DETECTION OF MUSHROOM INSALUBRITY BASED ON FEATURES EXTRACTED FROM IMAGES BY USING K-NEAREST NEIGHBOUR ALGORITHM COMPARING WITH CONVOLUTIONAL NEURAL NETWORK ALGORITHM

Sk Yaseen Ahmed[1], T.P.Anithaashri[2*]

Article History: Received: 12.12.2022 **Revised:** 29.01.2023 **Accepted:** 15.03.2023

Abstract

Aim: The aim is to improve the detection of mushroom insalubrity based on features extracted from images by using K-Nearest Neighbour (KNN) algorithm compared with Support Vector Machine (SVM) algorithm. **Materials and methods:** By using K-Nearest Neighbour algorithm and SVM algorithm both were identified and performed with the sample size of 45 each and the software tools that were used in this project are a jupyter notebook. Accuracy values for identification of toxicity in mushrooms are calculated to quantify the performance of K-Nearest Neighbour algorithm. Results **and**
Discussion : The analysis on train dataset and test dataset were successfully performed using SPSS and acquired accuracy for the Convolutional neural network compared to K-nearest neighbor algorithm method which gave more accuracy with the level of significance ($p<0.05$) and the resultant data depicts the reliability in independent sample tests.
Conclusion: On the whole process of prediction of accuracy the K-nearest neighbour algorithm gives significantly better performance compared with Convolutional neural network

Keywords: K-Nearest Neighbour Algorithm, Convolutional Neural Network, Machine Learning, Mushroom Toxicity, Image Processing.

[1]Research Scholar, Department of Computer Science and Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences. Saveetha University,Chennai,Tamilnadu.India, Pincode:602105
[2*]Department of Computer Science and Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Science, Saveetha University,Chennai,Tamilnadu,India, Pincode:602105.

1. Introduction

In this project, it will analyze data and create different literacy models that will explain whether a mushroom can be exposed or toxic by its specifications such as cap shape, cap color, gill color, etc. using different separators for image processing and machine reading (Al-Mejibli). The process below can be answered using machine learning methods (Auerbach, Paul S). can find that features can explain(Mushroom Toxicity). The toxic or emerging class of mushrooms we took a dataset from kaggle. There are many myths surrounding mushrooms and their consumption. He had done an analysis of the experimental data set in the python to remove those myths. Approach using image processing only works with delicacies (Yağan et al. 2020).

The existing system implemented through machine learning methods and image processing Neural Network and Adaptive Neuro Fuzzy conclusion systems are used for continuous brackets (Shuman and Li 1997). Different methods used in brackets are used to separate different mushrooms such as mushroom venom from or can not be eaten using an anaconda software(Auerbach et al. 2008). In addition, to monitor the complaint by the Bureau of Epidemiology, the health care system has reported that the number of cases and deaths in food poisoning is 6 people. (Hintikka 1978)The number of cases of eating poisonous mushrooms is increasing, because people living in the north and northeast of Thailand often prefer to collect wild mushrooms to feed them continuously. Wild mushrooms can thrive especially in businesses in the north and northeast of Thailand; therefore, the disease rate in the north-northeast regions is better than in the other Thai tunnel (Dogan et al. 2021). From a statistical analysis, it can be assumed that the incidence of morbidity and mortality is very high during May to November because this period is conducive to mushroom growth (Al-Mejibli, Intisar Shadeed).Our team has extensive knowledge and research experience that has translated into high quality publications(Pandiyan et al. 2022; Yaashikaa et al. 2022; Venu et al. 2022; Kumar et al. 2022; Nagaraju et al. 2022; Karpagam et al. 2022; Baraneedharan et al. 2022; Whangchai et al. 2022; Nagarajan et al. 2022; Deena et al. 2022)

The research gap in the existing system is implemented through machine learning and methods(Wang et al. 2020). This design is concerned with the broader perspective of recent experiments on emerging and edible mushroom brackets(Singh and Chauhan 2016). The purpose of this is to explain the pathophysiology of mushroom toxin. Review the donation of mushroom poisoning case. Consider mushroom treatment options(Masters 2002). Describe ways to improve care co-operation between members of a trained team to address issues of mushroom poisoning.(Brhane)

2. Materials and Methods:

The Study setting of the proposed work is done in the Compiler Design Lab, Saveetha School of Engineering, Saveetha University. The number of groups identified are two. Group 1 is the toxic mushroom and group 2 is the non-toxic mushroom.(Wagner, Heider, and Hattab 2021) Convolutional neural network was iterated various number of times with a sample size of 45 calculated from clinical website and SPSS analys(Wagner et al. 2021)is is carried out with level of significance $p<0.05$

The software tool used for detecting the accuracy score is using the python sklearn library. Hardware configuration was AMD RYZEN 5 (2.10 GHZ) processor with 16GB ram and 64 bit OS, x64 based processor system. The software configuration was Windows 11 professional. The data was pre processed after performing many steps such as removing noise data, feature engineering , and feature selection.

Convolutional neural network:

Convolutional neural network is a type of network model which allows us to extract higher representations for the image content unlike the classical image recognition where you define the image features yourself cnn takes the images raw pixel data trains the model then extracts the features automatically for better classification

Following are the steps for implementing the Convolutional neural network method

There are 8 steps included in the proposed method. Firstly extracting the data, data cleaning,removing nan values. Finally training data with the Convolutional neural network method

Step 1:Start importing the data from a csv file

Step 2:For training, some processing steps are required such as performing EDA

Step 3:Extracted features values.

Step 4:Apply the Convolutional neural network techniques.

Step 5: compare with knn algorithm

Step 6:Identify the toxicity poisonous(1) edible(0)

Step 7:calculate the accuracy values using SPSS Tool

Step 8: Then finally applying Convolutional neural network on the cleaned data and finally getting the accuracy of the model is 98.45%

K Nearest Neighbour:

According to this method system the mushroom is selected based on the authorized data that is divided into two groups.One is toxic and the other

one is edible based on the sample size and data .Using this knn algorithm we an train the data and compile the output.Knn is more accurate than cnn algorithm because of high efficiency

Following are the steps for implementing the k nearest neighbour method

There are 8 steps included in the proposed method. Firstly extracting the data, data cleaning,removing nan values. Finally training data with the k nearest neighbour method

Step 1: Start importing the data from a CSV file.
Step 2: Extract the features from the images using machine learning
Step 3: Importing Dataset. Here, we will see the dataset being imported.
Step 4: Split Dataset. Next step is to split our dataset into a test and train split.
Step 5: Training Model. Now in this step, we're going to see model training here we are using K neighbours
Step 6: Predict the accuracy
Step 7: Then finally applying Convolutional neural network on the cleaned data and finally getting the accuracy of the model is 99.05
Step 8: Sttop

Anaconda navigator is used for execution of the project code. It helps to manage and access notebook files and any kind of python files. By giving the python environment a command prompt can provide easy access to the code and execution. Main tools that need to be installed in the python environment are keras and tensorflow. Minimum of 4GB RAM is required to compile and execute the project code. Preferred operating systems are windows and ubuntu. The above mentioned method is for users using windows OS. Using anaconda navigator software and anaconda prompt that install the necessary modules. To check with the data and accuracy reliability SPSS is used with a default alpha value of 0.05.

3. Results

Mushroom toxicity classification is used to reduce the death caused by the poisonous mushrooms and to save human life.For This purpose Convolutional neural network is compared with knn algorithm.By applying these methods knn is giving significantly better accuracy of 95% than Convolutional neural network .(Berhane)The results are collected by performing multiple iterations of the experiment for identifying different scales of accuracy rate .Further performed the statistical calculations using the SPSS tool and obtained the accuracy from the experimented data and independent sample test is performed

Table1 represents the comparison of accuracy Convolutional neural network and knn(k-nearst neighbour), by iterating in intrusion detection systems for various numbers of times.

Table 2 represents the sample size(N=45), Mean, Standard deviation and Standard error mean are classified based on the accuracy and loss of the data. The accuracy 100% of knn algorithm is significantly higher compared to the convolution neural network

Table 3 represents the significance of the data and standard error difference, where significance of Convolutional neural networkand k nearest neighbour with the confidence interval as 95% and level of significance of 0.05.
Figure 1 represents the analysis of the accuracy of the Convolutional neural network and k nearest neighbour for better improvisation in the mushroom detection system.
Figure 2 represents the comparison of mean accuracy of the convolution neural network. The comparison of accuracy gained. The accuracy of group 1 is 100% and group 2 is 98.45%. The k nearest neighbour has significantly performed better when compared to convolutional neural network.Group 1 appears to produce the most consistent results with its standard deviation ranging from the lower 93's to higher 95's. Group 2 appears to produce the most variable results with its standard deviation ranging from 85's to 90's. There is a significant difference between Convolutional neural networkand k nearest algorithm

4. Discussion

In this study of Mushroom detection systems, the k nearest neighbour has higher accuracy of 100% in comparison to Convolutional neural network 96%. K nearest neighbour has better significance ($p<0.05$) than Convolutional neural network and while using the independent sample t-tests(Goetz 2003; Auerbach, Donner, and Weiss 2008)
Similar work has been carried out by the author (Goetz 2003; Auerbach, Donner, and Weiss 2008). To predict the mushroom toxicity I have used machine learning and gathered all the possible information to predict the outcome value which is the accuracy of the algorithm which we have used in this proposed system (Goetz, Christopher G.2003). In this article they proposed an automatic mushroom toxicity recognition technology and Convolutional neural network from a complex image of the mushroom in MATLAB with accuracy(92. 62%) (Bennett et al.).
In order to overcome the above drawbacks in the proposed system for mushroom toxicity and

classification (Seymour, Tom. 2017). I had developed a working model using machine learning by comparing k-nearest neighbour and convolutionl neural network. The limitations of this study is that a huge color database has to be created manually extracting colors from the mushrooms. The future work is to propose the algorithm for the mushroom classification system using image processing and machine learning (Al-Mejibli, Intisar Shadeed, and Dhafar Hamed Abdel. 2017). In the future we will implement this system on the open cv library with some other algorithms and also perform a performance check of the system design (Wagner, Dennis, Dominik Heider, and Georges Hattab. 2021).

5. Conclusion

Enhanced mushroom detection system using k nearest neighbour is compared with Convolutional neural network with respect to attributes in train datasets with test datasets. SPSS is used to depict the accuracy of 99.05% using k nearest neighbour is more significant than the Convolutional neural network with the accuracy of 98.45%.

Declarations

Conflict of Interests

No conflict of interest in this manuscript.

Author Contribution

Author SYA was involved in data collection, data analysis, manuscript writing. Author TPA was involved in conceptualization, data validation and critical review of manuscript.

Acknowledgement

The authors would like to express their gratitude towards Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences (formerly known as Saveetha University) for providing necessary infrastructure to carry out this work successfully.

Funding

We thank the following organizations for providing financial support that enabled us to complete this study:

RootQuotient Technologies P Ltd, Chennai

Saveetha School of Engineering.

Saveetha University.

Saveetha Institute of Medical Technical Sciences.

6. References:

Al-Mejibli, Intisar Shadeed, and Dhafar Hamed Abd. 2017. "Mushroom Diagnosis Assistance System Based on Machine Learning by Using Mobile Devices." Journal of Al-Qadisiyah for Computer Science and Mathematics. https://doi.org/10.29304/jqcm.2017.9.2.319.

Auerbach, Paul S., Howard J. Donner, and Eric A. Weiss. 2008. "Mushroom Toxicity." Field Guide to Wilderness Medicine. https://doi.org/10.1016/b978-1-4160-4698-1.50046-1.

Dong, Jinhua, and Lixin Zheng. 2019. "Quality Classification of Enoki Mushroom Caps Based on CNN." 2019 IEEE 4th International Conference on Image, Vision and Computing (ICIVC). https://doi.org/10.1109/icivc47709.2019.8981375.

Goetz, Christopher G. 2003. "Mushroom Toxicity." Encyclopedia of the Neurological Sciences. https://doi.org/10.1016/b0-12-226870-9/01195-3.

Ismail, Shuhaida, Amy Rosshaida Zainal, and Aida Mustapha. 2018. "Behavioural Features for Mushroom Classification." 2018 IEEE Symposium on Computer Applications & Industrial Electronics (ISCAIE). https://doi.org/10.1109/iscaie.2018.8405508.

Wagner, Dennis, Dominik Heider, and Georges Hattab. 2021. "Mushroom Data Creation, Curation, and Simulation to Support Classification Tasks." Scientific Reports 11 (1): 8134.Nadya Chitayae, Andi Sunyoto, "Performance Comparison of Mushroom Types Classification Using K-Nearest Neighbor Method and Decision Tree Method", Information and Communications Technology (ICOIACT) 2020 3rd International Conference on, pp. 308-313, 2020.Narumol Chumuang, Kittisak Sukkanchana, Mahasak Ketcham, Worawut Yimyam, Jiragorn Chalermdit, Nawarat Wittayakhom, Patiyuth Pramkeaw, "Mushroom Classification by

Physical Characteristics by Technique of k-Nearest Neighbor", Artificial Intelligence and Natural Language Processing (iSAI-NLP) 2020 15th International Joint Symposium on, pp. 1-6, 2020.Hui Zhao, Fuhua Ge, Pengfei Yu, Haiyan Li, "Identification of Wild Mushroom Based on Ensemble Learning", Big Data and Artificial Intelligence (BDAI) 2021 IEEE 4th International Conference on, pp. 43-47, 2021.

Tables and figures

Table 1.Accuracy table for K Nearest Neighbour and Convolutional neural network the accuracy of Method 1 is 99.05% and Method 2 is 98.45 %.

No.of iterations	Knearest Neighbour Algorithm	Convolutional neural network algorithm
1	99.05	98.45
2	99.01	98.01
3	99.03	98.02
4	99.05	98.03
5	99.07	98.04
6	99.09	98.05
7	99.11	98.06
8	99.13	98.07
9	99.17	98.08
10	99.19	98.09

Table 2. Statistical Analysis of Mean, Standard Deviation, and Standard Error of accuracy of Convolutional neural network and K Nearest Neighbour. There is a statistically significant difference in accuracy between the methods.K Nearest Neighbour has the highest accuracy (100%) and Convolutional neural network(98.45%).

Group	N	Mean	Std.Deviation	Std.Error Mean
Algorithms K Nearest Neighbour	45	99.1850	.29217	.09239
Convolutional neural network	45	98.0900	.12910	.04082

Table 3: Comparison of Significance Level with value p>0.05. Both Convolutional Neural Network and K Nearest Neighbour have a confidence interval of 95% with the significance level of accuracy is >0.05.

		F	sig.	t	df	sig.(2-tailed)	Mean difference	Std.Error Difference	95% Confidence interval of the differenceLower	95% Confidence interval of the difference Upper
Accuracy	Equal variance assumed	1.275	.027	10.841	18	.001	1.09500	.10101	.88279	1.30721
Accuracy	Equal variances not assumed			10.841	12.385	.001	1.09500	.10101	.87568	1.31432

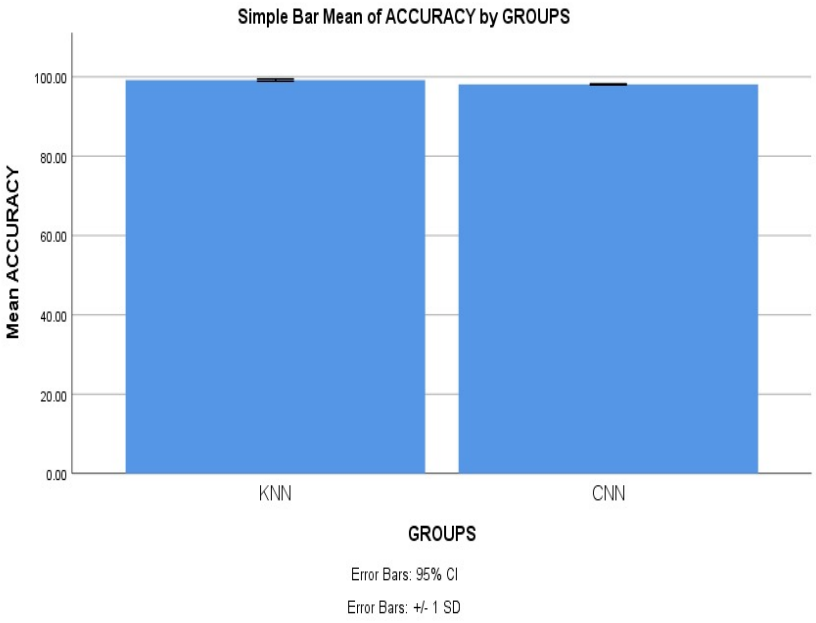

Fig. 1. Comparison of mean accuracy between K-nearest neighbor algorithm over Convolutional neural network algorithm, where the former is better than the latter with 98.45% increase. X axis gives the algorithms and Y Axis: Mean accuracy of detection ± 1 SD.