions and interpretations identified by a stochastic data source as the average amount of information[18].

Mean: helpful for evaluating potential losses and advantages. For example, we used the method proposed to determine the matrix of features, particularly for the calculation of image height and width.**Correlation**: Correlation is one of the most often employed in which a reciprocal interaction or associations between the quantities is referred to by the expression "correlation" [19].Homogeneity: is one of the broad categories of distributed data mining and applies, over all the participating nodes, to the method of mining the same attribute collection.**Diameter**: the actual or virtual mushroom stem diameter is large.**Less noise:**Noise reduction is a critical aspect in the accuracy of the image[20], since it works to reduce the issues of the image. We would use noise reduction in the proposed solution to delete useless parts from initial images like the backdrop of the images.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 3, March 2021 ||

| DOI: 10.15680/IJIRCCE.2021. 0903106|

FEATURES EXTRACTION IMAGES WITHOUT BACKGROUND: In the present step, Matlab was used to remove features fromEigen for modified photographs (i.e. pictures with no background). We calculated the height and width for each picture depending on the edges of the mushroom pictures, using the gray-scale detection edge to create new images.

Machine learning model: We use python after the elimination of the function to construct the model of a system and use various algorithms such as SVM, Neural Network, Decision Tree and KNN. We use random samples of 66% scale, but because of the limited number of 380 instances in the dataset, we did not get a successful result. Thus, we use 10 folds cross confirmation. We test the effects with precision, f-measurement, precision and retrieval, after constructing the qualified model. The percent of incorrectly categorised instances is calculated with the uncertainty matrix. The training pattern in python using google colab respectively as seen in Figures 3 and 4.

Layer (type)	Outer Shape	Param #	Trainable
Conv2d	[64, 128, 128]	9,408	True
BatchNorm2d	[64, 128, 128]	128	True
RELU	[64, 128, 128]	0	False
MaxPool2d	[64, 64, 64]	0	False
Conv2d	[64, 64, 64]	36,864	True
BatchNorm2d	[64, 64, 64]	128	True
RELU	[64, 64, 64]	0	False
Conv2d	[64, 64, 64]	36,864	True
BatchNorm2d	[64, 64, 64]	128	True
Conv2d	[64, 64, 64]	36,864	True
BatchNorm2d	[64, 64, 64]	128	True
RELU	[64, 64, 64]	0	False
Conv2d	[64, 64, 64]	36,864	True
BatchNorm2d	[64, 64, 64]	128	True
Conv2d	[64, 64, 64]	36,864	True
BatchNorm2d	[64, 64, 64]	128	True
RELU	[64, 64, 64]	0	False
Conv2d	[64, 64, 64]	38,864	True
BatchNorm2d	[64, 64, 64]	128	True
Conv2d	[128, 32,32]	73,728	True
BatchNorm2d	[128, 32, 32]	256	True

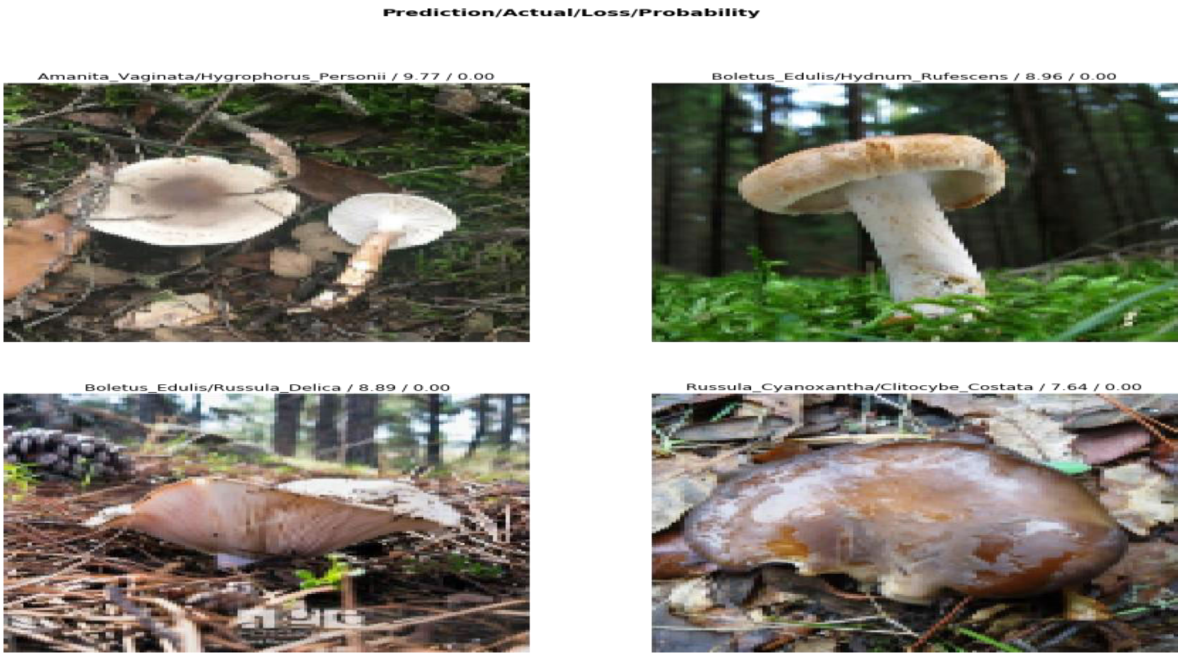
Figure 3: Machine Learning model using python

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 3, March 2021 ||

| DOI: 10.15680/IJIRCCE2021. 0903106|

We conclude that Knime is faster than orange,3 but orange3 is more user-friendly and usable than Knime after various tools for creating machine learning models.



IV. SIMULATION AND RESULTS

We connect it to real dimensions, after removing all Eigen features from images (cap diameter, stem tall). Figure 5 displays the accuracy findings for Eigen in various machine learning methods, with actual dimensions.

FIGURE 5

EVALUATION RESULT					
METHOD	AU C	CA	F1	PRECISIO N	RECAL
KNN	0.944	0.896	0.899	0.906	0.896
Tree	0.849	0.878	0.876	0.875	0.878
SVM	0.583	0.775	0.676	0.600	0.775
NEURAL NETWOR K	0.596	0.751	0.703	0.689	0.751

For Eigen with real dimensions, the experimental findings showed that KNN technique delivered maximum precision (0,944).Due to not possible to calculate the proportions when we have a mushroom image only, we have tried to derive virtual characteristics from photographs by measuring mushroom imagery heights and widths as seen in FIGURE 6.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 3, March 2021 ||

| DOI: 10.15680/IJIRCCE.2021. 0903106|

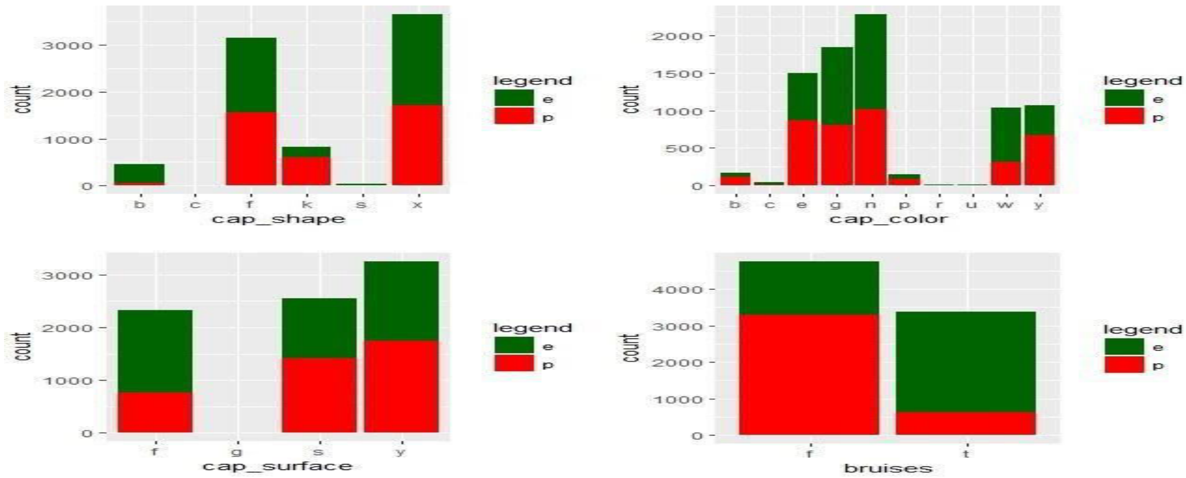


Figure 6: Width & height calculation

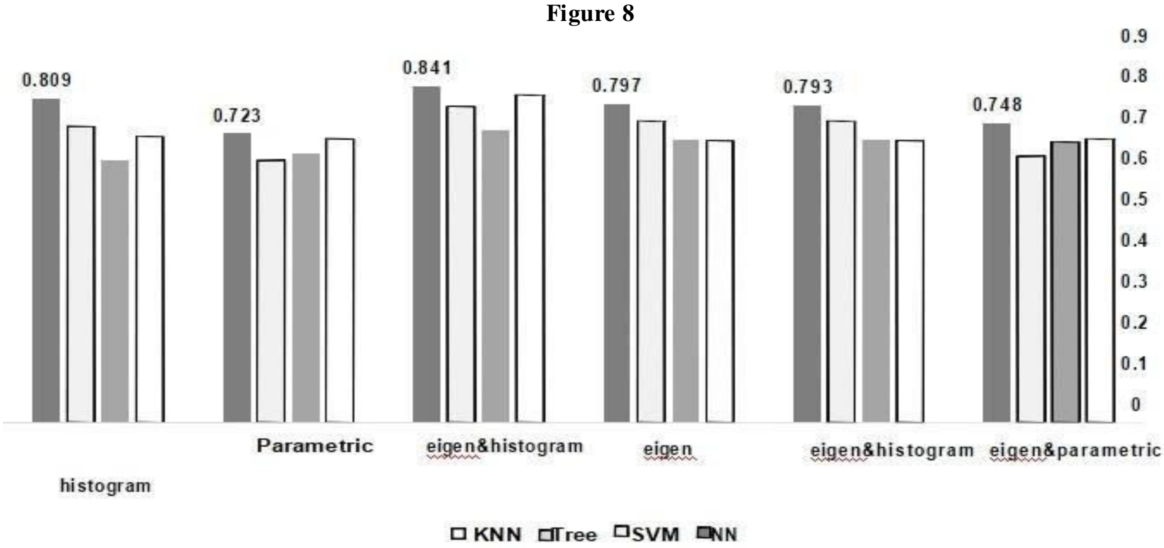
Figure 7 below shows the findings following calculation of virtual dimensions from images and extraction of Eigen's features from images. As we can see, KNN's most effective results are 87 percent accurate.

EVALUATION RESULT					
METHOD	**AUC**	**CA**	**F1**	**PRECISION**	**RECAL**
KNN	0.867	0.854	0.852	0.855	0.854
Tree	0.764	0.801	0.800	0.801	0.801
SVM	0.830	0.674	0.670	0.700	0.674
NEURAL NETWORK	0.858	0.851	0.848	0.866	0.851

Figure 7: Evaluation results for Eigen features

EVALUATION RESULT					
METHOD	**AUC**	**CA**	**F1**	**PRECISION**	**RECAL**
KNN	0.874	0.864	0.862	0.872	0.864
Tree	0.732	0.783	0.783	0.783	0.783
SVM	0.876	0.843	0.841	0.852	0.843
NEURAL NETWORK	0.845	0.815	0.814	0.816	0.815

Figure 8



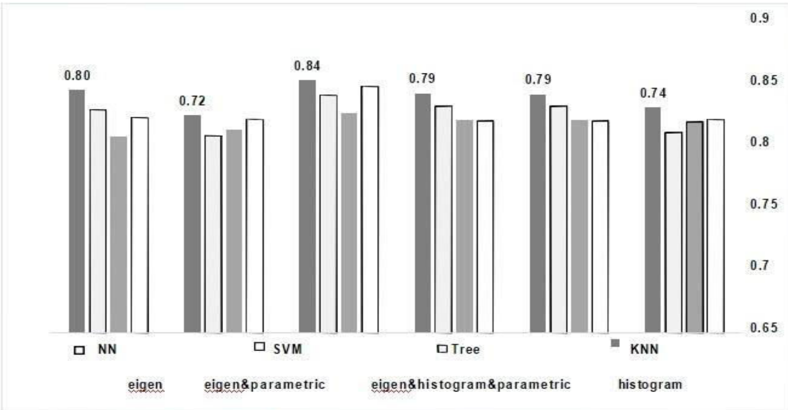


figure 11:evaluation result without background

V. CONCLUSION AND FUTURE SCOPE

We also used a number of algorithms to obtain better results in the classification of mushrooms and introduced each neural network (NN). We also removed certain characteristics from the mushroom pictures, such as Eigen's characteristics, histograms and parameters. In order to increase efficiency, we can subtract background frames, but unfortunately this step did not improve the result. Finally, the experiment results show benefits, especially in the case of the KNN algorithm for background images. While accuracy achieved 0.844, the real scale of the shapes (i.e. the height and width of the mushrooms) has been replaced by proportions simulated. This leads to 87%. The overall value for KNN after extracting the videos was 0.819. We are trying our work to obtain a certain physical dimension from the mung images such as cup diameters, high stems, colour and texture. In order to strengthen the classification process we will also aim to expand the dataset and use more image.

REFERENCES

1. I. Al-Mejibli and D. Hamed Abd, "Mushroom Diagnosis Assistance System Based on Machine Learning by Using Mobile Devices Intisar Shadeed Al- Mejibli 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.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 3, March 2021 ||

| DOI: 10.15680/IJIRCCE.2021. 0903106|

3. R. LaBarge, "Distinguishing Poisonous from Edible Wild Mushrooms," 2008.
4. 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.
5. https://doi.org/10.1016/S0933- 3657(01)00077-X
6. 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.
7. https://doi.org/10.1126/science.1160379
8. 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.
9. https://doi.org/10.30534/ijatcse/2019/058120 19
10. 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.
11. A. Deshpande and R. Sharma, "Multilevel Ensemble Classifier using Normalized Feature based Intrusion Detection System," *International Journal of Advanced*

INNO SPACE
SJIF Scientific Journal Impact Factor
Impact Factor: 7.488

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर
NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

Scan to save the contact details