IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

KRISHAKMITRA (कृषकमित्र) - GRADING OF MUSHROOM

Mannat Doultani[1] , Divya Khiani[2], Richa Bhatia[3] , Roma Bulani[4] , Khushboo Murjani[5]

[1] Assistant Professor, [2,3,4,5] Student
Department of Computer Engineering
Vivekanand Education Society's Institute of Technology
Mumbai, Maharashtra, India.

Abstract -- Mushroom is a type of fungi that is very commonly used in cooking. There are various species of mushrooms that can be found naturally. Most of the mushrooms are poisonous and hence it is very crucial to find out whether the mushroom is edible or not before consuming it. For this purpose, various characteristics of the mushroom can be used as a measure to determine their quality. Machine Learning-Classifying algorithms can be used for this which will help to determine the grade of the mushroom, based on its characteristics. The "A" graded mushroom can then be cultivated by the farmers for a living.

Keywords -- Naive Bayes, Gaussian function, Edibility.

1. INTRODUCTION

Mushrooms are actually wild fungi that were found to be a good delicacy in the early 1650 and hence due to increased consumption of mushrooms people started cultivating it on a large-scale from the business point of view. Sometimes poisonous mushrooms like Conocybe filaris, Webcraps etc. may result in the death of a person, hence determining the grade of the mushroom that one is consuming is very important. Using the Naive Bayes Classification algorithm we can teach the machine to classify the mushrooms into an 'edible' (grade A) and 'poisonous' (grade B) based on its various characteristics. This type of classification will help farmers to cultivate more good quality mushrooms and set their price accordingly. This module is an extension to the previously developed application "KrishakMitra" which is used to predict the type of crop and its variety which will give maximum yield for the farmers field based on the characteristics of the field soil and that location's climatic conditions using Machine Learning algorithms.

2. LITERATURE SURVEY

Mushroom quality is graded using hyperspectral image analysis in the wavelength range of 400 to 1000 nm. Different algorithms are used based on chemometric techniques and image processing methods in the paper[1] An automated system to grade the raw mushrooms and analyze them on the basis of the size of mushrooms and the level of opening cap. Proposed methods calculate the size of shiitake by using area sum of shiitake area and the rate of lamella areas is presented in this paper.[2] Support Vector Machine and Naive Bayes algorithms are used for the classification of mushrooms. Expert system is developed to classify the mushrooms on the basis of their characteristics. In the studied paper, performances of both algorithms are evaluated on mushroom data in fold cross-validation.[3] The classification of poisonous mushrooms is determined by three classification algorithms namely; Decision tree, Naive Bayes and Support vector machine. The Agaricus and Lepiota family is taken into consideration for mushroom data in the proposed paper. [4]

3. DATASET

class	cap-shape	cap-surface	cap-color	bruises	odor	gill-attachment	gill-spacing
p	x	s	n	t	p	f	c
e	x	s	y	t	a	f	c
e	b	s	w	t	l	f	c
p	x	y	w	t	p	f	c
e	x	s	g	f	n	f	w

Fig 4 : Snapshot of the dataset used

The Dataset consists of 23 columns out of which the first column describes the class of the mushroom which can be either edible or poisonous. Rest 22 columns describes various characteristics of the mushroom based on which the mushroom can be categorized into edible or poisonous categories. The characteristics included in the dataset are cap-shape, cap-surface, cap-color, bruises, odor, gill-attachment, gill-spacing, gill-size, gill-color, stalk-shape, stalk-root, stalk-surface-above-ring, stalk-surface-below-ring, stalk-color-above-ring, stalk-color-below-ring, veil-type, veil-color, ring-number, ring-type, spore-print-color, population, habitat. The dataset was obtained from www.kaggle.com.

4. IMPLEMENTATION

The module for grading uses Gaussian Naive Bayes to classify the mushroom quality. The module takes the characteristics of the mushroom as input and grades them accordingly. Some of the characteristics on which the mushrooms are graded as follows :

1. Cap shape, surface, color
2. Gill size, attachment, color
3. Veil type, color
4. Ring number, type
5. Stalk shape, the color above and below the ring
6. Odor,etc.

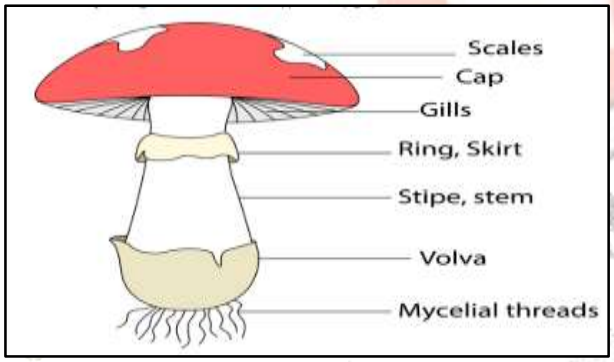

Fig 1 : Mushroom and its parts

```
G:\BE PROJECT(GROUP NO 16)>python naivetry.py
Enter Grades
'x'
's'
'n'
't'
'p'
'f'
'c'
'n'
'k'
'e'
'e'
's'
's'
'w'
'w'
'p'
'w'
'o'
'p'
'k'
's'
'u'
('Predicted Value:', 1)
```

Fig 2 : Grading of mushroom

5. EXPLANATION

Based on the various characteristics of the mushroom they can be classified into edible or poisonous. Considering 22 characteristics of mushroom like cap-shape, gill size, veil type, ring number, stalk shape etc and applying Machine Learning algorithm (Naive Bayes) we can predict whether the mushroom is safe to consume or poisonous. Each characteristic can have more than 1 value and they are represented as abbreviations. For e.g For cap-shape the values can be b/c/x/f/k/s each representing a particular shape like bell=b, conical=c, convex=x, flat=f, knobbed=k, sunken=s. After taking the input and applying the Algorithm we predict whether mushroom is edible and safe to be consumed or is poisonous. The output value **'1'** represents **'edible'** whereas **'0'** represents **'poisonous'**.

```
G:\BE PROJECT(GROUP NO 16)>PYTHON try1.py
('accuracy', 0.9453345900094251)
```

Fig 3 : Accuracy of Algorithm

6. COMPARISON OF VARIOUS ALGORITHMS

Algorithm	Bernoulli NB	Multinomial NB	Gaussian NB
Definition	The parameters that we use to predict the class variable take up only values yes or no i.e Bernoulli assumes that each predictor is binary valued.	The predictors used by this type of classifier is the frequency of a particular value of a particular characteristic in the whole dataset.	This type of classifier is useful when the values of the characteristics are assumed to be distributed normally and we can assume that these values are sampled from a gaussian distribution
Accuracy for System	79.68%	65.64%	94.53%

```
G:\BE PROJECT(GROUP NO 16)>python bernoullinb.py
('accuracy', 0.79688972667295)
```

Fig 5

```
G:\BE PROJECT(GROUP NO 16)>python multinomialnaive.py
('accuracy', 0.6564561734213007)
```

Fig 6

```
G:\BE PROJECT(GROUP NO 16)>PYTHON try1.py
('accuracy', 0.9453345900094251)
```

Fig 4

The above figures are the snapshots of cmd depicting accuracy of various Naive Bayes algorithms.

7. GAUSSIAN NAIVE BAYES FUNCTION

There are various extensions to the NB algorithm. The Gaussian algorithm is the easiest one to use because it only requires to estimate the mean and standard deviation from your training data. After calculating the probabilities using the following formula for input values for each class using their respective frequencies we can calculate the mean and standard deviation of input values (x) for each class to summarize the distribution.

$$\mathbf{MAP(h) = max(P(d|h) * P(h))}$$

Probabilities of new x values are calculated using the Gaussian Probability Density Function (PDF).
When making predictions these parameters can be plugged into the Gaussian PDF with new input for the variable, and in return, the Gaussian PDF will provide an estimate of the probability of that new input value for that class.

$$\mathbf{pdf(x, mean, sd) = (1 / (sqrt(2 * PI) * sd)) * exp(-((x-mean^2)/(2*sd^2)))}$$

Where pdf(x) is the Gaussian PDF, sqrt() is the square root, mean and sd are the mean and standard deviation calculated above, PI is the numerical constant, exp() is the numerical constant e or Euler's number raised to power and x is the input value for the input variable.
We are taking into consideration 22 characteristics of mushrooms to determine whether they are edible or poisonous. We train the algorithm using 80% of the available data and test it using the remaining 20% test data. We were able to achieve an accuracy of about 94%.

Using Android Studio , we have developed an application which takes characteristics as input and produces an output which shows whether the mushroom is of '**A**' quality or '**B**' quality.
This page takes input about the cap of Mushroom using Spinner
Cap Shape
Cap Size

Cap Color

Fig 7 : Screenshots of Developed App

Fig 8 : Final grade of Mushroom being displayed

Here, we have shown the input characteristics of CAP feature. Likewise, characteristics of all the other features are taken using the drop down menu. Using all these inputs to 22 characteristics we classify whether the mushroom is edible or poisonous. Fig. 7 represents the Cap features of the mushroom. Fig. 8 represents the final result generated by the app which displays the grade of the mushroom i.e grade '**A**' or grade '**B**'.

6. FUTURE SCOPE

The grading module can be extended for the selling of crops that are produced by the farmers. This will eliminate involvement of middlemen and their commission and will ensure that farmers get a reasonable price for the produce. After the crops are graded and its price is set by the farmer, people can bid online for the crop and the farmer can sell the crop to the highest bidder and ensure that he gains maximum profit.
Also, there are chances that farmers may be entering false data to ensure that he gets more price for its produce than the original quality of the crop. This issue can be solved using Image Processing so that the buyer can rely on the grade given to the crop by the app.

7. CONCLUSION

This paper presents a new application of the Naive Bayes algorithm. In this, NB algorithm is used for the classification of training data and predicting the class of mushrooms based on the classification model built by NB on training process .Experiments conducted on the mushrooms data shows that Naive Bayes gives accuracy of 94.5% as shown in fig 3 above. The performance of the NB classifier is evaluated using the in-built accuracy function. The results obtained shows that Naive Bayes is an efficient algorithm for this application domain.

8. ACKNOWLEDGMENT

We would like to acknowledge our Project guide Mrs Mannat Doultani for providing necessary guidance for writing the research paper.

9. REFERENCES

[1] A.A.Gowen, C.P. O'Donnell, J.M. Frias, G. Downey "Hyperspectral imaging for mushrooms (Agaricus bisporus) quality monitoring" Innovative Food Science and Emerging Technologies 11 (2010) 423–431.

[2] Takeshi Hazisawa, Masashi Toda, Teruvasu Sakoil, Kazuhiro Matumural, Masahito Fukuda "Image analysis method for grading raw shiitake mushrooms" FCV 2013 - Proceedings of the 19th Korea-Japan Joint Workshop on Frontiers of Computer Vision. 46-52. 10.1109/FCV.2013.6485458.

[3] Prasad Babu, Rajani Thommandru, K.Swapna, E.Nilima "Development of the mushroom expert system on SVM classifier and Naive Bayes classifier" International Journal of Computer Science and Mobile Computing - IJCSMC (Vol.3, No. 4) 1328-1335.

[4] Agung Wibowo, Yuri Rahayu, Andi Riyanto, Taufik Hidayatulloh "Classification algorithm for edible mushroom identification" 250-253. 10.1109/ICOIACT.2018.8350746. .

[5] Ms. Nirali A. Kansagara, Ms.Trupti M. Khurape, Ms. Jyoti S. Kamble, Ms. Manas M. Kulkarni, Prof.Mr.G.I.Rathod "An Android Application for Online Agri-Auction" International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395 -0056 Volume: 03 Issue: 02 | Feb-2016.

[6] Kamran, M. & Anjum, Maria & Rehman, Mariam & Ahmad, H. & Kamran, Muhammad. (2016). Classification of Information Systems in E-agriculture: A Mapping Study. International Journal of Computer Science and Information Security. 14.

[7] Chirag Shah, Aryan Yadav, Jeevesh Pachouri, Prakash Choudhary "AGROMART - The Market for Farmers".

[8] K.C., Saban & Timalsina, Arun. (2016). Challenges for Adopting E-Commerce in Agriculture in Nepalese Context — a Case Study of Kathmandu Valley.