Original Article

Prediction of Insufficient Accuracy for Mushroom Classification whether Poisonous or Eatable Food using Random Forest Training by comparing Logistic Regression to Improve Accuracy

Maheswara Reddy[1], Saravanan.M.S[2]

[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]Project Guide,Department of Computer Science and Engineering,Saveetha School of Engineering,Saveetha Institute of Medical and Technical Sciences,Saveetha University, Chennai, Tamilnadu, India. Pincode:602105.

Abstract

Aim: Mushrooms have a spot with development type and they contain principal supplements, for instance, proteins, nutrients, cell reinforcements, antioxidants and amino acids. There are a huge load of benefits of mushrooms. A wide range of mushrooms are not eatable. So before eating up mushrooms, it should be checked for consumable mushrooms. Careful affirmation and fitting distinctive confirmation of species are the super protected way to deal with ensuring not eatable mushrooms, and safeguard against expected accidents of consuming poisonous one.
Materials and Methods:The survey used 51 models with two get togethers of computations with the G-power worth of 80% and the Mushroom were accumulated from various web sources with late audit disclosures and edge 0.05%, sureness length 90% mean and standard deviation. To expect the Mushroom precision rate for at present the Logistic Regression computation has found 90.57% of accuracy, therefore this study needs to find better accuracy for accuracy decrease prediction with the novel Random Forest Training Machine Learning estimation.
Result: This investigation found 94.64% of accuracy for poisonous distinguishing proof using the Random Forest Training estimation with a statistically significant difference between the two groups ($p=0.047$; $p<0.05$) with 95% sureness stretch.
Conclusion:This investigation found 92.82% of accuracy for poisonous recognizable proof using the Random Forest Training estimation with a basic worth of two followed tests is 0.001($p<0.05$) with 95% conviction stretch.
Keywords: Machine Learning, Data Mining, Novel Random Forest Training, Logistic Regression, Poisonous, Mushrooms.

DOI: 10.47750/pnr.2022.13.S04.040

INTRODUCTION

Every year, mushroom poisoning kills or affects a large number of individuals all around the world. A person's most valuable asset is his or her health. Keeping this in mind, we've put together a quick overview of mushroom poisoning research. The term mushrooms macrofungi or Macromycetes has been differently depicted by several creators. They are the significant, spore bearing fruiting mixes of advancements, which ordinarily show up over the ground after storms. They generally fill in backwoods regions, seeing somebody on woody pieces of trees either as parasite, saprophyte or as symbionts in the dirt (Engel 2019). Mushrooms and parasites as a rule are non-green creatures lacking chlorophyll (Engel 2019; Sheldrake 2020). They can't make their own food from fundamental inorganic materials, like water, carbon dioxide, and nitrates, utilizing energy from the sun. The indications of mushroom harming can waver from slight gastro gastrointestinal inconvenience (Dye et al. 2017) , disgorging to death ("Death in Medieval Europe" 2016). If we eat our favourite mushroom without first determining if it is edible or Poisonous, we risk death ("Death in Medieval Europe" 2016; Engel 2019; Sheldrake 2020).

Mushrooms are less Eatable than different plants because of the way that their cell dividers are generally made out of the polysaccharide chitin impervious to the activity of stomach related liquids in the human creature. It could be utilized to sort styles or to expect values out of your data. Since it utilizes a managed concentrating on strategy, it requires your data to fuse objectives for training locally. Order is one of the commitments achieved through the data mining measure. It is named one of the expectation procedures for spider secret data. A form created through data schooling must have the capacity to separate the class of a fresh out of the box new data mushroom in those families (National Academies of Sciences, Engineering, and Medicine et al. 2018). Mushroom, Agaricaceae, has a place with members of a circle of family members of growths with gills (National Academies of Sciences, Engineering, and Medicine et al. 2018; Ponder, Lindberg, and Ponder 2019). It is continually ordered into two groups; appropriate for eatable and poisonous. The time-frame mushroom is continually used to counsel appropriate for eatable species, even as the time-frame toadstool is utilized for harm plausible. In any case, the toadstool is continually changed through a method for poisonous mushrooms (Gogoi, Rathaiah, and Borah 2019). This research related work was presented and published in more than 41 indexed journals.Our team has extensive knowledge and research experience that has translate into high quality publications (Bhansali et al. 2021; Jayanth et al. 2021; Sudhakar, Ravel, and Perumal 2021; Sathiyamoorthi et al. 2021; Deepanraj et al. 2021; Raju et al. 2021; Arun Prakash et al. 2020; Kamath et al. 2020; Shanmugam et al. 2021; Rajasekaran et al. 2020; Adhinarayanan et al. 2020; Rajesh et al. 2020; Aurtherson et al. 2021)

The research gap identified is that several works have demonstrated that the performance Logistic Regression provides less accuracy in prediction of mushroom. Hence the aim of this study is to increase the accuracy of predicting the likelihood of poisonous or eatable mushrooms and improve the prediction model using the Novel Random Forest Training.

MATERIALS AND METHODS

This investigation study was finished at the Machine Learning Laboratory, Department of Computer Science and Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Science, Chennai. This research study uses two groups of classification algorithms used for the study. Social events Group 1 and Group 2 are the Random Forest Training estimation and the Logistic Regression computation respectively (Muff, Signer, and Fieberg2020). Every model size was expected using the G-power gadget with version 3.1.10 and achieving 51 model sizes with 80% of G-power regards and the cut off two followed basic worth is set to 0.05 and the assurance stretch as 95%. The mushroom dataset which is to be credited for the proposed work is assembled from kaggle.com (Muff, Signer, and Fieberg 2020; Satapathy et al. 2016), one of the more notable online organizations for data specialists and Data Miningspecialists. It licenses customers to look and find changed datasets that they require.

The Mushroom harmful or not dataset is to be imputed for the proposed work is collected from kaggle.com, one of the more popular online communities for data scientists and machine learning practitioners. It allows users to search and find different datasets that they require. The dataset used here consists of 14 attributes and contains 12 features that can be used to predict mortality by harmful or not mushrooms. The mushrooms dataset has 305 rows which consists of data for the symptoms that are related to poisonous which also includes duplicate, null and missing values.

Novel Random Forest Training (RF) falls under the category of supervised algorithms. Random Forest Trainings can be used for both classification and Logistic regression. In the classification RF the decision variable is categorical. For implementing a RF, the cost function used to evaluate the binary splits called the Gini Index should be calculated. The split creation is done with help of calculating the gini score, splitting the dataset and evaluating the all splits and this is done by creating a root node. It can start building the tree by first deciding when to stop 3 of 12 the growth of the tree by creating the terminal node by setting the maximum tree depth and second, by using recursive splitting. The terminal node is used for the final prediction and recursive splitting is a method used to build the tree.

The help vector machine has been picked on the grounds that it addresses a structure both fascinating from an Data Mining point of view and from an inserted frameworks viewpoint. A Random Forest Training (RF) is a straight or non-direct classifier, which is a numerical capacity that can recognize two various types of items.

Pseudocode of the Random Forest Training algorithm:

```
Inputs : Determine the various training and test data.
Outputs: Determine the calculated accuracy.
Select the optimal value of eatable and poisonous for RF.
while(stopping condition is not met )do
        Implement RF train steps for each data point.
        Implement RF clasity for testing data points.
end while
Return accuracy
```

Logistic Regression has been chosen because it represents a framework both interesting from an Machine Learning perspective and from an embedded systems perspective. A Logistic Regression is a linear or nonlinear classifier, which is a mathematical function that can distinguish two different kinds of objects.

Pseudocode of the Logistic Regression algorithm

```
Step1: Initialize all parameters (Bo. B1, etc.)
Step 2: Calculate (predict) dependent variable (h(x))
Step 3: Calculate poisonous function (poisonous(ha(z), y))
Step 4: Calculate gradient for poisonous function
Step 5: Update all parameters
Step 6: Repeat steps 2 to 5.
```

STATISTICAL ANALYSIS

Finally, the results assembled from the get-together one and get-together two estimations will be applied using Statistical Package for Social Sciences (SPSS) with variation 26. The independent model t-test was performed to break down the shows of the Logistic Regression and novel Random Forest Training computations. The independent variable is the forest attribute in the dataset and other twenty attributes such as odor, population, bruises etc are dependent variables for our study for mushroom prediction.

RESULTS

The Random Forest algorithm and Logistic Regression algorithm are compared with 51 samples by applying various 70% of training and 30% of testing datasets by varying the number of records of the dataset and the outcomes are depicted in Table 1 and the dataset consists of 100 rows where the accuracy of both the RF and LR algorithm are obtained for 10 iterations. It is observed that the Novel Random Forest algorithm is significantly better than the Logistic Regression algorithm.

From Table 1, the statistical analysis of RF Algorithm and LR. Mean accuracy value, Standard deviation and accuracy for RF and LR algorithms are obtained for 10 iterations. It is observed that the RF algorithm performed better than the LR algorithm.

From Table 2, the Independent sample t-test for significance and standard error are determined. There exists a statistically significant difference between the two groups ($p=0.047$; $p<0.05$) with 95% confidence intervals were calculated. By statistical analysis RF gives 1.46 standard deviation with 0.46 standard error and LR gives 2.91 standard deviations with 0.92 standard error with a significant value less than 0.5 that shows the proposed work gives better results.

In Table 3, observed after performing statistical analysis of 10 samples, RF obtained 3.24 standard deviations with 1.03 standard error while the LR algorithm obtained 3.24 standard deviations with 1.03 standard error. There exists a statistically significant difference between the two groups ($p=0.047$; $p<0.05$) showed our hypothesis holds a good value. For changes in the input values (independent variable) the corresponding output values are changed (dependent variables).

The independent sample t-test was used to compare the accuracy of two algorithms and a statistically significant difference was noticed with $p < 0.05$. The RF model obtained 89.6% accuracy. When compared with the other algorithm's performance, the proposed RF classifier achieved better performance than LR.

The Red dots shown poisonous mushrooms and Green dots shown its eatable mushrooms. The symptom shown here is the rest of the models accuracy comparison as shown in Fig. 1.

The Red dots showed poisonous mushrooms and Green dots showed its eatable mushrooms as shown in Fig. 2.

The data taken for this research study will undergo preprocessing, a step to remove the unwanted, missing values and null values, it has been preprocessed. The results of mushroom analysis classification has accuracy for RF algorithm having 89.6% and LR having 86.9% across the samples as shown in Fig. 3, and it shows that RF is significantly better than the LR algorithm.

Finally, compare the mean accuracy of the Novel Random Forest and the Logistic Regression algorithm as shown in Fig. 4, and it shows that the Novel Random Forest is significantly better than the Logistic Regression algorithm.

DISCUSSION

In this article, we proposed a modified mushroom destructiveness unmistakable evidence method. The Random Forest Training strategy proposed by Zhou has an affirmation precision of more than 98%. Taking into account the incredible precision essentials, we separated three model plan models. Vital backslide yielded collecting results by looking at the reasonable components essential to perceive noxiousness. The exactness of the Random Forest Training technique is better than that of determined backslide. Appeared differently in relation to Random Forest Training and determined backslide, Novel Random Forest Training achieved better results similar to ID precision. Therefore, Random Forest Training is a nice technique for normally recognizing whether a mushroom is not edible (Ren et al. 2020) .

Distinctive ordinary mushroom harmfulness insistence systems are at this point being used (Suetsugu, Okamoto, and Kato 2019). These techniques utilize various responsibilities regarding picking harmfulness, yet they have distinctive cutoff focuses, like low accuracy, unsatisfactory affirmation of dim poisonous substances, the essential for a super exploratory climate, and adequate expert information and complex test testing procedures. To stay away from the limitations of these frameworks and apply them to minimal model information appraisal, we utilized Data Mining. Maybe than critical neural affiliations, which require mind blowing exertion in hyperparameter tuning, Random Forest Training is considerably less hard to prepare and can be applied to various kinds of information in various districts. The Novel Random Forest Training assessment participates in the going with benefits: (1) it has a major arrangement; (2) it will overall be applied to datasets of various sizes; (3) the testing systems and managing are immediate; and (4) for our fundamentals, mushroom destructiveness is seen rapidly (Rockefeller 2015). Part based learning and iterative classifiers in the Random Forest Training method have the best show among the three frameworks for Machine Learning utilized in this article. This altered ID methodology is reasonable for non professional obvious affirmation and for dull mushroom assortments (Burtsev, Shagdurov, and Demkin 2019).

Among the three Machine Learning methodologies, Random Forest Training yielded the best precision. Regardless, the vigor of the classifier should be improved. A diversion for this good might come from the difficulty of parts, the part checked stem surface above the ring, or it might come from the genuine assessment. In like way, further making dependability is a first concern when attempting to manage the accuracy of the classifier. Since Novel Random Forest Training can be utilized for various types of datasets and certification dependent upon picture highlights is more helpful than with different classifiers, this procedure for seeing whether a mushroom is harmful can be reached with picture insistence. In any case, there is now no dataset of mushroom pictures (Sin 2020).

CONCLUSION

The main aim of the study is to measure the accuracy in the classification of poisonous or eatable mushrooms. This research study applied the RF algorithm for detection of mushrooms harmful or not achieved and has been compared with the LR algorithm. The results obtained show that the Random Forest Training algorithm has found 92.86% of accuracy on mushrooms data than the 90.37% of the LR algorithm.

DECLARATIONS

Conflict of Interests

No conflict of interest in this manuscript.

Authors Contributions

Author TMR was involved in data collection, data analysis, and manuscript writing. Author SMS was involved in conceptualization, data validation, and critical review of the manuscripts.

Acknowledgements

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

Funding

We thank the following organizations for providing financial support that enabled us to complete the study.
1. AAIC Technologies Private Limited, Hyderabad.
2. Saveetha University.
3. Saveetha Institute of Medical and Technical Sciences.
4. Saveetha School of Engineering.

REFERENCES

1. Adhinarayanan, Rajesh, AravindhRamakrishnan, Gopal Kaliyaperumal, Melvinvíctor De Poures, Rajesh Kumar Babu, and DamodharanDillikannan. 2020. "Comparative Analysis on the Effect of 1-Decanol and Di-N-Butyl Ether as Additive with diesel/LDPE Blends in Compression Ignition Engine." *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects*, June, 1–18.
2. Arun Prakash, V. R., J. Francis Xavier, G. Ramesh, T. Maridurai, K. Siva Kumar, and R. Blessing Sam Raj. 2020. "Mechanical, Thermal and Fatigue Behaviour of Surface-Treated Novel Caryota Urens Fibre–reinforced Epoxy Composite." *Biomass Conversion and Biorefinery*, August. https://doi.org/10.1007/s13399-020-00938-0.
3. Aurtherson, P. Babu, Bhanu Teja Nalla, Karthikeyan Srinivasan, Kulmani Mehar, and Yuvarajan Devarajan. 2021. "Biofuel Production from Novel Prunus Domestica Kernel Oil: Process Optimization Technique." *Biomass Conversion and Biorefinery*, May. https://doi.org/10.1007/s13399-021-01551-5.
4. Bhansali, Karan J., Kamlesh R. Balinge, Subodh U. Raut, Shubham A. Deshmukh, M. Senthil Kumar, C. Ramesh Kumar, and Pundlik R. Bhagat. 2021. "Visible Light Assisted Sulfonic Acid-Functionalized Porphyrin Comprising Benzimidazolium Moiety for Photocatalytic Transesterification of Castor Oil." *Fuel* 304 (November): 121490.
5. Burtsev, N. I., V. C. Shagdurov, and I. O. Demkin. 2019. "Application of the Random Forest Machine Learning Algorithm for Recognizing Patient Arm Movements While Using a Bionic Prosthesis." *XIV RUSSIAN-GERMANY CONFERENCE ON BIOMEDICAL ENGINEERING (RGC-2019)*. https://doi.org/10.1063/1.5121935.
6. "Death in Medieval Europe." 2016. https://doi.org/10.4324/9781315466859.
7. Deepanraj, B., N. Senthilkumar, D. Mala, and A. Sathiamourthy. 2021. "Cashew Nut Shell Liquid as Alternate Fuel for CI Engine—optimization Approach for Performance Improvement." *Biomass Conversion and Biorefinery*, February. https://doi.org/10.1007/s13399-021-01312-4.
8. Dye, Leslie R., Christine Murphy, Diane P. Calello, Michael D. Levine, and Aaron Skolnik. 2017. *Case Studies in Medical Toxicology: From the American College of Medical Toxicology*. Springer.
9. Engel, Michael S. 2019. "Termite Evolution: A Primal Knock on Wood or a Hearty Mouthful of Dirt." *Current Biology: CB*.
10. Gogoi, R., Y. Rathaiah, and T. R. Borah. 2019. *Mushroom Cultivation Technology*. Scientific Publishers.
11. Jayanth, BellappuVenkat, Melvin Victor Depoures, Gopal Kaliyaperumal, DamodharanDillikannan, DilipsinghJawahar, KumaranPalani, and Ganesha Prasad MeravanigeeShivappa. 2021. "A Comprehensive Study on the Effects of Multiple Injection Strategies and Exhaust Gas Recirculation on Diesel Engine Characteristics That Utilize Waste High Density Polyethylene Oil." *Energy Sources, Part A: Recovery,*

Utilization, and Environmental Effects, June, 1–18.
12. Kamath, Manjunath, Subha Krishna Rao, Jaison, Sridhar, Kasthuri, Gopinath, Sivaperumal, and Shantanu Patil. 2020. "Melatonin Delivery from PCL Scaffold Enhances Glycosaminoglycans Deposition in Human Chondrocytes – Bioactive Scaffold Model for Cartilage Regeneration." *Process Biochemistry* 99 (December): 36–47.
13. Muff, Stefanie, Johannes Signer, and John Fieberg. 2020. "Accounting for Individual-Specific Variation in Habitat-Selection Studies: Efficient Estimation of Mixed-Effects Models Using Bayesian or Frequentist Computation." *The Journal of Animal Ecology* 89 (1): 80–92.
14. National Academies of Sciences, Engineering, and Medicine, Division of Behavioral and Social Sciences and Education, Board on Science Education, Board on Behavioral, Cognitive, and Sensory Sciences, and Committee on How People Learn II: The Science and Practice of Learning. 2018. *How People Learn II: Learners, Contexts, and Cultures*. National Academies Press.
15. Ponder, Winston Frank, David R. Lindberg, and Juliet Mary Ponder. 2019. *Biology and Evolution of the Mollusca, Volume 1*. CRC Press.
16. Rajasekaran, S., D. Damodharan, K. Gopal, B. Rajesh Kumar, and Melvin Victor De Poures. 2020. "Collective Influence of 1-Decanol Addition, Injection Pressure and EGR on Diesel Engine Characteristics Fueled with diesel/LDPE Oil Blends." *Fuel* 277 (October): 118166.
17. Rajesh, A., K. Gopal, De Poures Melvin Victor, B. Rajesh Kumar, A. P. Sathiyagnanam, and D. Damodharan. 2020. "Effect of Anisole Addition to Waste Cooking Oil Methyl Ester on Combustion, Emission and Performance Characteristics of a DI Diesel Engine without Any Modifications." *Fuel* 278 (October): 118315.
18. Raju, P., K. Raja, K. Lingadurai, T. Maridurai, and S. C. Prasanna. 2021. "Glass/Caryota Urens Hybridized Fibre-Reinforced nanoclay/SiC Toughened Epoxy Hybrid Composite: Mechanical, Drop Load Impact, Hydrophobicity and Fatigue Behaviour." *Biomass Conversion and Biorefinery*, March. https://doi.org/10.1007/s13399-021-01427-8.
19. Ren, Dian, Nathanael Aubert-Kato, Emi Anzai, Yuji Ohta, and Julien Tripette. 2020. "Random Forest Algorithms for Recognizing Daily Life Activities Using Plantar Pressure Information: A Smart-Shoe Study." *PeerJ*. https://doi.org/10.7717/peerj.10170.
20. Rockefeller, J. D. 2015. *The Guide in Growing and Harvesting The Best Mushrooms in Your Own Home*. J.D. Rockefeller.
21. Satapathy, Suresh Chandra, Yogesh Chandra Bhatt, Amit Joshi, and Durgesh Kumar Mishra. 2016. *Proceedings of the International Congress on Information and Communication Technology: ICICT 2015, Volume 1*. Springer.
22. Sathiyamoorthi, Ramalingam, Gomathinayakam Sankaranarayanan, Dinesh Babu Munuswamy, and Yuvarajan Devarajan. 2021. "Experimental Study of Spray Analysis for Palmarosa Biodiesel-diesel Blends in a Constant Volume Chamber." *Environmental Progress & Sustainable Energy* 40 (6). https://doi.org/10.1002/ep.13696.
23. Shanmugam, Rajasekaran, DamodharanDillikannan, Gopal Kaliyaperumal, Melvin Victor De Poures, and Rajesh Kumar Babu. 2021. "A Comprehensive Study on the Effects of 1-Decanol, Compression Ratio and Exhaust Gas Recirculation on Diesel Engine Characteristics Powered with Low Density Polyethylene Oil." *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects* 43 (23): 3064–81.
24. Sheldrake, Merlin. 2020. *Entangled Life: The Phenomenal Sunday Times Bestseller Exploring How Fungi Make Our Worlds, Change Our Minds and Shape Our Futures*. Random House.
25. Sin, Julie. 2020. "Recognizing Whether You Are in Simple, Complicated, or Complex Problem Territory." *Commissioning and a Population Approach to Health Services Decision-Making*. https://doi.org/10.1093/oso/9780198840732.003.0012.
26. Sudhakar, M. P., Merlyn Ravel, and K. Perumal. 2021. "Pretreatment and Process Optimization of Bioethanol Production from Spent Biomass of GanodermaLucidum Using Saccharomyces Cerevisiae." *Fuel* 306 (December): 121680.
27. Suetsugu, Kenji, Tomoko Okamoto, and Makoto Kato. 2019. "Mushroom Attracts Hornets for Spore Dispersal by a Distinctive Yeasty Scent." *Ecology*. https://doi.org/10.1002/ecy.2718.

TABLES AND FIGURES

Table 1. Comparison between Random Forest and Logistic Regression algorithm with N=10 samples of the dataset with the highest performance of 92.84% and 90.57% in the sample (when N=1) using the dataset size = 100 and the 70% of training and 30% of testing data.

Sample (N)	Dataset size / Rows in %	Random Forest algorithm Accuracy in %	Logistic Regression algorithm Accuracy in %
1	8125	92.84	90.57
2	7613	92.23	90.12
3	7041	91.87	89.76
4	6582	91.32	89.45
5	5893	91.01	89.12
6	4931	90.73	88.67
7	4078	90.21	88.12
8	3421	89.45	87.67

9	2067	89.04	87.09
10	1924	88.34	86.54

Table 2. Statistical results of Random Forest and Logistic Regression algorithms. Mean accuracy value, standard deviation and standard error mean for Random Forest and Logistic Regression algorithms are obtained for 10 iterations. It is observed that the Random Forest (90.70%) algorithm performed better than the Logistic Regression (88.37%) algorithm.

Algorithms (Accuracy)	Sample (N)	Mean	Standard Deviation	Standard Error Mean
Random Forest algorithm	10	90.70	1.44	0.458
Logistic Regression algorithm	10	88.37	1.53	0.485

Table 3. The Independent sample t-test of the significance level Random Forest and Logistic Regression algorithms results with a statistically significant difference between the two groups (p=0.047; p<0.05).. Therefore both the Random Forest and the Logistic Regression algorithms have a significance level less than 0.05 with a 95% confidence interval.

Accuracy	Levene's Test for Equality of Variances		T-test of Equality of Means					95% of the confidence interval of the Difference	
	F	Sig.	t	df	Sig (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Equal Variance Assumed	0.116	0.047	3.489	18	0.001	2.32	0.66	0.926	3.729
Equal Variance Not Assumed	-	-	3.489	17.94	0.001	2.32	0.66	0.925	3.730

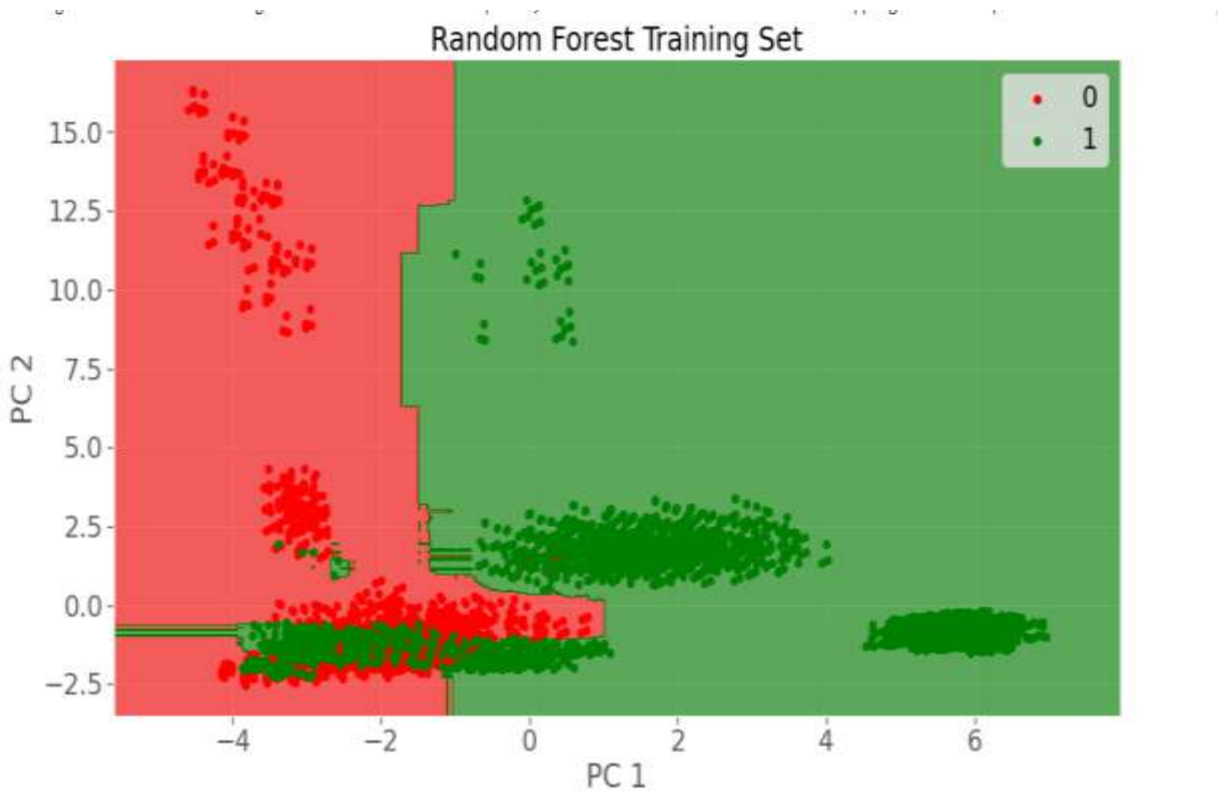

Fig. 1. Comparison of accuracies of Random Forest Training (Poisonous in Red dots and Eatable in Green dots) across the samples

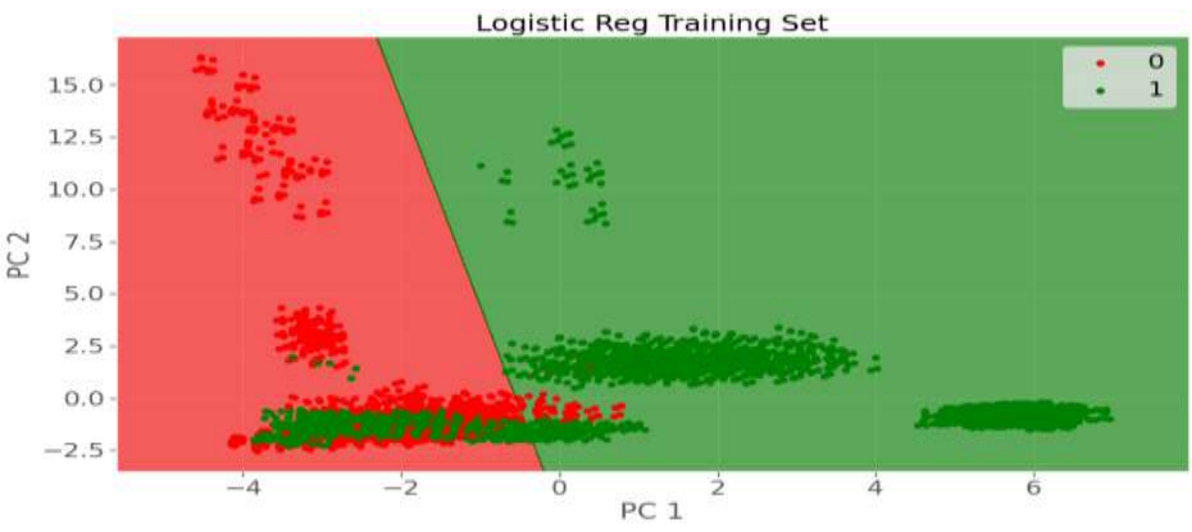

Fig. 2. Comparison of accuracies of Logistic Regression (Poisonous in Red dots and Eatable in Green dots) across the samples.

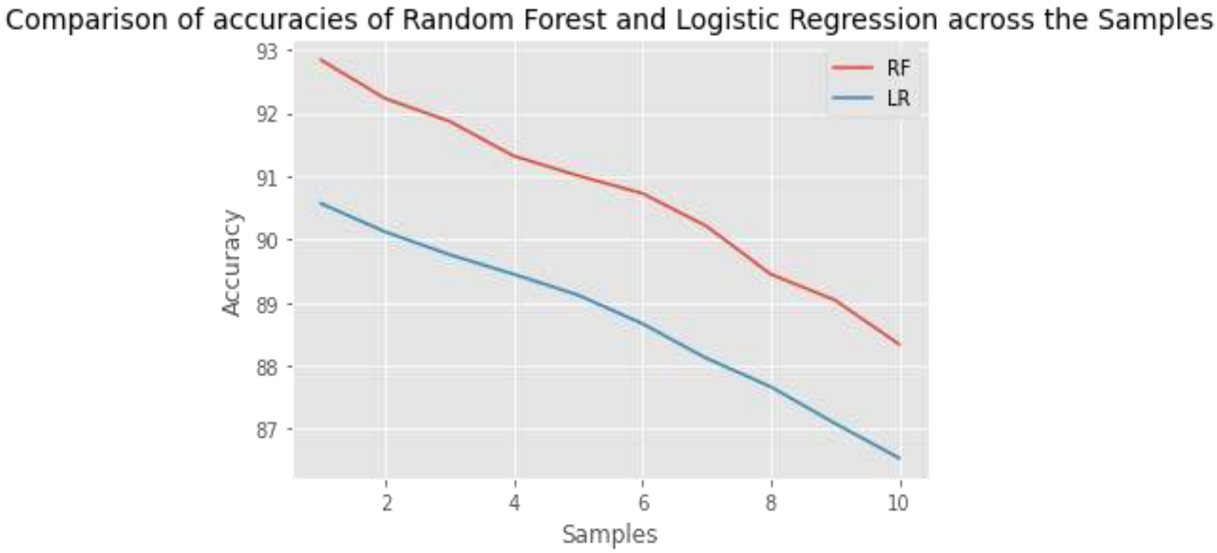

Fig. 3. Result of eatable mushroom classification using RF (92.84) and LR (90.57) across the samples.

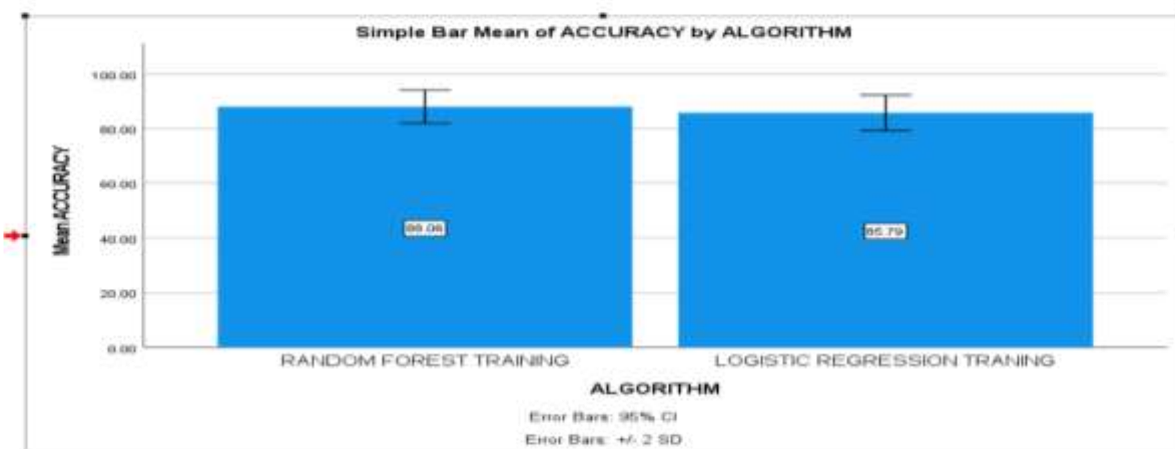

Fig. 4. Comparison of Random Forest Training algorithm and Logistic Regression in terms of mean accuracy. The mean accuracy of Random Forest Training is better than Logistic Regression and the standard deviation of Random Forest Training is slightly better than Logistic Regression algorithm. X-axis: (GROUPS) Random Forest Training vs Logistic Regression algorithm and Y axis: Mean accuracy of prediction ±2 SD.