Transactions on Computer Science and Intelligent Systems Research
ISSN: 2960-1800, eISSN: 2960-2238 | Volume 5
AIDML 2024

Research and Analysis of the Application of Machine Learning in Agricultural Development

Yimin Yuan *

Assumption Catholic Secondary School, Burlington, L7N 3P1, Canada

* Corresponding author: yimin.yuan.580@gobluego.ca

Abstract. Agriculture is the most basic, fundamental and important industry. Now, amid global climate change and resource shortages, agriculture must deal with the challenges of growing demand as the world's population increases This article organizes three aspects of agriculture that need improvement: anticipatory preparation before production, improvement of production methods, and detection and classification of agricultural products, and analyzes how machine learning can help agricultural progress in these three aspects. Residual deep convolution and spatial pyramid pooling algorithms in machine learning can be used to help detect plant pests and diseases. The RF algorithm, XGBoost algorithm, LightGBM algorithm and CatBoos in machine learning can generate landslide susceptibility maps. Deep learning, convolutional neural networks, and support vector machines can identify hybrid wheat. Through this research, it can be determined that machine learning can be of great help to agricultural development, and this help and development is mutual. The significance of this study lies in how machine learning can help agricultural development and face these problems.

Keywords: Machine learning; Agricultural development; application.

1. Introduction

Agriculture is cultivating the soil, growing crops, or raising livestock for human use, including producing food, feed, fiber, fuel, or other valuable products. Agriculture is really important to any nation or society. It supports people's livelihoods, provides raw materials for food and other products, helps countries develop their economies, creates jobs and increases incomes. However, agriculture faces challenges such as increased demand and lack of resources and energy. In order to face and solve these challenges, agriculture needs adequate anticipatory preparation and improvement of production methods and inspection and classification of agricultural products.

Machine learning is a very popular and important part of artificial intelligence. Machine learning algorithms use data to identify patterns and trends to make future predictions and analyses. Machine learning algorithms can improve model performance by practicing on large amounts of data. Machine learning can also automatically learn and optimize models, which not only reduces the cost and time of manual intervention, but also improves prediction efficiency and accuracy. For example, Mesut Ersin Sonmez, Kadir Sabanci, and Nevzat Aydin used convolutional neural networks-support vector machines in machine learning to identify wheat hybrids, and Sourav Bhadra, Vasit Sagan, Juan Skobalski, and Fernando Grignola used multi-temporal drone-based RGB Image and end-to-end 3D CNN for soybean yield prediction, Awais, M., Naqvi, S.M.Z.A., and Zhang, H. Soil analysis through machine learning.

This paper aims to prove that machine learning can be applied to agriculture and promote the development of agriculture. In order to achieve this goal, this article briefly introduces machine learning and agriculture, analyzes the three main aspects that need to be improved in agriculture, and summarizes the research on machine learning algorithm models applied by other scholars to address these three aspects. For example, machine learning in Applications in agricultural product yield forecasting. This article analyzes these studies and their results, and compares the differences and advantages and disadvantages between machine learning algorithm models and traditional methods. This article also puts forward some challenges and problems in the current application of machine

learning in the agricultural field, and discusses the potential and development direction of machine learning in agricultural development in the future. Through these steps, this article can conclude that machine learning can greatly help agricultural development.

2. Machine Learning Overview

Machine learning is a field of study in artificial intelligence and computer science. Machine learning enables computers to learn from parse data and algorithms, improve their accuracy autonomously, and make decisions or predictions without direct code. It is a science of managing the learning of computers to create computer systems that can automatically improve based on experience, just like humans do. By machine learning, a computer can improve its learning ability and intelligence, so the computers can parse data it never met in the training before.

In 1943, the computational model theory of neural networks was proposed, laying the foundation for the development of machine learning. In 1950, the proposal of the Turing test made artificial intelligence an important research topic. Subsequently, there have been countless breakthroughs in this field, and people have promoted its development. Deep learning has also achieved impressive results in many fields in recent years.

Machine learning has been applied in quite a few places, including supply chain management, big data sorting and recommendation, fraud prevention, personalized functions and so on.

Deep learning is a research direction in the field of machine learning. It is based on artificial neural networks to simulate human intelligence, analyze and classify data, and enable it to have the ability to discover potential patterns. Deep learning includes supervised learning, unsupervised learning and reinforcement learning. Its application fields include NLP natural language processing, CV computer imaging, autonomous driving, medical image analysis.

Regression analysis is a reliable way to identify which variables influence the results. Most of them aim to confidently determine which factors are most important, which can be ignored, and how these factors interact with each other. The most common form of regression analysis is linear regression, but there are many other varieties that can match different types of problems in daily life.

In machine learning, support vector machines are supervised maximum margin models with associated learning algorithms. It is a two-class classification model, the most basic one is a linear classification with maximum margin defined on the feature space. The learning strategy of SVM is to maximize the interval, which is formalized as a problem of solving convex quadratic programming, which is also equivalent to the minimization problem of the regularized hinge loss function. Its basic idea is to solve the separation hyperplane that can correctly divide the training data set and have the largest geometric interval. The SVM learning algorithm is very suitable for solving convex quadratic programming.

The K-Nearest Neighbor algorithm is a supervised learning algorithm that can be used for both classification and regression. The KNN algorithm does not have a learning process in the general sense. Its algorithm uses training data to divide the feature vector space, and uses the division results as the final algorithm model. The KNN algorithm essentially outputs the category label or predicted value corresponding to the feature vector for any n-dimensional input vector corresponding to a point in the feature space. The principle of KNN is to determine which category a new sample belongs to based on the categories of the K closest sample points to the new sample, which is somewhat similar to a minority-majority voting.

3. Application

3.1. Current status of agricultural issues

Agriculture today faces many problems. For example, global population growth has led to increased demand, climate change, environmental changes, plant diseases, insect pests, natural disasters, overly rigid irrigation and sowing methods, and insufficient mechanization of agricultural machinery. These problems have caused excessive losses in agricultural products and have brought problems to agricultural development. Big inconvenience.

Faced with these problems, the traditional method of relying on personal and other people's experience is no longer enough to solve it. If the method is not improved, it will be impossible to meet the demand, let alone develop agriculture stably. Therefore, there is an urgent and important need to optimize and improve from all aspects and find new methods.

One direction that can be worked hard and explored is to use machine learning to help agricultural development. Using machine learning to prepare before agricultural production, improving efficiency during agricultural production, and improving and ensuring quality after agricultural production, the combination of these three methods can provide help and direction to solve this problem.

3.2. Production forecasting and environmental analysis

The first direction of machine learning application is predicting and analyzing environmental resources, climate change and crop yields before agricultural production.

Water is a necessary component of organisms' daily metabolic activities and is an indispensable component of plant growth and maintenance of transport mechanisms. The main function is to dissolve nutrients and transport them to all parts of the plant.

Soil is a key determinant of agricultural production. There are subtle connections between precipitation, soil structure, water availability and plant growth. The interaction of these aspects results in the complex properties of soil moisture. The slightest difference can trigger a series of biological and chemical reactions.

The interaction of soil and water affects plant uptake of nutrients. When soil pores are saturated, the diffusion of water is blocked, reducing the soil's ability to sustain crop growth. When soil pore saturation is too low, the microbial flora in the pores is susceptible to serious damage. When the soil pores are too dry, the water channels in the soil pores are easily disconnected [1]. These conditions will affect plant growth, so we need to predict the soil.

One of the properties that has a strong influence on soil parameters, soil moisture content, is the amount of water contained per unit volume of soil. This is closely related to carbon allocation, plant growth, nutrient cycling, photosynthetic rates and microbial activity. The most important thing in predicting soil is to find out its soil moisture content.

There are many methods for measuring SWC, each with its own advantages and disadvantages, such as remote sensing inversion of soil moisture content based on hyperspectral features, long short-term memory neural network (LSTM) model combined with random sampling learning methods for multiple predictions. One of the methods for measuring SWC is the gravimetric method (GM), which is the most basic method and the so-called absolute method [1].

Machine learning has considerable advantages in soil prediction because it is smart enough. Moreover, compared with destructive methods such as GM and neutron methods that may lead to false positive results, machine learning predictions are obviously more accurate and effective .

In terms of soil texture prediction, traditional soil texture analysis draws conclusions through experiments and statistical analysis. However, this method is not stable and accurate, time-consuming and expensive.

But artificial intelligence technologies such as machine learning clearly have considerable advantages. These algorithms make predictions from data sets, reducing cost, time and labor, and are more stable and accurate.

In addition to this, the geological structure surrounding the planting area is also very important.

In order to predict whether a landslide will occur, the number of landslides that occurred in previous years, geology, landforms, topography, climate and other related factors can be used as parameters to predict and generate models to find different solutions through the prediction results. Landslide susceptibility maps are a good way to predict landslides. In machine learning, the random forest (RF) algorithm is an ML algorithm designed for non-parametric multivariate classification, extreme gradient boosting (XGBoost) is a combination of gradient boosting algorithm and decision tree model, and light gradient boosting (LightGBM) is a A gradient boosting method, classification boosting (CatBoost) is another improved gradient boosting technology, these algorithms can generate landslide susceptibility maps [2-4].

Predicting crop yields is equally important. Crop yield prediction can help predict crop yields and determine when to sow and when to harvest based on the prediction results to obtain better crop yields.

Traditional crop yield prediction methods include reference to empirical knowledge, crop growth models and remote sensing methods. Although these methods are quite stable and applicable to many different crops and locations, they are difficult to apply on a large scale [5].

The combination of machine learning and drones has more advantages than traditional methods.

Yield estimation via an end-to-end 3D CNN framework using multi-temporal drone-based RGB images. The authors' research shows that multi-temporal observations of UAV RGB images provide sufficient information for a 3D CNN architecture to not only accurately predict yields, but also do so in a non-destructive manner.

3.3. Improve the quality of agricultural products

The second direction of machine learning application is in agricultural production to improve the quality and quantity of agricultural products and the efficiency of agricultural production by optimizing multiple links in the crop growth process.

Machine learning uses data to monitor crop growth, detect plant diseases and pests, improve irrigation methods, predict crop yields, and improve agricultural machinery.

Plant diseases are the destruction of plant tissue structure due to certain influences. Plants generally show discoloration, necrosis, rot, wilting or deformity. Pest pests are the damage caused to crops by various insects including mites. Plant diseases and insect pests will be extremely destructive to agricultural production and harvest, seriously damage the quality of agricultural production and crop yields, and are a very serious threat to food security.

In terms of plant diseases and insect pests, the traditional method is manual analysis and judgment. This method relies on personal experience and cannot guarantee sufficient accuracy. Machine learning has already achieved considerable research results in this area.

An example of machine learning for plant pest and disease detection is an ultra-lightweight and efficient network for image-based detection of plant pest and disease infections. The architecture of this network uses residual deep convolution and spatial pyramid pooling to extract plant disease-related features and classify deep features [3]. Compared with other advanced architectures, although this architecture has lower classification accuracy, the computing power consumption of the network is significantly reduced.

Irrigation is a vital part of agricultural production. Irrigation can replenish moisture in the soil, provide necessary growth moisture for crops, improve soil aeration, and promote the dissolution and release

of nutrients in the soil. Reasonable irrigation can also regulate soil temperature and create a suitable growing environment for crops.

However, a huge challenge facing agriculture is water shortage. The world is facing a water crisis, and it is necessary to improve irrigation methods to protect the environment.

The traditional irrigation method is surface irrigation, which is too extensive and wastes a lot of water resources. Machine learning can help agriculture predict irrigation water volume and schedule irrigation.

Evapotranspiration is one of the most important indicators in irrigation water management. In terms of predicting evapotranspiration, machine learning not only has high accuracy, but also has a variety of prediction methods, such as support vector machine (SVM), random forest (RF) and generalized regression neural network (GRNN), hybrid LTSM and CNN model [6].

Automation and intelligence of agricultural machinery are very important. Automated and intelligent agricultural machinery are indispensable for improving agricultural production efficiency, effectively utilizing and saving resources, and promoting agricultural development.

Today's agricultural machinery still requires human control, which not only increases costs, but also may cause problems due to human error.

Automated and intelligent agricultural machinery based on machine learning can reduce manpower, help optimize and conserve resource use, increase productivity and profitability, and reduce environmental impact.

For example, detection systems monitor changes in collected sensor data to automatically detect anomalies in machines and processes. When a disc mower is deployed and used in a location where it may hit a hard solid object, the disc mower is easily damaged.

Recent advances in sensing technology and improvements in computing power have accelerated the development of condition monitoring systems based on different methods. Leveraging machine learning, there are two deep learning models for intelligent condition monitoring of disc mowers to notify the machine operator in the event of a malfunction, namely Convolutional Neural Network (CNN) and Residual Neural Network (ResNet), By comparing the ResNet model, it can better distinguish monitoring and have higher accuracy.

This example illustrates that machine learning can be of considerable help in the automation and intelligence of agricultural machinery.

3.4. Agricultural product classification detection and price prediction

The third application direction of machine learning is the classification and quality detection of agricultural products and the price prediction of agricultural products. Using machine learning to deal with post-agricultural production problems can effectively reduce the time and money costs of manual processing.

The classification of agricultural products is very important. Traditional methods require manual inspection of each sample, which is laborious, time-consuming, and prone to human error. Machine learning classification methods have more advantages. Not only is it highly accurate, it takes less time, and it is non-destructive.

An example of the application of machine learning to the classification and detection of agricultural products is that Emrah Dönmez conducted a comparative study on the hazelnut variety identification experiment of three models: BigTransfer(BiT)-M R50×1, BiT-M R101×3 and BiT-M R152×4. result. Among them, the BiT-M R152×4 model has the highest accuracy [7].

There is also wheat hybrid identification, which is more difficult than ordinary crop classification because the differences between different hybrids are much smaller than the differences between different types of crops.

Both deep learning models, MobileNetv2 and GoogleNet, have high accuracy for classifying wheat varieties and hybrids. Mesut Ersin Sonmez innovatively used deep learning for feature extraction, combining convolutional neural networks (CNN) and support vector machines (SVM) to further improve the accuracy of wheat hybrid identification [8].

Emrah Dönmez developed a hybrid model based on the visual transformer (VT) method and the convolutional neural network (CNN) model, modified the ResMLP architecture and fine-tuned and improved the EfficientNetV2b0 model. By combining the two methods, the hybrid transformer model developed, the accuracy of wheat hybrid identification is also very high [9].

Judging from these studies, machine learning is of great help in classifying agricultural products.

Pesticides play an important role in modern agriculture and are used to protect crops from pests and diseases. However, pesticides can also cause environmental pollution and leave toxic residues on crops. Therefore, in order to protect individuals and ecosystems from the harm of residual pesticides, agricultural products Quality inspection is extremely important.

Traditional detection methods are spectrometry, chromatography, and in vivo detection. These methods will not only cause certain damage to agricultural products. Machine learning algorithms are more beneficial than traditional methods for predicting pesticide toxicity.

Many algorithms in machine learning can detect toxicity. Ganesan Anandhi and M. Iyapparaja compared SVM, k-NN, ANN, CNN, LDA, DQA and RF, which are all algorithms based on regression models [10]. These methods that leverage machine learning are more efficient and less disruptive than traditional methods.

Agricultural product price prediction is also an important need in agriculture. By predicting agricultural product prices, farmers can decide on agricultural production planning.

Murat Sari combined the extreme learning machine method and the genetic algorithm to create a novel commodity price prediction model, the GA-ELM model [11,12]. This model not only has the accurate prediction ability of machine learning, but also requires less training data than some other deep learning models.

Prilly Oktoviany, Robert Knobloch and Ralf Korn proposed a two-step hybrid model that, in addition to identifying historical price states and predicting external factors, predicts future price states through K-nearest neighbors and random forest algorithms.

Both models have been used well in predicting prices in multiple fields, demonstrating that machine learning has many advantages for predicting prices.

4. Discusstion

As the world's population continues to increase, the demand for agricultural production also increases, so people are trying to find and adopt newer and better methods and technologies to develop agriculture.

Machine learning can learn and progress like humans by continuously analyzing large amounts of data, improving the accuracy of predictions bit by bit, thereby making predictions for more complex and unknown data. Therefore, machine learning has great advantages to help agricultural development and progress.

Machine learning can optimize all aspects of agricultural production through prediction and detection, thereby improving the efficiency and quality of agricultural production. Agriculture can also be made more precise by collecting and analyzing soil conditions, climate conditions, crop growth, etc., thereby saving resources, reducing waste, and protecting the environment. Machine learning can also predict the occurrence of pests and diseases, weather changes, natural disasters, wild animals, and other risk factors, and it can respond in time and reduce losses through early warning.

With the continuous development of machine learning, machine learning may even promote the advancement of breeding technology in agriculture and collaboratively improve current crop varieties, thereby helping agriculture take a big step forward.

However, while machine learning can help agricultural progress, there are still some challenges that need to be solved.

One problem is that applying statistics and quantification in agriculture is very difficult. Even if the areas are adjacent, factors such as weather and soil may vary greatly, and thousands of influencing factors need to be taken into consideration.

Similarly, agricultural data is difficult to collect. Even if the data can be collected, temporal data is still scarce and difficult to manage due to its seasonality. Machine learning requires a large amount of data practice learning to make accurate predictions.

Another problem is that current research on machine learning and agriculture is still expensive, and the difficulty in popularizing it means that it is difficult to make progress in agriculture as a whole.

In addition, machine learning technology is quite complex, and related maintenance requires personnel with professional knowledge. This will also make it difficult to promote new research and development and make it difficult to make overall progress in agriculture.

However, despite many challenges, machine learning can still help agricultural development. And I believe that with the development of machine learning, these challenges will be solved.

5. Conclusion

This article explores how the application of machine learning in agriculture can advance agricultural development. In terms of preparation before agricultural production, machine learning can predict crop yield, soil prediction, and landslide disaster prediction. In terms of agricultural production, the residual deep convolution and spatial pyramid pooling algorithms in machine learning can be used for ultra-lightweight and efficient plant disease and pest detection networks, significantly reducing the computing power consumption of the network. In terms of agricultural production, machine learning can also predict the amount of irrigation water and realize the intelligence and automation of agricultural machinery. In terms of post-agricultural production, machine learning can perform quality inspection, classification and price prediction of agricultural products.

Machine learning has been applied to many aspects of agriculture and promoted the development of agriculture. However, these applications still have more or less defects and areas for improvement, and these applications need to be researched, improved and optimized in the future. The significance of this study is that the application of machine learning in agriculture can not only save resources, protect the environment, and promote agricultural development, but also put forward new demands for machine learning and promote the development of machine learning. Therefore more further research on this type of research should be conducted.

References

[1] Awais, M., Naqvi, S.M.Z.A., Zhang, H. et al. AI and machine learning for soil analysis: an assessment of sustainable agricultural practices. Bioresour. Bioprocess. 10, 90 (2023). https: //doi.org/10.1186/s4064 3- 023 - 00710 - y.

[2] Usta, Z., Akıncı, H. & Akın, A.T. Comparison of tree-based ensemble learning algorithms for landslide susceptibility mapping in Murgul (Artvin), Turkey. Earth Sci Inform 17, 1459 – 1481 (2024). https://doi.org/10.1007/s12145-024-01259-w.

[3] Wang, B., Zhang, C., Li, Y. et al. An ultra-lightweight efficient network for image-based plant disease and pest infection detection. Precision Agric 24, 1836 – 1861 (2023). https: //doi.org/10.1007/s11119 - 023 - 10020 - 0.

[4] Mokhtar, A., Al-Ansari, N., El-Ssawy, W. et al. Prediction of Irrigation Water Requirements for Green Beans-Based Machine Learning Algorithm Models in Arid Region. Water Resour Manage 37, 1557 – 1580 (2023). https: //doi.org/10.1007/s11269 - 023 - 03443 - x.

[5] Bhadra, S., Sagan, V., Skobalski, J. et al. End-to-end 3D CNN for plot-scale soybean yield prediction using multitemporal UAV-based RGB images. Precision Agric 25, 834–864 (2024). https://doi.org/10.1007/s11119-023-10096-8.

[6] Michael Jaumann, Ertug Olcay, Timo Oksanen. Condition Monitoring using Convolutional Neural Network in Agricultural Machinery - Use Case: Disc Mower. IFAC-PapersOnLine. Volume 55, Issue 32. 2022. Pages 235 - 240.

[7] Dönmez, E., Kılıçarslan, S. & Diker, A. Classification of hazelnut varieties based on bigtransfer deep learning model. Eur Food Res Technol 250, 1433–1442 (2024). https: //doi.org/10.1007/s00217 - 024 - 04468 - 1.

[8] Sonmez, M.E., Sabanci, K. & Aydin, N. Convolutional neural network-support vector machine-based approach for identification of wheat hybrids. Eur Food Res Technol 250, 1353 – 1362 (2024). https: //doi.org/10.1007/s00217 - 024 - 04473 - 4.

[9] Dönmez, E. Hybrid convolutional neural network and multilayer perceptron vision transformer model for wheat species classification task: E-ResMLP+. Eur Food Res Technol 250, 1379 – 1388 (2024). https: //doi.org/10.1007/s00217 - 024 – 04469 - 0.

[10] Ganesan Anandhi, M. Iyapparaja. Systematic approaches to machine learning models for predicting pesticide toxicity. Heliyon. Volume 10, Issue 7. 2024. e28752.

[11] Oktoviany, P., Knobloch, R. & Korn, R. A machine learning-based price state prediction model for agricultural commodities using external factors. Decisions Econ Finan 44, 1063 – 1085 (2021). https: //doi.org/10.1007/s10203 - 021 - 00354 - 7.

[12] Sari, M., Duran, S., Kutlu, H. et al. Various optimized machine learning techniques to predict agricultural commodity prices. Neural Comput & Applic (2024). https: //doi.org/10.1007/s00521-024-09679 - x.