Design Engineering

ISSN: 0011-9342 | Year 2021
Issue: 7 | Pages: 7041-7048

Implications of Technology in Agriculture – Review on Smart Agriculture.

Prajwal R
Research Scholar, Jain-Deemed-to-be-University
gajja.prajwal@gmail.com

Dr. Stalin Alex R
Associate Professor, Computer Science & Engineering(Data Science), School Of Computer Science & IT, Jain -Deemed-To-Be-University, Bengaluru, Karnataka, India
drstalinalex@gmail.com

Abstract:

Smart farming is an epicentre of the new upcoming agricultural methodologies. It is the amalgamation of new technology-oriented procedures wrapped with sustainable practices that enhances productivity by not harming mother earth through overuse of pesticides, fertilizers, and other natural resources. Technologies such as Machine learning, IoT, Sensor-based agriculture, etc. are gearing up the pace in making their presence in modern-day farming practice. Our farmers can improve the production capabilities, reduce the wastage of yield, and maximize their productivity by adapting to the technology-oriented farming. The usage of these technologies will have a positive edge. Agriculture combined with the advantages of sensor based data collections can be used to monitor soil health, water level of soil, crop quality, weed detection, disease detection, etc. This review paper indicates the usage of the machine learning techniques in the field of agriculture in order to optimize the yield.

Keywords IoT, Smart Agriculture, Machine Learning, Precision Agriculture

I. INTRODUCTION

Machine Learning a broader approach of Artificial intelligence facilitates machines to learn and produce the output that we the users expect. It might be as simple as a keyboard learning that takes place in mobile phones to much complex learning such as stock predictions, early detection systems, etc. It is a methodology in which machines are trained to produce appropriate results based on the data sets that are available. The data sets given for training for the machines are known as training sets. Aforesaid technology is now being used in the area of agriculture for various purposes such as early detection systems, yield optimization techniques, intelligent irrigation systems, sensor-based farming, intelligent cold storage systems, etc. Soil management, Water management, Disease management, etc are some of the areas where we can find the implementation of Machine Learning in the field of agriculture.

The system or framework accepts data from the sensors that are placed in the farms. The data as such is not useful, but when the same data is subjected to analysis, it provides insights about the

ISSN: 0011-9342 | Year 2021
Issue: 7 | Pages: 7041-7048

farm and crop grown in the farm. The data collected from the sensors cannot be processed just by one algorithm or one process. We must find out a framework that works out for the situation or the problem that we are resolving. Amalgamation of multiple algorithms will lead us to a framework that can be designed which will be efficient, practical and might help farmers optimize yield by adapting the sustainable approaches.

The organization of this document is as follows. In Section 2 (**Frequently used algorithms**), in Section 4(**Conclusion**) a conclusion that is the last part of the article.

II. ALGORITHMS

1. **SVM:** Support Vector Machine a supervised machine learning model is based on the classification algorithms. By using a Hyperplane to classify the elements, this algorithm helps us in making machine learning easy. We can use multidimensional transformations if the complexity of the data increases and imposes challenges for introducing a hyperplane. In agriculture, SVM can be used for segregating between the crop and the weeds that would grow in the farm. It thereby enhances the fertile growth of the plant.

As P.J.Ramos in his paper titled "Automatic fruit count on Coffee branches using Computer Vision"[3] utilizes this methodology to segregate coffee seeds of the plant using digital images. He segregates the seeds into harvestable, semi ripe and non harvestable or unripe seeds[3]. The results showed improvement towards 87.83% accuracy for harvestable coffee seeds.

In [5], the author uses the above mentioned method for identification of the number of immature green citrus fruits which are grown under natural outdoor circumstances. Features such as coarseness, contrast, directionality, regularity, roughness, granularity, etc of an image of 20*20 pixels were observed. The accuracy was obtained to be 80.4%.

S. Ying-xue, X. Huan et. al, in [6]observe various features like, agricultural surface weather, and soil physico-chemical data with yield and development records and predict the yield at different development stages.

2. **Boyer Moore:** A pattern recognition algorithm which recognizes patterns in strings has been serving as a standard benchmark in this category. This algorithm applies a 'backward' approach. The pattern string is compared with the string from right to left from the rightmost character.

This algorithmic approach was implemented in detection of cherry branches with full foliage by Suraj Amatya et. al.[4] The accuracy of this method was observed at 89.6%.

ISSN: 0011-9342 | Year 2021
Issue: 7 | Pages: 7041-7048

3. **DBSCAN:** Density-Based Spatial Clustering of Application with Noise, is one which is usually used for clustering in machine learning or in data mining. This algorithm accepts 2 basic parameters; eps - which specifies the distance between each point to be considered as neighbors; minPoints - this specifies the minimum number of points that has to be considered to name it a dense region. In early detection systems, this algorithm can be used to classify the parameters that will classify the given qualities[1]. The algorithm can be effectively used for separating the high density cluster with a low density cluster [2]. For example, in detection for finding the disease or the pest infested plant at an early stage. It is easy for us to identify a good versus the infested crop, as we are interested in finding a healthy crop. This algorithm can be backed by the data collected by the sensors or UAV spatial images against the images that showcase good and bad plants. This will enhance productivity and also in adopting sustainable practices. In [10], authors have expressed that data mining techniques are very much necessary for accomplishing the solutions for the broadest economic sector that is agriculture. Jharna Majumdar et. al in [10], have worked on the Indian context. They have taken seven features such as rainfall, area of cultivation, crop being produced, maximum and minimum temperatures, speed of wind, humidity, and cloud coverage into consideration for their work. In their proposed work, they have used a modified approach of DBSCAN. PAM and CLARA have been coupled together in their approach that directs them to obtain optimal parameters to produce the maximum crop production. The crop under consideration in [10] is wheat.

4. **Divisive Analysis (DIANA):** It is a hierarchical clustering technique that can be used in clustering. This algorithm uses a top-down approach for clustering. Initially, the entire data is considered as a single cluster and further divided into 2 least similar clusters. This process is repeated till all the elements are left as singles. For example, if a spatial image of a farm is taken using a UAV, classification of the plants and the weed that are grown in the farm can be made using the said classifier algorithm.

5. **K-Means:** This algorithm is a clustering algorithm. It accepts n data points as inputs and classifies them into k clusters. This process of grouping is the training phase and the output of the same would be clusters that new points would belong to, based on the training given to the model. The overall idea behind this algorithm is, we add k new points to the given data. Each of the k new points is termed as “centroid”, and will be going around trying to center itself in the middle of the k clusters. Once these points stop moving, our clustering algorithm reaches its end. For instance, if we were to grade the plants on the level of effect of disease on it, we can apply this algorithm and classify it to the nearest category.

The Clustering/EM method tries to assign the observations made to those clusters so that the means across clusters available are as different from each other. In other words, EM assigns observations by computing probabilities of cluster memberships based on one or more

ISSN: 0011-9342 | Year 2021
Issue: 7 | Pages: 7041-7048

probabilistic distributions, unlike K Means that tries to maximize the difference for continuous variables.[7] An algorithm similar to K Means clustering, EM(Expectation Maximization) technique was incorporated by J. Senthilnath et.al, in detecting tomatoes. The authors in the articles have worked on the High spatial imagery that was captured using UAVs. The precision was noted at 0.9191 and F-Measure was noted at 0.7308 by the authors.

6. **ANN:** Artificial Neural Networks. These networks are designed to mock the working of the human brain and neurons. They are mainly used in the field of economics, forensics and for pattern recognition. It basically contains 3 layers.

a. Input Layer: This layer accepts the input that represents raw data which would be fed into the network.
b. Hidden Layer: The activities of the input units and the weights on the connections between the input and the hidden units. There may be one or more hidden layers.
c. Output Layer: The output is based on the activity of the internal hidden layers and the weights between hidden layers and the output units.

Iftikhar Ali et.al, implemented ANN for estimation of biomass that could be produced in grasslands. Vegetation indices, spectral bands of the same were collected and fed into the neural network to obtain RMSE value at 15.35.[8]

On the similar lines, X.E. Pantazi et.al, works on wheat and yield prediction within field variations were analyzed by considering the features like normalized values of On-line predicted soil. Their work exposes 81.65% accuracy.[9]

On the similar lines, ANN can be further implemented for using the results that are obtained from the other classifiers to determine the health condition of a crop. The overall growth quality can be monitored using this method. Thereby having a check on the quality of the yield.

7. **Image Processing:** In [13], the author quotes some of the works from Cruz et. al.. The author emphasizes on Integrated Pest Management (IPM). He opines that most of the recommended solutions for the Fall Armyworm were through the methods of applying the pesticides and chemicals which was proportional to the damage that was caused by the worm. The author uses an image of 30 plants that were cultivated in pots in a greenhouse in Brazil. Random plants were chosen which were infested with the worm. The author uses 720 images totally using an RGB configuration through a digital camera.

ISSN: 0011-9342 | Year 2021
Issue: 7 | Pages: 7041-7048

III. DATA ANALYSIS AND INTERPRETATION

TABLE I

Method Name	Usage in Field of Agriculture	Paper Referred
SVM	Coffee seeds were classified based on this algorithm.[3] The author used "Computer Vision" to capture digital images of the plant. These images were fed to the SVM algorithm to segregate them on the basis of harvestable, semi-ripe and non-harvestable seeds. The method adopted in the paper by the author was able to achieve 87.83% accuracy in segregation of harvestable seeds. Also in [5], the author has used the same algorithm in identification of immature green citrus. An image of 20*20 pixels were collected and features such as roughness, coarseness, contrast, granularity, regularity were observed. The author claims to have achieved 80.4% accuracy. S. Ying Xue et. al, in [6] have used SVM in understanding the condition of the agricultural field, surface weather, soil condition to predict the yield at different stages.	[3], [5], [6]
Boyer Moore	It is a pattern recognition algorithm that has been the benchmark for the string pattern recognition problem. This algorithm applies a backward approach. Suraj Amatya et. al. in their paper titled "Detection of cherry tree branches with full foliage in planar architecture for automated sweet-cherry harvesting" has used the algorithm in detection of cherry branches with full foliage. The author was able to achieve 89.6% by using this methodology.	[4]
DBSCAN	Density Based Spatial Clustering of Application with Noise. This algorithm comes handy when we need to clusterize the data in machine learning or in data mining. This algorithm uses two parameters - One is the distance between each point that is considered as neighbours, and the second parameter being the minimum number of points that are required to call a region a dense region. This algorithm can be implemented in systems and applications that	[1], [2], [10]

ISSN: 0011-9342 | Year 2021
Issue: 7 | Pages: 7041-7048

	perform an early detection. It can be used to identify a plant that is infested by a pest at an early stage by using this algorithm. In [10], the authors have taken 6 years of data for the areas of Karnataka, along with 7 parameters such as rainfall, area of cultivation, crop being produced, maximum and minimum temperatures, speed of wind, humidity, and cloud coverage which would help them to determine the optimality for maximizing the yield. A modified DBSCAN coupled with PAM and CLARA was used to obtain the optimal parameters to maximize yield.	
Divisive Analysis (DIANA)	The algorithm classifies clusters in a top-down approach. Initially the entire data set is considered as a single cluster. Which is further divided into 2 clusters based on the similarities. This process is repeated till data elements are left as single data items. In [11], Alan Gavioli et.al., have investigated this algorithm along with 20 other clustering algorithms. The evaluation was performed on data collected for a span of 5 years from the period of 2010 to 2015. The main objective behind their work was to identify and manage sub areas within an agricultural field processing various characteristics.	[11]
K-Means	This algorithm accepts n data points and classifies them into k clusters. J.Arno et. al., in [12], the authors have carried cluster analysis using the values interpolated from 3 grape fields. Initially, the clustering was made through k-means and later was compared by repeating the same using fuzzy logic and c-means algorithms. The experiment showed larger variations. In k-means algorithm, the data or observation can be grouped into a single cluster, whereas in fuzzy or c-means, the same could be grouped into more than one cluster. A similar method to K-Means, that is EM clustering assigns the observations to the clusters based on the probabilistic distribution, unlike K-Means that tries to identify and maximize the difference among the continuous variables. In [7], EM (Expectation Maximization) method was implemented by authors on detecting tomatoes. The work was on High spatial imagery that was captured using UAVs. The work shows 0.9191 for precision and 0.7308 for F-	[12], [7]

ISSN: 0011-9342 | Year 2021
Issue: 7 | Pages: 7041-7048

	Measure value.	
ANN	These are designed to mock the functioning of the human brain. The functionality of this is divided into 3 layers - input layer, hidden layer and output layer. In [8], Iftikhar Ali et. al. use ANN methodology to estimate the amount of biomass that could be produced from a grassland. Similarly X. E. Pantazi et. al. presents their work on yield prediction of wheat using ANN [9]. Their work exposes an accuracy of 81.65%	[8],[9]
Iterative Approach	720 images of Maize plants were captured. The saplings were infested using fall armyworm and the images were captured in RGB. The images were first transformed into 256 grayscale images. Then the images were thresholded using an iterative method [13]. The author was able to observe 94.72% of accuracy in classifying the images as infested crops.	[13]

III. CONCLUSION

With the aid of technology, farmers can precisely monitor aspects including growth, the quality of the farm, the soil content and its quality. This helps in maintaining the health of the crop, the soil at a top notch level. Advancement of technology such as IoT, Machine Learning and Deep Learning, has been a boon to farmers. As seen in the tabulation, many learning algorithms are used in smart agriculture that have helped a lot in the way we can monitor and extract the yield from the farms. These technologies have helped in identifying insects, weeds grown in the farm. This will have a direct and proportional impact on the yield expected out of a farm.

REFERENCES

1. https://towardsdatascience.com/how-dbscan-works-and-why-should-i-use-it-443b4a191c80, Internet Source
2. https://medium.com/@elutins/dbscan-what-is-it-when-to-use-it-how-to-use-it-8bd506293818#:~:text=DBSCAN%20works%20as%20such%3A,this%20shape%20as%20a%20cluster., Internet Source
3. P.J. Ramos, F.A. Prieto, E.C. Montoya, C.E. Oliveros, “Automatic fruit count on coffee branches using computer vision” Comput. Electron. Agric. 2017
4. Suraj Amatya, Manoj Karkee, Aleana Gongal, Qin Zhang, Matthew D. Whiting “Detection of cherry tree branches with full foliage in planar architecture for automated sweet-cherry harvesting”. Biosyst. Eng. 2015

ISSN: 0011-9342 | Year 2021
Issue: 7 | Pages: 7041-7048

5. Subhajit Sengupta, Won Suk Lee, “Identification and determination of the number of immature green citrus fruit in a canopy under different ambient light conditions”. Biosyst. Eng. 2014
6. S. Ying-xue, X. Huan, Y. Li-jiao, Support Vector Machine-Based Open Crop Model (SBOCM): Case of Rice Production in China, Saudi Journal of Biological Sciences (2017)
7. https://docs.rapidminer.com/latest/studio/operators/modeling/segmentation/expectation_maximization_clustering.html#:~:text=The%20EM%20(expectation%20maximization)%20technique,to%20the%20K%2DMeans%20technique.&text=Instead%20of%20assigning%20examples%20to,one%20or%20more%20probability%20distributions., Internet Source
8. Iftikhar Ali, Fiona Cawkwell, Edward Dwyer, and Stuart Green, “Modeling Managed Grassland Biomass Estimation by Using Multitemporal Remote Sensing Data—A Machine Learning Approach”. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 2016
9. X.E. Pantazi, D. Moshou, T. Alexandridis, R.L. Whetton, A.M. Mouazen, “Wheat yield prediction using machine learning and advanced sensing techniques”. Comput. Electron. Agric. 2016
10. Jharna Majumdar, Sneha Naraseeyappa, Shilpa Ankalaki, “Analysis of Agriculture data using data mining techniques: application of big data”, Journal of Big Data, 2017
11. Alan Gavioli, Eduardo Godoy de Souza, Claudio Leones Bazzi, Kelyn Schenatto, Nelson Miguel Betzek, “Identification of management zones in precision agriculture: An evaluation of alternative cluster analysis methods”, Elsevier, Biosystems Engineering, 2019, pp. 87
12. J. Arno, J. A. Martinez-Casasnovas, M. Ribes-Dasi and J. R. Rosell, “Clustering of grape yield maps to delineate site-specific management zones, Spanish Journal of Agricultural Research, 2011, pp. 722-723
13. D. G. Sena Jr, F. A. C. Pinto, D. M. Queiroz, P. A. Viana, “Fall Armyworm Damaged Plant Identification using Digital Images”, Biosystems Engineering, pp. 449