AN INVESTIGATION OF VARIOUS MACHINE LEARNING APPLICATIONS TO IMPROVE FOOD AGRICULTURE SECTORS

[1*]Dr. Deshmukh Narendra P, [2]Prof. Sphurti Birajdar, [3]Prof. Deshmukh Megha N

[1]Assistant Professor, Dattakala Group of Institutions, Dattakala Faculty of Management, A/P Swami Chincholi, Tal Daund, Dist Pune
npdeshmukh76@gmail.com

[2]Assistant Professor, Dr. D.Y.Patil Vidyapeeth's Global Business School & Research Centre, Tathawade, Pune.
sphurti.birajdar@dpu.edu.in

[3]Assistant Professor, A. G. Patil Institute of Technology, Vijapur Rd, Opp. SRP Camp, Soregaon, Solapur, Maharashtra, India.
megha.deshmukh0@gmail.com
Corresponding Author: Dr. Deshmukh Narendra P (npdeshmukh76@gmail.com)

ABSTRACT

In the field of machine learning, there is a wide number of approaches that may be used to develop a model that is capable of anticipating data that has not yet been seen and to learn prediction rules based on historical data. Machine learning involves analysing data samples in order to recognise patterns and determine how to make decisions. This is done with the goal of predicting future data. The current model of agriculture, known as "smart agriculture," views a farm as a collection of distinct units and searches for inconsistencies in the levels of output and demand at each individual unit. The ultimate goal of "smart farming" is to reduce overhead costs while simultaneously increasing crop yields and business revenue. Farmers who are on the cutting edge of agricultural practises employ them. The predictive power of machine learning algorithms paves the way for the implementation of intelligent agricultural practises.

Keywords: Machine Learning, Classification, Smart Agriculture, Prediction

1. INTRODUCTION

In general, the term "intelligence" refers to the mechanical reproduction of the human cognitive skills of learning natural language, planning in that language, perceiving in that language, and comprehending in that language. These cognitive abilities are referred to as "intelligence capacities." As a result of advancements in theory and research, computer systems are now capable of doing activities that previously required human intelligence, including visual perception, voice recognition, decision making, and language translation. The information technology sector has made significant investments in gear that is intended to simulate human performance. John McCarthy provided the definition of artificial intelligence as the use of scientific research and technological development in the production of intelligent computer programmes. [1]

Machine learning and deep learning are two of the most common applications of artificial intelligence today. When it comes to forecasting, a broad variety of businesses and individuals make use of these models that are driven by data. Researchers are working on various machine learning approaches in an effort to solve this issue, which is complicated by the unpredictability and complexity of the data in the food industry [2]. The implementation of AI in the food sectors has become necessary as a result of a number of factors, the most important of which is the development of standardised and reliable product quality control procedures. These factors include improved customer experience, efficient management of the supply chain, increased operational efficiency, and reduced workforce size. [3]

The Passing Along of Information and Messages The study of gathering, preserving, manipulating, analysing, and distributing certain types of information for specified objectives is what we mean when we talk about technology in today's world. The advancement of computer and communication technologies are the primary pillars around which modern technology is built. These developments have had a disproportionately huge impact on the ways in which we store and communicate data. ICT, or information and communication technology, is essentially what enables electronic communication to take place by processing and transporting data. It is possible to see information and communications technology (ICT) as both a driver and an enabler due to its ability to bridge the gap between persons and geographical locations [3].

Recent advancements in information technology, computer technology, telecommunications, and the internet have created an environment that is amenable to the implementation of innovative agricultural extension tactics for the growth of agriculture and rural regions. Since information and communication technology has become so widespread in today's society, literacy is often associated with the ability to use computers. The singular and the plural forms of the abbreviation "ICT," which stands for "information and communication technologies," may be

used interchangeably. ICT is defined as "the integration of microelectronics, computer hardware and software, telecommunications, and the storage of huge amounts of information, as well as its quick transfer through computer networks" by the United Nations Development Program (UNDP). As a consequence of the widespread availability of ICTs such as the Internet, several communication and collaboration tools have been developed. Some examples of these tools include e-mails, online discussion forums, net meetings, webinars, and video conferencing. It has been crucial in both enabling and facilitating the administration of organisational knowledge [3], which has been a critical function for it. [4]. Information and communication technology (ICT) has been cited as a primary contributor to the achievements of a number of businesses, including IBM, HP, TI, and Buckman Labs [5] [6]. These businesses believe that the assistance they obtained from ICT in terms of knowledge management was a significant factor in the achievement of their goals.

2. Related Work

When dealing with classification problems, statisticians often turn to a technique known as logistic regression (LG). There is a substantial amount of overlap between linear regression (LR) and logistic regression; however, although LR is utilised for numerical concerns, logistic regression is used for categorization problems. The output calculation coefficient value is the ultimate goal of logistic regression, which aims to get it. The output of a nonlinear function is converted into a 0 or 1 by employing the logistic function as the transformative tool. To draw conclusions from a collection of data, one may utilise logistic regression (LG), which involves first analysing the data and then basing one's predictions on one or more independent variables. It is a powerful and effective algorithm for the learning process. [7]

A classifier known as Nash Bayes (NB) applies the Bayes algorithm to the data. Assume that any correlations found between the characteristics in the dataset are the result of pure coincidence. Finding the parameters that are required to generate a forecast only requires a very little amount of the available data. It is efficient in the categorization jobs it does. NB creates a forecast by using the probabilities that are derived for each feature in isolation and then combining them. Due to the simplicity of its application, this model is well suited for modelling large datasets. [7] As its name suggests, the Decision Tree (DT) adopts a structure similar to a tree in order to identify answers to problems involving categorization and regression. The enormous dataset is reduced by DT to a manageable size, consisting of a tiny group of relatively few values that can be readily segmented over time. There are many different kinds of nodes in DTs, the most important of which are the root node, the node at which choices are made, and the leaf nodes. In contrast to leaf nodes, which are where choices are made, root nodes may branch out in several directions. This method of data analysis may be used on any sort of information, and it is not difficult to learn how to use it properly. Random Forest (RF) [7] is a machine learning technique

that has some characteristics with DT. However, in contrast to DT, RF uses a forest of trees to avoid overfitting to the training data. When it comes to activities associated with machine learning, the RF may be used for regression, classification, or pretty much anything else you can think of. The RF often performs better than the DT. [8]

Because of the great degree of accuracy it offers, the K-Nearest Neighbor (KNN) algorithm was selected for use in this study's classification problems. It is an algorithm for machine learning with supervision. It begins by examining a K-group of tagged data as a starting point. It determines the class that has the greatest number of nearest neighbours to a new point and assigns it to that class on the basis of the distance from the K-labeled data. This process begins with the new point. During training, the number of neighbours that are taken into consideration is denoted by the letter K. [8]

Support Vector Machine (SVM) is a technique for label categorization that places things into the categories that are most appropriate for them. This approach makes use of a hyperplane. It has a broad range of applications, including the identification of fraudulent activity in locations like databases that track credit card transactions. In addition, SVM performs effectively even when dealing with data of a large dimension. In machine learning, overfitting is a prevalent issue; nevertheless, support vector machine (SVM) is one of the few techniques that can deal with it. As a result, this study investigates the effectiveness of support vector machines (SVMs) in detecting fraudulent activity in datasets including information on credit card transactions. The findings indicate that SVMs do well in comparison to other machine learning methodologies.

At the moment, neural networks (NN) are the machine learning algorithms that are proving to be the most successful, and the vast majority of data mining tools make use of them. It functions using a system that is analogous to that of the human brain. A network consisting of several neurons that are linked to one another processes the input. It performs functions that are analogous to those of the human brain. It makes no difference how much data you feed the neural network; it will train quickly and accurately regardless. In addition, NN ensures the development of patterns and trends for the purpose of analysis, which is something a single person would be unable to perform with a dataset that was sufficiently complicated.

The use of NN in data mining applications such as decision assistance is becoming more widespread. Computers are pretty adept at following instructions, and just like people, they can learn from their mistakes. Computers are also quite good at following instructions. A neural network (NN) is a model that can be created inside of a computer that simulates the activity of neurons found in the human brain by integrating the most beneficial aspects of both. The effectiveness with which NN can categorise data and the ease with which it can recognise complicated patterns are two of its selling features. When compared to other machine learning

models, neural network models, also known as artificial neural networks (ANN), perform very well in terms of accuracy and efficiency. [8]

3. Applications of Machine learning in Agriculture

Artificial neural networks performed well in tests in which they were required to differentiate between cultivated crops and weeds in photos taken in fields [9]. First, an artificial neural network based on back propagation was taught to identify corn plants and weeds in digital photos by analysing the intensity patterns of the objects in the images. After it had been properly educated, the neural network was put to use to differentiate between the weed plant and the maize plant. In this particular investigation, a score of one hundred was attained.

For the purpose of this investigation, we developed an artificial neural network in order to make projections regarding the amount of soybeans and maize that will be harvested in the state of Maryland. The network implemented a back propagation method and gave us the ability to modify both the learning rate and the number of neurons present in the hidden layer. The study indicated that the ANN model produced more trustworthy forecasts of both soybean and corn yields when compared with the results of the regression model. This was determined by comparing the outcomes of the two models.

In the process of evaluating product quality, one of the most important steps is product grading. In addition to ranking the work, it determines the amount of effort, both in terms of time and money, that was put into obtaining the outcomes. Grading requires both specialised knowledge and a meticulous examination of the products that are being supplied. After taking into account each of the factors that go into the grading, an overall quality is determined. Things of a high quality go rapidly off the shelves but are also often returned. Quality assurance is the last phase, and perhaps the most important one, that each industrial or agricultural production unit must complete in order to fulfil the requirements of their respective industry or market.

There was an attempt [11] made to classify apples into one of four distinct grade levels using a fruit-grading system that was based on the internet. According to the findings of the research, almost ninety percent of the apples were accurately classified on the basis of their external look. When it comes to horticulture products, the fruit business represents a significant financial outlay. Additional study on the subject of rating quality may be found at this location. According to the findings of the study, machine vision has been used effectively in a variety of fruit grading systems.

In accordance with the standards established in Europe, Author [12] conducted study on the classification of apple fruits and selected one of four groups ('I,' 'II,' 'Extra,' and 'Reject'). Apples were classified according to their appearance, taking into account characteristics such as blemishes, colour, and size. According to the findings of the study, the success rates for correctly recognising Golden Delicious apples and Jonagold apples were, respectively, up to 78 and 72

percent. In another piece of research, the quality of the apples was assessed using a fuzzy logic method.

The importance of the agriculture industry to the overall economy makes broad use of cloud computing an absolute need. There is a good chance that agriculture contributes a significant amount of value to the economy. The suggested agricultural conceptual model that is based on cloud computing has the overarching objective of optimising benefits while taking into consideration all essential elements, such as the area geography and weather patterns. The results of ML's influence on farming are shown in Table 1.

Applications of Machine learning in Agriculture

1. **Crop Related Information**

2. **Soil Information**

3. **Monitoring Growth**

4. **Farmers Data**

5. **Expert Consultation**

6. **E-Commerce**
 7. **Object Detection and Classification**

8. **Sorting and Grading**

9. **Predictive Maintenance and Monitoring**

10. **Machine Learning Based Analytics**

Table 1: **Applications of Machine Learning in Agriculture**

If farmers make use of the knowledge included in a comprehensive database of all prior crops cultivated, they may get some value from having access to such a database. [10] The ability to store data in the cloud enables the preservation of both long-term predictions and localised

meteorological information. They provide the same function for farmers, assisting them in the decision-making process about their crops.

The knowledge that farmers have about their land has a significant role in the selection of crops that they will plant. It is also possible that it will offer a pattern of the soil from the past, which may be used to predict future trends in the soil. Do you notice any differences in the composition or kind of the soil that you've examined? Have there been any changes in the acidity or alkalinity of the soil?

The evolution of different types of crops in different environments and across time This enables a comparison to be made between the rates of growth experienced by the present generation and those experienced by previous generations. At the regional level, it is feasible to collect, keep track of, and analyse data pertaining to the farmers in a certain region. These data might be used by the planning authorities in order to narrow down the locations that have the highest potential agricultural effect. [12]

In this section, you can find explanations of the most typical difficulties that farmers experience on. In addition, specialists may be able to provide prompt support for certain issues. They are unable to compete in the market because of the geographical isolation in which they find themselves. There are often several individuals operating as middlemen between the customer and the producer, taking advantage of both sides in the process. Farmers have the ability to cut out the middlemen and sell their goods directly to customers by using the agricultural management information system that is accessible on the cloud. It is possible that having access to a web-based AMIS that provides real-time data on elements such as precipitation, market pricing, fertiliser recommendations, and crop planting periods might be extremely helpful for farmers located in rural or outlying regions. [13] [14]

4. CONCLUSION

The practise of analysing data in order to make inferences about the future is referred to as "machine learning," and the word is used to cover a broad range of different methods. The ultimate objective of "smart agriculture" is to boost earnings in the agricultural industry while simultaneously lowering the sector's operating expenses. New, forward-thinking farming practises are often adopted by wealthy farmers. Because machine learning algorithms are inherently predictive, they make it possible to practise "smart farming." This article examined some of the applications of machine learning that have been made in the agriculture sector. Because machine learning algorithms are able to generate predictions, they are the best option for solving a diverse range of practical problems in a number of industries.

REFERENCES

[1] Y. Kumar, K. Kaur and G. Singh, "Machine Learning Aspects and its Applications Towards Different Research Areas," 2020 International Conference on Computation, Automation and Knowledge Management (ICCAKM), 2020, pp. 150-156, doi: 10.1109/ICCAKM46823.2020.9051502.

[2] A. Negi and K. Rajesh, "A Review of AI and ML Applications for Computing Systems," 2019 9th International Conference on Emerging Trends in Engineering and Technology - Signal and Information Processing (ICETET-SIP-19), 2019, pp. 1-6, doi: 10.1109/ICETET-SIP-1946815.2019.9092299.

[3] R. Sil, A. Roy, B. Bhushan and A. K. Mazumdar, "Artificial Intelligence and Machine Learning based Legal Application: The State-of-the-Art and Future Research Trends," 2019 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS), 2019, pp. 57-62, doi: 10.1109/ICCCIS48478.2019.8974479.

[4] Manne, Ravi, and Sneha C. Kantheti. 2021. "Application of Artificial Intelligence in Healthcare: Chances and Challenges". Current Journal of Applied Science and Technology 40 (6), 78-89. https://doi.org/10.9734/cjast/2021/v40i631320

[5] I. Khan, X. Zhang, M. Rehman and R. Ali, "A Literature Survey and Empirical Study of Meta-Learning for Classifier Selection," in IEEE Access, vol. 8, pp. 10262-10281, 2020, doi: 10.1109/ACCESS.2020.2964726.

[6] P. Samuel S., K. Malarvizhi, S. Karthik and M. Gowri S.G., "Machine Learning and Internet of Things based Smart Agriculture," 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS), 2020, pp. 1101-1106, doi: 10.1109/ICACCS48705.2020.9074472.

[7] B. Sharma, J. K. P. S. Yadav and S. Yadav, "Predict Crop Production in India Using Machine Learning Technique: A Survey," 2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), 2020, pp. 993-997, doi: 10.1109/ICRITO48877.2020.9197953.

[8] A. Muniasamy, "Machine Learning for Smart Farming: A Focus on Desert Agriculture," 2020 International Conference on Computing and Information Technology (ICCIT-1441), 2020, pp. 1-5, doi: 10.1109/ICCIT-144147971.2020.9213759.

[9]K. N. Bhanu, H. J. Jasmine and H. S. Mahadevaswamy, "Machine learning Implementation in IoT based Intelligent System for Agriculture," 2020 International Conference for Emerging Technology (INCET), 2020, pp. 1-5, doi: 10.1109/INCET49848.2020.9153978.

[10] S. Namani and B. Gonen, "Smart Agriculture Based on IoT and Cloud Computing," 2020 3rd International Conference on Information and Computer Technologies (ICICT), 2020, pp. 553-556, doi: 10.1109/ICICT50521.2020.00094.

[11] F. Nolack Fote, S. Mahmoudi, A. Roukh and S. Ahmed Mahmoudi, "Big Data Storage and Analysis for Smart Farming," 2020 5th International Conference on Cloud Computing and Artificial Intelligence: Technologies and Applications (CloudTech), 2020, pp. 1-8, doi: 10.1109/CloudTech49835.2020.9365869.

[12] D. V. Ashoka, C. B. Abhilash and M. B. Bharath, "A Framework for Enhancing the Agriculture yield using Cloud Clusters," 2019 1st International Conference on Advanced Technologies in Intelligent Control, Environment, Computing & Communication Engineering (ICATIECE), 2019, pp. 162-166, doi: 10.1109/ICATIECE45860.2019.9063770.

[13] M. S. Mekala and P. Viswanathan, "A Survey: Smart agriculture IoT with cloud computing," 2017 International conference on Microelectronic Devices, Circuits and Systems (ICMDCS), 2017, pp. 1-7, doi: 10.1109/ICMDCS.2017.8211551.

[14] D. K. Baghel, A. Singh and P. K. Deka, "Agricultural management using cloud computing in India," 2017 International Conference on Computing, Communication and Automation (ICCCA), 2017, pp. 801-806, doi: 10.1109/CCAA.2017.8229905.