FOREX Publication
Open Access | Rapid and quality publishing

International Journal of Computer Science Research (IJCSR)
Research Article | Volume 3, Issue 3 | Pages 125-130 | ISSN XXXX-XXXX

Mushroom Classifier System Using Machine Learning Algorithm

M.Jasmine Pemeena Priyadarsini[1*], S. Rajalakshmi[2], ShabareeshParlakot[3], Naveen Kumar Reddy[4] Ernest Bravin Clinton S[5], Rizee P.J[6] and Ganesan Subramanian[7]

[1]*School of Electronics Engineering, Vellore Institute of Technology, Vellore, 632014, India; jasmine@vit.ac.in*
[2]*School of Electronics Engineering, Vellore Institute of Technology, Vellore, 632014, India;*
[3]*School of Electronics Engineering, Vellore Institute of Technology, Vellore, 632014, India; p.shabareesh2017@vitstudent.ac.in*
[4]*School of Electronics Engineering, Vellore Institute of Technology, Vellore, 632014, India; naveenkumar.reddy2017@vitstudent.ac.in*
[5]*School of Mechanical Engineering, Vellore Institute of Technology, Vellore, 632014, India; Ernestbravin.clinton2020@vit.student.ac.in*
[6]*School of Computer science and Engineering, Vellore Institute of Technology, Vellore, 632014, India; rizee.pj2022@vitstudent.ac.in*
[7]*School of Engineering and IT, Manipal Academy of Higher Education, Dubai, ganesan.subramanian@manipaldubai.com*

***Correspondence:** *jasmine@vit.ac.in*

ABSTRACT- This paper studies the classification method for mushrooms based on Machine Learning algorithms and Graphic Processors. As mushrooms have many species, classification of mushrooms turns out to be difficult. These large numbers of species contain some edible and some poisonous or deadly poisonous mushrooms. Sometimes mushroom recognition is difficult through our naked eyes and due to a lack of knowledge of the identification of edible mushrooms, recognition turns out to be complicated. Although there are experts in distinguishing poisonous mushrooms from the list of edible mushrooms, where 70-80 species are reported as poisonous, occasional cases occur of misidentification of fatal mushrooms. Also, mushroom collectors have no formal discipline for a testimony of mushrooms, due to which people consume such wild mushrooms misidentified as nutritious mushrooms resulting in life-threatening Categorical features, support vector machine, convolution neural network, graphics processor, machine learning g disease or death-causing illnesses. The main aim of this project was to apply a method to detect whether the mushrooms fit human consumption or not. This paper proposes a method to determine the quality of the mushrooms using a categorical dataset that has 23 distinct characteristics. To solve the complication of the classification of mushrooms, a supervised learning model with the associated learning model is used. This method achieves a good result through the comparison of total time and speed between GPU and CPU.

Keywords: Categorical Features, Support Vector Machine, Convolution Neural Network, Graphics Processor, Machine Learning.

ARTICLE INFORMATION

Author(s): M.Jasmine Pemeena Priyadarsini, S.Rajalakshmi, ShabareeshParlakot, Naveen Kumar Reddy, Ernest Bravin Clinton S, Rizee P.J and Ganesan Subramanian ;

Received: 15/08/2024; **Accepted**: 25/08/2024; **Published**: 15/09/2024;
e-ISSN: XXXX-XXXX;
Paper Id: IJCSR-030306;
Citation: 10.37391/IJCSR.030306

crossref member
CROSSREF.ORG THE CITATION LINKING BACKBONE

Publisher's Note: FOREX Publication stays neutral with regard to Jurisdictional claims in Published maps and institutional affiliations.

1. INTRODUCTION

A mushroom also known as a toadstool, is a fleshy spore-bearing fruiting body of a fungus that has been widely used since ancient times. It has been regarded as the supernatural food and becomes the global dominant across the globe. Due to its different varieties and characteristics, it is emphasized as the main ingredient of gourmet cuisine [1]. Mushroom has both nutritional and medicinal value. It is quite ubiquitous in the diets of rural and urban people 2]. Hence, the utilization of mushrooms is proliferating day by day. It is the source of selenium, riboflavin, potassium, carbohydrates, protein, and many other important nutritional elements which are beneficial to human health [1,3,4]. Along with primary metabolites like peptides, proteins, oxalic acid, fatty acids, and other significant elements, it also contains new anti-microbial secondary metabolites like steroids, terpenes, anthraquinones, quinolones, and derivatives of benzoic acid [5].

The high range of extraordinary health benefits of mushrooms contain cancer-fighting properties, and immunity boosters, that help lower cholesterol, and have anti-inflammatory powers, which help fight aging and strengthen bones. It also provides energy, protects the immune system, and prevents damage to cells and tissues [3,6]. It is also an essential attribute for recycling agro-waste to provide nutritional and medicinal benefits to the vegetarian population using high protein conversion efficiency [7]. They are considered and accepted because of their organoleptic merit, medicinal properties, and economic significance1. The total world production of mushrooms is 40 million tons contributed by the following

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of
Computer Science Research (IJCSR)
Research Article | Volume 3, Issue 3 | Pages 125-130 | ISSN XXXX-XXXX

countries: China, UK, India, USA, The Netherlands, Poland, France, Spain, Italy, Ireland, and Canada [8,9].

The global button mushroom market is likely to increase 9.4% during the period from 2020 to 2030 which is depicted in Figure 1. Indonesia is among those agricultural countries which are recognized as a lack of professional knowledge, above methods are not viable. It is also not feasible to pick the right one through conventional manual classification by just observing color, shape, odor, and other features as it mainly depends on labor. So, errors may occur due to incessant exhaustion and emotional instability of workers, different personal grading systems, and precision criteria, which lead to classification mismatch. Hence, it is advisable to use a computerized method which is precautionary over a manual approach to attain more accurate results.

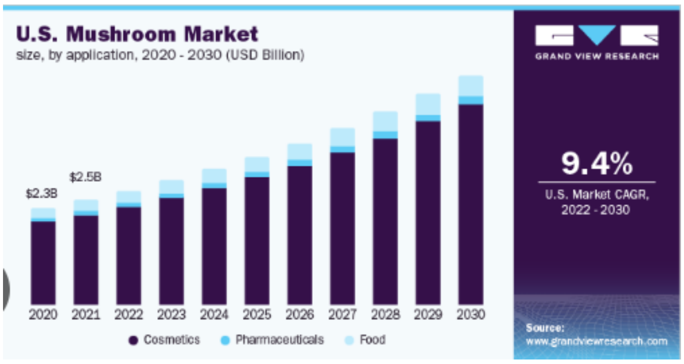

Figure.1 Global button mushroom market Statistics

Lastly, three distinct toxicity classifiers for mushrooms were created using three pattern-recognition techniques: multigrained cascade forest, logistic regression, and support vector machines. With an accuracy of almost 98%, the multigrained cascade forest classifier performed better than the logistic regression and support vector machine classifiers [10-12]. The international edible mushroom trade currently generates $3 billion year, and by 2019 it is expected to reach $5 billion annually. Europe holds the highest market share, with China and the US following closely behind. The main mushroom-producing countries ship tons of mushrooms each year to countries whose traditional national dishes heavily rely on mushrooms [13].

Toxic ingestion from wild mushrooms is a rare occurrence in Europe. Depending on the type of wild mushroom consumed, a wide range of clinical signs could appear. With frequent gastrointestinal issues, these early symptoms are frequently nonspecific and unrelated to the result. Emergency physicians face difficulties in managing mushroom poisoning and assessing danger [14]. Most mushroom gatherers in this region have no formal training in identifying poisonous and nonpoisonous mushrooms. They largely depend on experience and traditional wisdom passed on over generations to differentiate between edible and nonedible-mushrooms. Perhaps, a hundred mushroom species or more cause symptoms when ingested; of which only about fifty or so species worldwide are potentially fatal [15].

Many Indian ethnic groupings rely heavily on mushrooms in their diets. Despite their proficiency in differentiating between dangerous and edible mushrooms, there are sporadic reports of poisoning incidents brought on by inadvertent ingestion of hazardous mushrooms. We report a family experiencing amanita-like poisoning following ingestion of wild mushrooms, with tragic results [16]. Mushroom poisonings are a common case type, including accidental poisonings as well as those caused by crimes and suicide attempts. Additionally, there has been a rise in interest in hallucinogenic mushrooms in recent years. For legal processes, it is crucial to accurately identify fungal species that contain hazardous compounds. Morphological analysis is used for classifying different types of mushrooms [17].

By utilizing supervised machine learning models on the dataset of different mushroom properties provided by UCI, we are able to create a prediction system that is capable of classifying mushrooms. This project is motivated by the goal of determining which machine learning models perform the best on the dataset and which characteristics are most suggestive of toxic mushrooms [18]. Due to advancements in technologies, machine learning has become quite popular and plays a vital role in a classification purpose that falls under Artificial Intelligence. Basically, it is the study of how a computer can learn from experience and make predictions.

In this paper, our goal is to accurately classify the mushrooms into edible and poisonous classes using a support vector machine. Categorical features as well as images are used as the foundation to extract the various features like color, shape, etc. The Dataset has been taken from the UCI Machine Learning Repository which is a conglomerate of Agaricus and Lepiota Families containing 8124 samples having 23 distinct attributes. There are several operations that have been performed prior to SVM and various parameters are tuned and optimized to achieve better accuracy.

2. MATERIALS AND METHODS

The datasets used, the preprocessing and data augmentation methods and the various algorithms used are discussed in detail in the next subsections. The workflow for the work is presented in a flowchart form in *figure 2.*

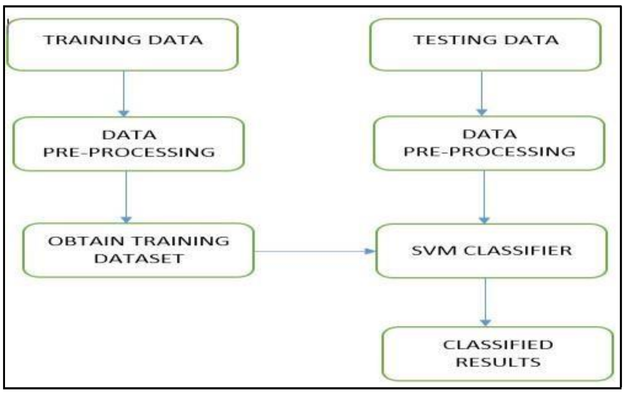

Figure 2. Flow chart of the work

2.1. Datasets

The Center for Machine Learning and Intelligent Systems, University of California, Irvine" is the source of the dataset. After that, a group of researchers uploaded the dataset, defined the input and output variables, separated it into training and validation sets, and computed the proper hidden layers. After that, the sets underwent testing and qualification to guarantee the best accuracy. The input variables selected are those from the Mushroom Database that can be quickly retrieved. veil type, stalk shape, stalk root, stalk surface above ring, stalk surface below ring, stalk color above ring, and veil attachment, spacing, size, and color of the gills veil-color, ring-type, ring-number, spore-principle, spoilsport principle, color, population, and habitat of the pore print. These variables were afterwards transformed into a format suitable for neural network research.

2.2. Preprocessing and Data Augmentation

Data must be cleaned and processed to be modeled. The most popular stages in data pre-processing are:

• Filling up the missing values using feature imputing. Our dataset has no missing values.)

• Feature Encoding: This process turns numeric data into categories. There is an abundance of this.

• Data splitting (dividing data into splits for testing and training).

2.3. Algorithms

SVM-support vector machines: A support vector machine is used to determine the decision boundary to separate different classes and maximize the margin. The basic principle of SVM is to create a hyperplane that separates the data into classes. There are two main criteria taken into consideration.

Finding a hyperplane: To choose the best plausible hyperplane that eliminates the optimization problem, we must consider the specific set of points called difficult points close to the decision boundary as shown in *figure 3*. Lagrange Multipliers technique is used to convert the problem into such a win that it can be solved analytically.

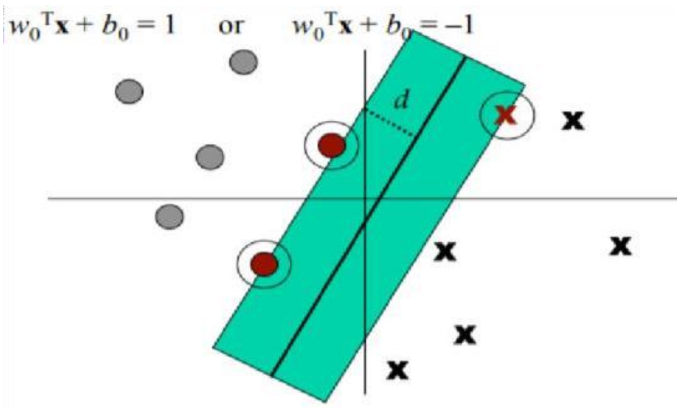

Figure 3. Decision boundary using a Hyperplane

Maximizing the margin: Minimizing is necessary to maximize the margin. |w|| There are no data points between H1 and H2, given the conditions in *equation 1* and *equation 2*. Therefore, combine Equations 1 and 2 to create $yi(xi, w) \geq 1$ One can tune the SVM by changing the tuning parameters of SVM algorithm that are mentioned as: kernel, regularization, gamma as shown below in *figure 4.*

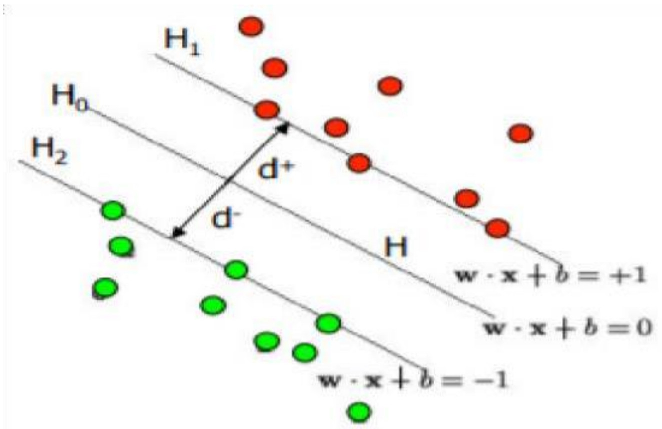

Figure 4. Maximizing the margin using SVM algorithm

Confusion matrix: The number of accurate and inaccurate predictions a classifier made is displayed in a table called a confusion matrix. It serves as a statistic to assess the performance of a classification model. The output of an algorithm in supervised learning, specifically for the statistical classification problem in machine learning, can be shown using a confusion matrix, which is a precise table design. Each row of the matrix represents an instance from a projected class, while each column represents an instance from a real class (or vice versa).

Terms we can measure from a confusion matrix are:
Precision: The precision metric shows the accuracy of the positive class. It is used to measure the correctness of a positive class.

Mathematical expression: $precision = \frac{TP}{(TP+FP)}$ (1)

Accuracy: Accuracy shows out of all classes how much we predicted correctly.

Mathematical expression: $accuracy = \frac{TP}{(TP+FN)}$ (2)

Recall: It states how much we predicted correctly out of all positive classes. It should be as high as possible.

Mathematical expression: $recall = \frac{TP+TN}{Total}$ (3)

F-measure: The F-score helps to measure recall and precision at the same time. Mathematical expression:

$$F - measure = \frac{2*recall*presicion}{recall+precision} \quad (4)$$

3. RESULTS

For each attribute, the corresponding value is obtained. Hence, we can easily identify which attribute contributes the most/least to make the prediction. The outcome shows the accurate classification of edible and poisonous mushrooms. The plot shows the two labeled data indicating red as edible and blue as poisonous. We can analyze no such drastic difference between actual labels and predicted labels which assure that our proposed method has efficiently done the categorization by optimizing and tuning various parameters. The classification of poisonous and edible mushroom is shown in below *figure 5*.

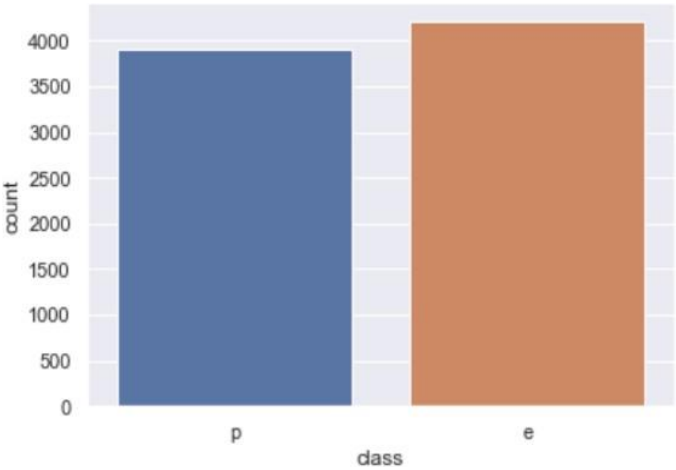

Figure 5. Classification of poisonous and edible mushroom

Logistic Regression Confusion Matrix
Test Accuracy: 94.96%

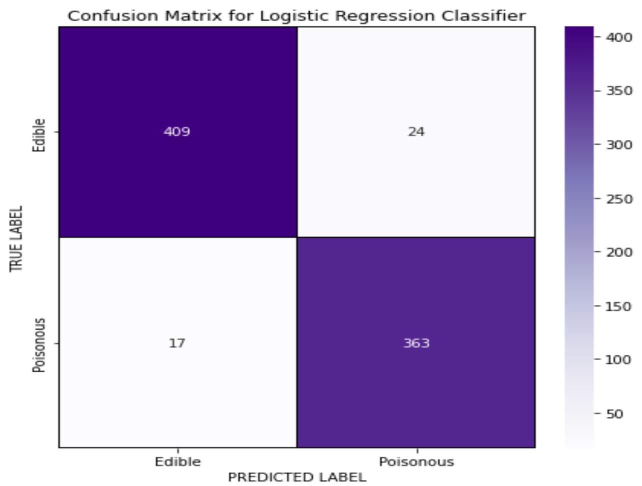

Figure 6. SVM Confusion Matrix

Test Accuracy: 97.1%

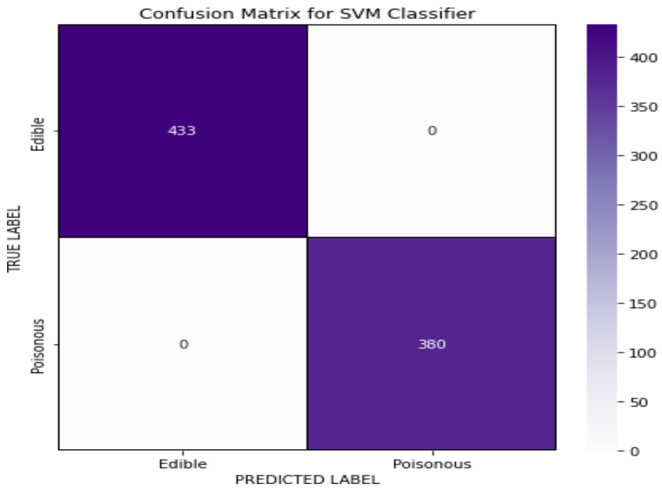

Figure 7. Naïve Bayes Confustion Matrix

Test Accuracy: 92.66%

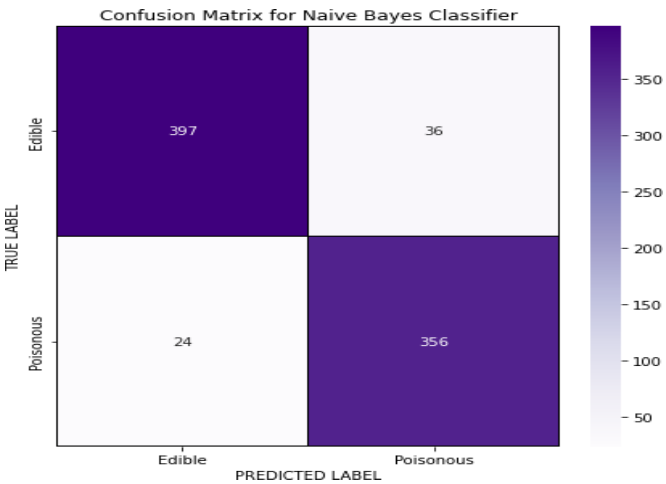

Figure 8.

Table 1. Comparison of different platform performing mushroom classification

Platform used for testing	Configuration	Time required to test data
CPU	Intel® Core (TM) with the processor configuration i5-6200U CPU, frequency is @2.30 GHz, Memory is 8 GB RAM.	1.14685 399e-05 Second
GPU Jetson Nano	NVIDIA Maxwell architecture which has memory of 128 NVIDIA CUDA® cores, 4 GB 64bit LPDDR4, frequency is 1600MHz with the speed of 25.6 GB/s	4.15928 936e-05 Second
Raspberry Pi 4 Model B	Pi 4 B is upgraded with Latest High-Performance Quad-Core 64-bit Broadcom 2711, Cortex A72 processor clocked at 1.5GHz speed.	1.11745 634e-05 Second

4. CONCLUSION AND FUTURE WORK

By summing up all the facts and proven results, there is no denial why we are striving to build some intelligent applications. The significant breakthrough in machine learning have not just provided us a better perspective to approach the problem but also helped us in optimizing and eliminating those parameters that can be a barrier to attaining more accurate outcomes. Our proposed methodology of implementing the SVM algorithm has improved the accuracy of actual and predicted data, optimization of algorithm parameters, and reduced the time consumption rate. By varying the kernel parameters of SVM we are able to compare the accuracy of the outcome. By leveraging different visualization libraries, we have plotted proportion levels of 23 distinct features that have helped us predict which feature is prominent or not in deciding the class of mushrooms. It is shown that veil the –type feature of the mushroom does not play any role in the prediction of the mushroom category, so it has been eliminated as there is no variation in its proportion level. Later, we produced a classification report and confusion matrix to compute different entities – precision, recall, f1-score, and support to ensure that the trained model is able to classify the mushroom categories accurately. Further, it is recommended to research more to find the quality of mushrooms with the same accuracy but within less amount of time for speedy production and selling in the global market. This may increase the economy of mushroom harvesting countries as well as it is also benefited for production of varieties of medicines.

Author Contributions: The research article have several authors and their dividual contributions are provided. As follows "Conceptualization, M.Jasmine Pemeena Priyadarsini and Ganesan Subramanian ; methodology, ShabareeshParlakot; software, Naveen Kumar Reddy; validation, M.Jasmine Pemeena Priyadarsini, and Ganesan Subramanian; formal analysis, Ernest Bravin Clinton S ; investigation, S.Rajalakshmi.; resources, Rizee P.J ; data curation, ShabareeshParlakot,; writing—original draft preparation, Ernest Bravin Clinton S.; writing—review and editing, S.Rajalakshmi.; visualization, Rizee P.J; supervision, M.Jasmine Pemeena Priyadarsini,; project administration, Ganesan Subramanian. All authors have read and agreed to the published version of the manuscript"

Funding: We would like to state the "This research received no external funding".

Acknowledgments: We would like to acknowledge our gratitude for the VIT management, DEAN and HOD to carryout this research in a successful manner.

Conflicts of Interest: We would like to declare that "The authors has no conflict of interest." We also wish to state the "The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results".

REFERENCES

[1] Valverde, M.E., Hernández-Pérez, T. and Paredes-López, O. 2015. Edible mushrooms:Improving human health and promoting quality life. Int. J. Microbiol. 2015:376-387

[2] Abulude, F. O., Ndamisto and Muhammed, M. 2013. Fungi: A review on mushrooms. In Gangully, S. (ed.). Current Trends in Advancement of Scientific Research and Opinion in Applied Microbiology and Biotechnology. Science and Education Development Institute, Nigeria, pp.18-31

[3] Cheung, P. C. K. 2008. Nutritional value and health benefits of mushrooms. In Cheung, P.C.K. (ed.). Mushrooms as Functional Foods. John Wiley & Sons, Inc., New York, NY, pp. 71-109

[4] Sharma, K. 2018. Mushroom: Cultivation and processing. International Journal of Food Processing Technology 5(12):9-12.

[5] Chatterjee, S., Sarma, M. K., Deb, U., Steinhauser, G., Walther, C. and Gupta, D. K. 2017. Mushrooms: from nutrition to mycoremediation. Environmental Science and Pollution Research 24:19480-19493.

[6] Badalyan, S. M., Barkhudaryan, A. and Rapior, S. 2019. Recent progress in research on the pharmacological potential of mushrooms and prospects for their clinical application. In Agrawal, D.C. and Dhansekaran, M. (eds). Medicinal Mushrooms, Recent Progress in Research and Development. Springer Nature, Singapore, pp.1-70.

[7] Thakur, M.P. 2020. Advances in mushroom production: key to food, nutritional and employment security: A review. Indian Phytopathology 73:377–395.

[8] Wibowo, A., Rahayu,Y., Riyanto, A. and Hidayatullo, T. 2018. Classification algorithm for edible mushroom identification. International Conference on Information and Communication Technology (ICOIACT).DOI:10.1109/ICOIACT.2018.8350746.

[9] Wang, Y., Du, J., Zhang, H. and Yang, X. 2020. Mushroom toxicity recognition based on multigrained cascade forest. Scientific Programming 2020(1):1-13.

[10] Diaz, J.H. 2016. Mistaken mushroom poisoning. Wilderness and Environmental Medicine 27:330–335.

[11] Schmutz, M., Carron, P.-N., Yersin, B. and Trueb, L. 2018. Mushroom poisoning: a retrospective study concerning 11-years of admissions in a Swiss Emergency Department. Internal and Emergency Medicine 13(1):59-67.

[12] Barman, B., Warjri, S., Lynrah, K.G., Phukan, P. and Mitchell, S.T. 2018. Amanita nephrotoxic syndrome: First case report in India. Indian Journal of Nephrology 28(2):170-172.

[13] Verma, N., Bhalla, A., Kumar, S., Dhiman, R.K. and Chawla,Y.K. 2014. Wild mushroom poisoning in north India: case series with review of literature. J. Clin. Exp. Hepatol. 4:361–365.

[14] Zuber, A., Kowalczyk, M., Sukela, A., Mleczko, P. and Kupiec, T. 2011. Methods used in species identification of hallucinogenic and other poisonous mushrooms in forensic investigations. Problems of Forensic Sciences 86:151–161.

[15] Chelliah, B.J., Kalaiarasi, S., Anand, A., Janakiram, G., Rathi, B. and Warrier, N. K. 2018. Classification of mushrooms using supervised learning models. International Journal of Emerging Technologies in Engineering Research (IJETER) 6(4):229-232.

[16] Prasad Babu, M.S.,Thommandru, R., Swapna, K. and Nilima, E. 2014. Development of mushroom expert system based on SVM classifier and Naïve Bayes Classifier. International Journal of Computer Science and Mobile Computing (IJCSMC) 3(4):1328-1335.

[17] Sopyła, K., Drozda, P. and Gorecki, P. 2012. SVM with CUDA accelerated kernels for big sparse problems. International Conference on Artificial Intelligence and Soft Computing, Part I.DOI:10.1007/ 978-3-642-29347- 4_51.

International Journal of
Computer Science Research (IJCSR)
Research Article | Volume 3, Issue 3 | Pages 125-130 | ISSN XXXX-XXXX

[18] Drozda, P. and Sopyła, K. 2016. Accelerating SVM with GPU: The State of the Art. International Conference on Artificial Intelligence and Soft Computing