Performance Comparison of Machine Learning Algorithms for Crop Seed Detection. An Efficiency Analysis

Mr. Abhishek Singh[1], , Dr. K.P Yadav[2],Dr. Sunil Kumar[3], Dr. Harish Nagar[4]

[1]Research Scholar CSE, Sangam University, Bhilwara, Rajasthan, India

[2] President SangamUniversity Margadarshak :AICTE (Govt. of India), Sangam University, Bhilwara, Rajasthan, India

[3]Department of Computer Science, Sangam University, Bhilwara, Rajasthan, India

[4]Department of Applied Science, Sangam University, Bhilwara, Rajasthan, India

Abstract:

In the current generation of Artificial Intelligence ML (Machine Learning)is used for identification and detecting crop major problems. However, our work is focused on the identification of seed with accuracy. Because crop production depends on the seed. For that, identification of seed according to currently soil efficacy and weather condition is the primary need. In this paper, we take some grains with 7 different parameters and analysis of the grain. The kernel of grain is used for finding differences (detail taken from Govt. resource).In conclusion, we find the best grain for a particular soil and an algorithm which gives better prediction in this condition. For the ML algorithm, we use Logistic Regression, SVC, Decision Tree, KNN, and Random Forest for comparison purposes.

Key Word: Artificial Intelligence, Machine Learning, Grain, Algorithm.

Introduction: Wheat is one of the major grain form wheat, rice, and corn and Grain cereals such as wheat, barley, rice, and maize are the nutritional basis of humans and animals worldwide, provides the most calories and protein for the global food supply [2]. The crop yields prediction with accurate in advance plays an important role in the grain food security, famine prevention, and circulation market [3]. valuable for managing many field activities, one of them is fertilization [4]. India is one of the top wheat producers in the world, concerning total wheat acreage of the world's India contributes to 11.26% and it's world's total production of 17.98% [5].

Especially for winter wheat, India's production assumes a dominant role, with nearly85% of the total summer grain production [6]. The estimating of India's winter wheat yield is highly required because it, accurately and timely, considering its significant influence on agricultural development and national food security, or even global scale.

The development of biological materials in cereals grains depends on different practices in features due to many factors cultivar or genotype, soil fertility, growing conditions, and agronomic[4]. The classification and grading play an important and critical role in the market because assure quality control guidelines.

Also, grain quality standardization helps in searching for better marketing and processing of grain for the production of various products. The chemical composition, functionality, and optimum industrial end-use are the physical properties, they affect the selection of grain types for specific uses. The standardization process of grain quality is done by process a grain's lot with similar quality or grade of grains[5].

The classification of wheat grains is based on Different methods like growth habits, physical characteristics, and end-use. The use of these approaches can be referred to as technological classification [6]. Wheat-based products require different classes of grain for their processing. Industrial quality is characterized by physical, chemical, and rheological analysis related with roller milled refined flour, used especially in the manufacture of yeast-leavened

bread (hard wheat) or chemically leavened cookies and cakes (soft wheat) or to produce semolina to elaborate long and short pasta products (durum wheat).

First and significant quality control for physical examination of the wheat kernels is provided for Choice as raw materials because the physical characteristics of the kernel are related to the design of end-use, machinery, handling, aeration, and storage, etc. Increasingly, experiments are to better grasp their qualities, the intrinsic characteristics of the grain are measured.

The analysis of the features of the wheat kernel is important since new cultivars are constantly being bred and produced.[7]

2. Previous Work

Through machine learning algorithms and image processing, many Studies have been conducted out on the classification of wheat. These studies extracted two or more classes of features.

In this study [8], by combining machine learning and artificial neural networks, different categories of wheat were identified using tools of the Mat-Lab Technique. Various color and morphological characteristics of wheat were collected for the identification process. The testing was done under ANN using these characteristics. It was concluded that wheat could not be identified by color attributes or morphological features alone, so a mixture of both has been used. It was found that the average accuracy was 95.86 percent.

In another study [9], The rain-fed cultivars of wheat grain were classified using the ANN. Color features, morphological features, and textural features were identified for this reason. These features have been fed as input to the multilayer perceptron neural network. The accuracy of the description was concluded to be 86.58%. The accuracy was lifted to 87.22 percent by using the UTA algorithm for function extraction.

Another study by [10] For classifying wheat and barley grain kernels, Discriminant Analysis, and K Nearest Neighbor were used. The machine testing was done with only morphological characteristics, the only color characteristics, and a mixture of morphological characteristics, color characteristics, and textural characteristics. It was concluded that while morphological, color, and textural attribute forms are used together relative to using them alone, accuracies greater than 99 percent can be obtained.

In this study [11] Support Vector Machine and NeuralNetwork was used. The SVM accuracy is 86.8% and the Neural Network accuracy is 94.5%. It is assumed that this algorithm performed much better than that algorithm based on these results. This algorithm's precision has emerged to be above the algorithm.

3. Methodology of Proposed Work

The technique needs multiple measures that are as follows:

In the field of agriculture, machine vision is generally used for identifying the varieties of distinct food crops and for Identification of their performance, too. A machine vision system (MVS) offers an alternative to the physical process of biological products. Machine learning algorithms and digital images have been used for this process. With the use of a digital camera, these photographs are captured and then preserved for future work on the device.

The camera serves as an eye and the system acts as the brain in a Machine Vision System. Image analysis algorithms process digital images stored in the system. From the digital images, the features are extracted and use them to create the pattern. The machine learning algorithms process the input pattern on the basis of the objects

and categorized them into their respective groups. These computer algorithms used for object classification are referred to as classifiers of patterns.

This study deals with the wheat grain classification. Five machine learning algorithms, named as SVC (Support Vector Classifier), Logistic Regression, Decision Tree, KNN, and Random Forest, are used to do this classification.

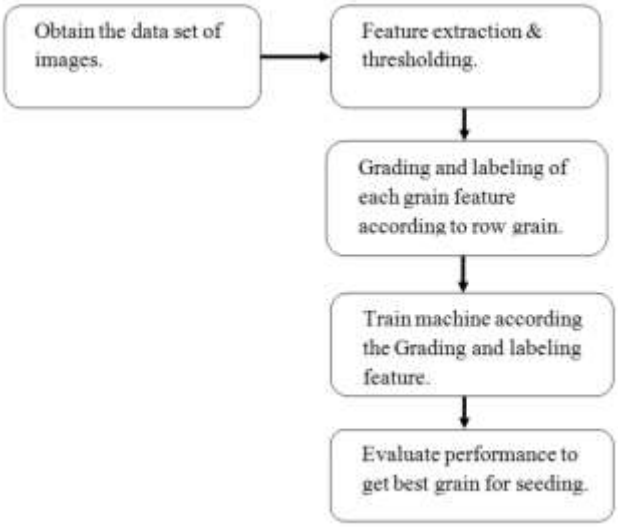


Figure 1: Methodology of wheat classification

3.1 Obtain wheat grain images dataset

Acquiring pictures of wheat grains is the first step in the classification of wheat. The wheat grains have been put on a black background to obtain the image of wheat grains and special measures have been taken so that no light gets reflected. These steps are taken in order to capture a better picture of the grains of wheat. the computer is used to store these images and a digital camera was used to obtain them.

S.No	Features of Grain	Criteria	Formula
1	Damage grain	The roundness ratio of Grain≈ 0.98	No. of damaged grain ----------*100 total quntative of grain in the sample
2	Foreign matter	Grain Pixel value < 5 and volume > 50	No. of particles with pixel value small than 5 ---------- *100 total no. of grain in the sample
3	Slightly damaged grains	Roundness of grain Ratio of ≈ 0.8	No. of slightly damaged grain ----------*100 total no. of grain in the sample
4	Broken grains	Length < ¾ of full length	No. of broken grain ---------- *100 total no. of grain in sample
5	Weevilled grains	Volume < 15	No. of grains with volume smaller than ---------- *100 total no. of grain in the sample

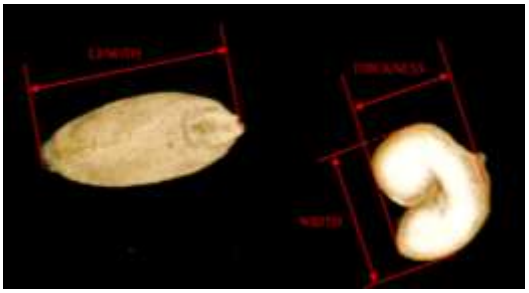


Figure 2: The captured image of Wheat grains

3.2. Feature extraction and Thresholding

The next phase in classifying wheat is to perform image thresholding the simplest method of image segmentation is called thresholding [7]. The holding can be used from a grayscale image to construct, binary pictures by turning both pixels below a few thresholds to zero and all pixels below zero to a few thresholds. The extraction of features describes a collection of features that reflect the information that is essential for wheat classification[5].

Table 1: Features extracted from images of wheat Features Formula

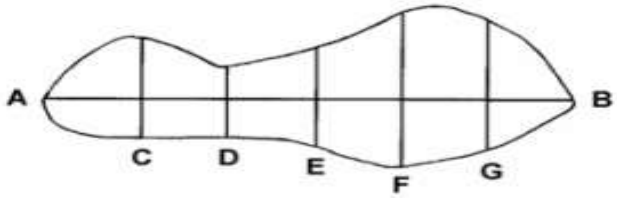
Features	**Formula**
Area	Using the offset method: Firstly, calculate the length of the longest axis of the area (line AB). This is called the line of length. Next, divide the long line into equal sections, for example, 10 ft. These lines are called ***offset lines***.  Length of line(AB) = 60 pixels, distance between offset lines are 10-pixels apartLength of each offset line; C = 15 pixels, D = 10 pixels, E=15 pixels, F = 25pixels, G =20 pixels Total length of offset lines = C + D + E + F +G = 15 + 10 + 15 + 25 + 20 = 85 pixels Area to measure = Distance between offset lines x sum of the length of the offset lines =10 pixel x 85 pixel= 850 pixel
Perimeter	No. of horizontal and vertical links+(No. of diagonal links*√2)
Volume	volume=1/ 2[x1(y2 - y3) + x2(y3- y1) + x3(y1-y2_](z1+ z2 + z3) / 3 Where, (x1, y1), (x2, y2) and (x3, y3) are the planar coordinates of the triangle vertices and z1, z2 and z3 are the respective elevations
Roundness Ratio	4*pi*area/perimeter^2 field
Percentage of broken grain	No. of broken grain/no. of the whole grain*100

Table 2: AGMART STANDARDS [10] for grading wheat grains

Grade Designation on	Foreign matter (% by wt.)	Damaged grains (% by wt.)	Slightly damaged grains (% by wt.)	Immature shriveled and Broken Grains	Weevilled Grains (% by wt.)
I	1	1	2	2	1
II	1	2	4	4	3
III	1	4	6	10	6
IV	1	5	10	10	10

3.4. Implementing Machine Algorithms [6]

Five machine learning algorithms, named as SVC (Support Vector Classifier), Logistic Regression, Decision Tree, NN, and Random Forest are use for classification. NN (Neural Network) [8] has been successfully used in numerous fields as a classifier. So, for wheat grain classification, it can be used. It can be modelled on the brain of a person. The neuron that operates identically in NN is the basic processing unit of the brain. The Support Vector Machine (SVM)[9] is a machine learning algorithm that learns to assign object labels by example.

It maximizes a specific mathematical function to concerning a given data set. The numeric parameters are calculated to allow access to the quality of each grain in order to implement the algorithms. Such parameters are separately fed to each algorithm and then the performance can be evaluated. In order to begin with classification algorithms, logistic regression is the correct algorithm.

Eventhough the name 'Regression' comes up, it is not a regression model, but a classification model. To frame binary output model, it uses a logistic function. The output of the logistic regression will be a probability ($0 \leq x \leq 1$), and can be used to predict the binary 0 or 1 as the output (if $x<0.5$, output= 0, else output=1). The decision tree is a tree-based algorithm which is used to solve problems with regression and classification. An inverted tree is framed which is branched off from a homogeneous probability distributed root node, to highly heterogeneous leaf nodes, for deriving the output. Regression trees are used for dependent variable with continuous values and classification trees are used for dependent variable with discrete values.

3.5. Evaluate of Performance

After the algorithms are implemented, their performance is measured to see whether or not desired outcomes are achieved. By this out come, farmer know about the seed which gives better result in the Bhilwara regoin. And, farmer get better productivity and low cultivation requirement because of best seed selection.

4. Results

After implementing the algorithms and the results of their evaluated performance were obtained. The results show that the accuracy of Logistic Regression is 98%, Support Vector Machine (SVM) comes out to be 96 %, Decision Tree accuracy is 94% , Neural Network (NN) comes out to be 92% and Random forest have accuracy 90% only.

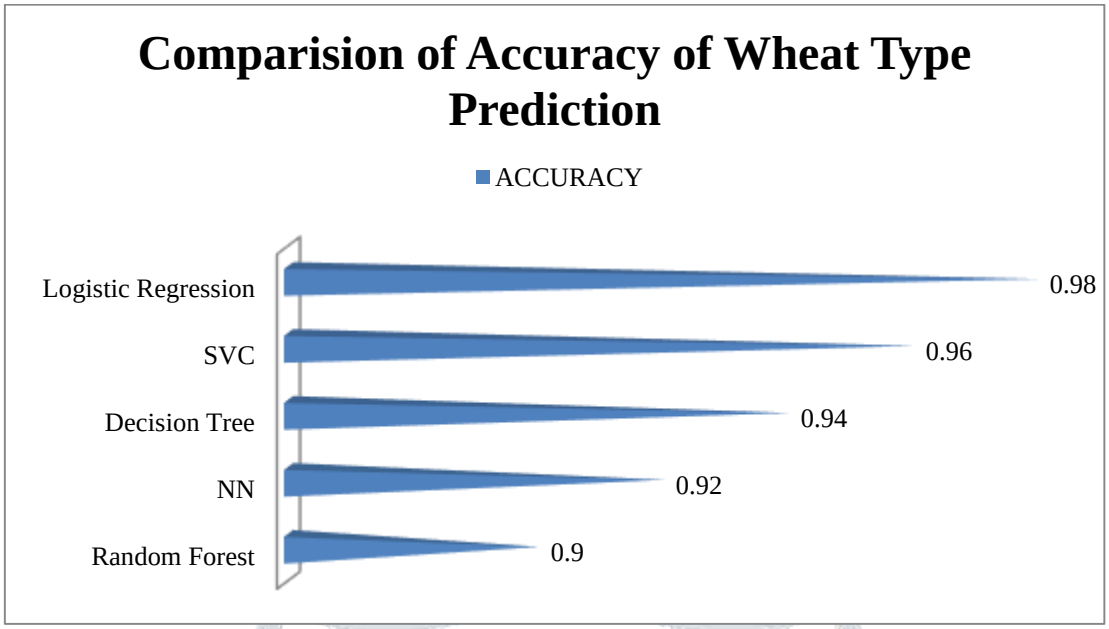


Figure 3: Overall accuracy of Wheat Grain

5. Conclusion

The main objective is to classify wheat grains efficiently using machine learning algorithms. For this purpose, five machine algorithms are used i.e the Logistic Regression is 98 percent performed better than all other algorithms. Based on these results it is concluded that Logistic Regression algorithm performed better than all other algorithm. The accuracy of this algorithm is best out to be more than all other algorithms.

6. Future Scope

This study can be carried out by using some other machine learning algorithms, which have higher accuracy and has a low computational cost. Moreover, research can be done on different varieties of wheat grains.

Reference

[1] FAOSTAT [Internet]. 2015. Available from: http://faostat3.fao.org/download/Q/QC/E [Accessed: 2016-08-02].

[2] Paulsen GM, Shroyer JP. Wheat/Agronomy. In: Wrigley C, Corke H, Walker CE, editors. Encyclopedia of Grain Science. Elsevier Academic Press. 2004. pp. 337–347.

[3] Morris CF. Cereal/Grain-Quality Attributes. In: Wrigley C, Corke H, Walker CE, editors. Encyclopedia of Grain Science. Elsevier Academic Press. 2004. pp. 238–254.

[4] Wilhelm LR, Suter AD, Brusewitz GH. Physical Properties of Food Materials. Chapter 2: Food & Process Engineering Technology. ASAE. 2005. pp. 23–49.

[5] Serna-Saldivar SO. Cereal Grains: Properties, Processing, and Nutritional Attributes. CRC Press, Taylor & Francis Group. 2010. 752 p.

[6] Carson GR, Edwards NM. Criteria of Wheat and Flour Quality. In: Khan K, Shewry PR, editors. Wheat: Chemistry and Technology. 4th edition. AACC International Inc. 2009. pp.

97–118.

[7] Babić L, Babić M, Turan J, Matić-Kekić S, Radojčin M, Mehandžić-Stanišić S, Pavkov I, Zoranović M. Physical and Stress-Strain Properties of Wheat (Triticumaestivum) Kernel. Journal of the Science of Food and Agriculture. 2011;91:1236–1243. DOI: 10.1002/jsfa.4305.

[8] Arefi, A. ModarresMotlagh, and R. FarrokhiTeimourlou, "Wheat class identification using a computer vision system and artificial neural networks",2011.

[9] AlirezaPazoki and ZohrehPazoki, "Classification system for rainfed wheat grain cultivars using artificial neural network", 2011.

[10] F. Guevara-Hernandez and J. Gomez-Gil, "A machine vision system for classification of wheat and barley grain kernels" 2011.

[11] P. Meesha and B. Nidhi, "Classification of Wheat Grains Using Machine Algorithms" International Journal of Science and Research (IJSR), India Online ISSN: 2319-7064 Volume 2 Issue 8, August 2013.