Review

Knowledge Mapping of Machine Learning Approaches Applied in Agricultural Management—A Scientometric Review with CiteSpace

Jingyi Zhang, Jiaxin Liu, Yaqi Chen, Xiaochun Feng * and Zilai Sun

College of Economy and Management, Northwest A&F University, Xianyang 712100, China; 2018012805@nwafu.edu.cn (J.Z.); 2018013443@nwafu.edu.cn (J.L.); chenyaqi@nwafu.edu.cn (Y.C.); sunzilai@mail.dlut.edu.cn (Z.S.)
* Correspondence: xiaochunf@nwafu.edu.cn

Abstract: With the continuous development of the Internet of Things, artificial intelligence, big data technology, and intelligent agriculture have become hot topics in agricultural science and technology research. Machine learning is one of the core topics in artificial intelligence, and its application has penetrated every aspect of human social life. In modern agricultural intelligent management and decision making, machine learning plays an important role in crop classification, crop disease and insect pest prediction, agricultural product price prediction, and other aspects of management and decision-making processes in agriculture. To detect and recognize the latest research developing features in a quantitative and visual way, and based on machine learning methods in agricultural management, the authors of this paper used CiteSpace bibliometric methods to analyze relevant studies on the development process and hot spots. High-value references, productive authors, country and institution distributions, journal visualizations, research topics, and emerging trends were reviewed and analyzed. According to the keyword visualization and high-value references, machine learning approaches focus on sustainable agriculture, water resources, remote sensing, and machine learning methods. The research mainly focuses on six topics: learning technology, land environment, reference evapotranspiration, decision support systems for river geography, soil management, and winter wheat, while learning technology has been the most popular in recent years.

Keywords: emerging trends; visualization; research hotspot; Web of Science

1. Introduction

Smart agriculture is a highly integrated product of agriculture and technology. The main applied technologies include sensor networks, big data, cloud computing, the Internet of Things, and so forth. Their main functions are to collect, process, and analyze the relevant data in agricultural production and operations so as to provide guidance and intelligent decision-making for the corresponding agricultural production. Machine learning algorithms are a common technology in the application of intelligent agriculture. Their main function is to use people's prior knowledge to build different algorithm models so as to realize the prediction of unknown things. The classification of crops, the identification and prediction of pests and diseases, and the price prediction of agricultural products are three crucial aspects for the development of smart agriculture. A large number of research results show that machine learning models (mainly including artificial neural networks, Bayesian models, deep learning, dimensionality reduction, decision trees, ensemble learning, and support vector machines) are widely used in the above three aspects [1].

Currently, smart agriculture management relies on data collected with high accuracy, and wireless sensor networks have been widely used with the Internet of Things to monitor the soil, water, and crop situation [2]. Data is collected through the use of the Internet of Things, and is then processed by machine learning techniques to increase the quantity and

quality of crop production [3]. Thus, agribusiness can better monitor the status of crops and receive suggestions to help reduce pollution and pesticide use. These technologies not only help create more sustainable agriculture but also make it more productive and increase profits [4]. Except for wireless sensor networks, remote sensing technology [5], information systems [6], and unmanned aerial vehicles [7] all need the support of machine learning methods to contribute to a smart agriculture system. The main limitation of machine learning in agriculture management is the environment, such as greenhouses and large-scale lands. There have been some research reviews on the technical method of agricultural management in academic circles. Hu et al. [8] reviewed the current research on operational management in precision agriculture and digital agriculture and found that most of the research was on technological design, the precision management of agricultural production, and the traceability management of product quality and safety. Sharma et al. [9] summarized the application of machine learning in agricultural supply chains and proposed that decision-makers should cultivate machine learning technology and maintain sustainable input costs to keep companies growing healthily. Jordan et al. [10] believe that machine learning methods have improved the fields of crop management, livestock management, water management, and soil management, which are interrelated and have established knowledge-based intelligent agricultural systems. Mazzia et al. [11] indicate that machine learning could be crucial to precise agriculture and could help achieve cost-effective agricultural mergers with unmanned aerial vehicles. Moreover, core machine learning approaches, such as supervised learning, unsupervised learning, and reinforcement learning, contribute to the emergence of computer vision, speech recognition, and robot control [10], which have supported many agricultural management studies. Additionally, the advantages and disadvantages of different machine learning approaches vary, and each approach has particular features and applications [12], which make them indispensable, revolutionary, and widely used in agricultural management.

The existing research has provided a rich theoretical basis for this paper. However, there is only one review article about machine learning approaches in agricultural management. It reviewed specific machine learning approaches in detail and several applications in agriculture [1]. However, this review did not reveal other important research information. Who are the famous scholars on the application of machine learning methods in agricultural management? Which countries and institutions have close exchanges? What are the research subjects and development trends? These problems need further analysis. Therefore, in order to gain a deep understanding of the current research progress of machine learning algorithms in the application of intelligent agriculture, this paper focuses on the main research results of machine learning algorithms in intelligent agricultural management. CiteSpace, a literature analysis tool, was adopted to analyze 1950 articles and 88,984 citations collected from the Web of Science. The analysis revealed six research aspects: high-value references, productive authors, distributions of countries and institutions, journal visualization, research topics, and emerging trends. It is expected that this study will provide a reference for the exploration of machine learning algorithms more suitable for the development of intelligent agriculture.

2. Research Methods and Data Collection

2.1. Research Methods

Scientometric analysis is an effective method to deal with data and information visualization, which can be used to detect the research frontier and research hotspots as well as to track vital changes in one area [13]. To avoid subjectivity, scientometric analysis mainly uses statistical approaches to analyze, process, and evaluate the quality and the feature of research materials. Based on a large number of articles in the literature database, this method is regarded as a popular tool to learn about one subject, summarize the development path, and predict the future trend. Scientometric analysis is used for many literature analysis tools, including CiteSpace.

CiteSpace, a Java-based application, was created by Chaomei Chen, a famous scholar focused on information visualization, knowledge mapping, and science frontier atlases. CiteSpace is a strong tool to systematically learn about one field rapidly. The principle of CiteSpace is to label co-citation clusters and then use time-sliced snapshots to form timeliness and pivotal points [14]. CiteSpace 5.7.R2 (64-bit), which is run in the Java 8 environment, was used for this paper. The new version is equipped with 3 central concepts: Kleinberg's burst detection, Freeman's betweenness centrality metric, and heterogeneous networks. This means that three key problems can be better solved: (1) identifying research fronts, (2) labeling specialties, and (3) detecting emerging trends and abrupt points [15]. Moreover, Professor Chaomei Chen provided us with an excellent research paradigm in regenerative medicine, demonstrating the optimistic future of this field [16]. The number of publications using CiteSpace has drastically increased in recent years, and the program can be attained at http://cluster.ischool.drexel.edu/~cchen/citespace/download/ (accessed on 20 March 2021).

2.2. Data Collection

The most crucial component of a review article is data collection, and the quantity and quality of relevant article material directly determine the visualization effect and the quality of the review article. The Web of Science is one of the main citation index databases. It mainly includes the science citation index, social science citation index, and arts and humanities citation index, providing an effective platform to conduct scientific research. In order to find the relevant and proper articles, we followed specific steps to collect and select data: (1) choose the Web of Science Core Collection; (2) use the advanced search setting; (3) follow the search conditions listed in Table 1 and select the final research materials.

Table 1. Detailed search setting parameters.

Source	Web of Science Core Collection
Citation	SCI-EXPANDED, SSCI, CCR-EXPANDED, IC
Search steps	#1 = (TS = ((machine learning AND agricultur* manag*) OR (neural network AND agricultur* manag*) OR (bayes AND agricultur* manag*))) AND language: (English) #2 = (TS = ((machine learning AND agricultur* manag* AND IOT) OR (machine learning AND agricultur* manag* AND Internet of Things) OR (neural network AND agricultur* manag* AND IOT) OR (neural network AND agricultur* manag* AND Internet of Things) OR (bayes AND agricultur* manag* AND IOT) OR (bayes AND agricultur* manag* AND Internet of Things))) AND language: (English) #3 = (TS = ((supervis* learn* AND agricultur* manag*) OR (reinforc* AND agricultur* manag*) OR (decision tree AND agricultur* manag*) OR (algorith* AND agricultur* manag*) OR (boost* AND agricultur* manag*) OR (deep learning AND agricultur* manag*))) AND language: (English) #4 = #3 OR #2 OR #1
Time span	2000–2020
Qualified Records	1950
References	88,984

* denotes any group of characters, including null characters. "agricultur*" stands for agriculture, agricultural. "manag*" stands for manage, management.

After removing duplicates, 1950 qualified records and their 88,984 references were retained, and 17 records and 521 invalid references were deleted.

3. Visualization Results

First, in order to exactly describe the co-occurrence, relationship, and citation information, basic parameters needed to be set properly before systematically analyzing the

trends and research status. The selection criteria were set as the top N strategy, which means that the top 50 levels of the most cited and high-frequency items were selected from each slice. Moreover, the fundamental value of examining map significance could be measured by two values: the degree of modularity (Q value), which is the modularity value assessing the structure of certain material, and the mean silhouette value (S value), representing the degree of clustering. If the Q value is >0.3 and the S value is >0.7, it means that the community structure is favorable and the clustering result is acceptable. The results indicate that the Q value is 3.463 and the S value is 0.746, demonstrating that the visualization effect is fine [17].

The following specific functions were chosen to assess the related review results: (1) reference visualization, (2) productive authors, (3) country and institution distribution, (4) published journals, (5) keywords and their timeline trends, and (6) category visualization.

4. Reference Visualization

A total of nine typical and influential articles about machine learning methods applied in agricultural management were selected through citation frequency, where the threshold was set as 18. These articles could, to some extent, represent a trend. The visualization result is shown in Figure 1.
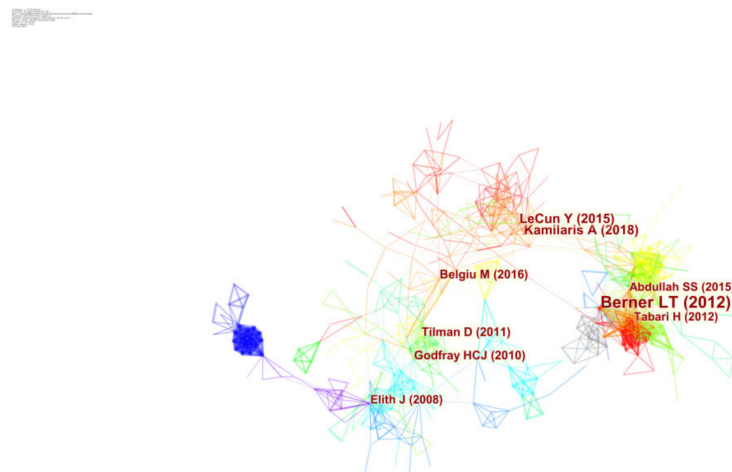
Figure 1. Reference visualization.

Among the 1950 selected articles, the articles were mainly compiled by Berner, who has the highest citation frequency. Among the articles, Berner et al. [18] were the first to use a machine learning method to address agriculture and environmental problems. For the article, published in *Biogeosciences* in 2012, a typical machine learning approach, the random forest, was used to process and predict a large-scale dataset and solve the specific forest problem in northern Siberia. The results showed that machine learning approaches could be applied to solve difficulties in various fields, and that they have advantages, such as low bias, the ability to handle large datasets, and not overfitting models compared to traditional methods. This article is significant in leading the trends of machine learning approaches applied in the agricultural field.

An excellent work published in *Nature*, the article "*Deep learning*", written mainly by LeCun, a scholar from New York University, is one of the most influential and initial works in this field; and the article led to a new approach to address speech recognition, visual object recognition, and other fields. Moreover, the principle and classification of deep learning are interpreted clearly [19]. To some extent, the deep learning neural network method has been widely used to solve specific agricultural problems.

An excellent work mainly by Kamilaris conducted a real application survey of deep learning approaches in agriculture. The advantages of deep learning, such as its high accuracy, use of simulated datasets, and shorter testing time, and the disadvantages of

deep learning, such as its inability to handle large datasets and being unable to exceed the boundaries of datasets, are discussed. In addition, as a promising machine learning approach, Kamilaris thinks that deep learning would have more application functions than land cover classification and weed detection [20].

It goes without saying that the constantly increasing population and consumption will lead to a lack of food and excessive greenhouse gas emissions for decades. The excessive exploration of land, water, energy, and fishing will not ensure sustainable agricultural development and food safety. Godfray outlined several aspects waiting to be solved in agriculture and food fields [21], while Tilman et al. [22]. stated that the sustainable intensification of agriculture is also of great significance. Determining how to get along well with the environment will have a profound influence on our future, which motivates many scholars to further probe specific aspects with various approaches, including machine learning.

Remote sensing is one of the key technologies in agricultural management, especially in the agricultural IOT [23]. Published in the *ISPRS Journal of Photogrammetry and Remote Sensing*, Belgiu's article "*Random forest in the remote sensing: A review of applications and future directions*", which expounded on the principles of the random forest approach and its application in remote sensing, has received wide attention. In addition, a comparison with other machine learning approaches and the future trends are discussed to achieve better applications and further research [24].

Elith et al. published an article named "*A working guide to boosted regression trees*" that introduced a strong algorithm with powerful ecological insight [25]. Boosted regression trees act as an automatic approach to recognize and predict relevant variables. Compared with traditional methods, boosted regression trees have huge predictive advantages and might be broadly used in agricultural management.

Finally, regarding evapotranspiration in arid environments, Tabari et al. stated that evapotranspiration becomes imperative when considering water resource management. Moreover, he evaluated climate data using machine learning approaches such as the Support Vector Regression, Adaptive-Network-Based Fuzzy Inference System, and Mixed Logistic Regression, and then expressed the criteria to examine the errors [26]. Similarly, the article "*Extreme Learning Machines: A new approach for prediction of reference evapotranspiration*" published in the *Journal of Hydrology* is a key point for machine learning approaches use in agriculture. The authors strongly recommend the Extreme Learning Machine algorithm due to its efficiency, simplicity, speed, and compatibility. This approach proved valid to accurately solve the reference evapotranspiration problem [27].

As seen above, the leading papers are mainly classified in the following four areas:

(1) Specific machine learning approaches, especially the random forest, deep learning, and extreme learning. These new approaches are now leading a hot research trend in agricultural applications.

(2) Sustainable agricultural development. How to address food shortages and how to solve intensive agricultural problems are becoming increasingly more significant for humans to survive well with limited resources.

(3) Remote sensing. Remote sensing technology is now widely used in agricultural management, especially in the agricultural Internet of Things and precise agriculture.

(4) Water resources. As the fundamental factor of agricultural operations, water resources are very important for crop yields and plant quality. Further research on water could greatly improve and push the progress.

5. Contributing Authors Analysis

The ten most productive authors were detected through CiteSpace. In this section, the ten most productive authors are introduced to help better understand the current research status. The results are shown in Figure 2.
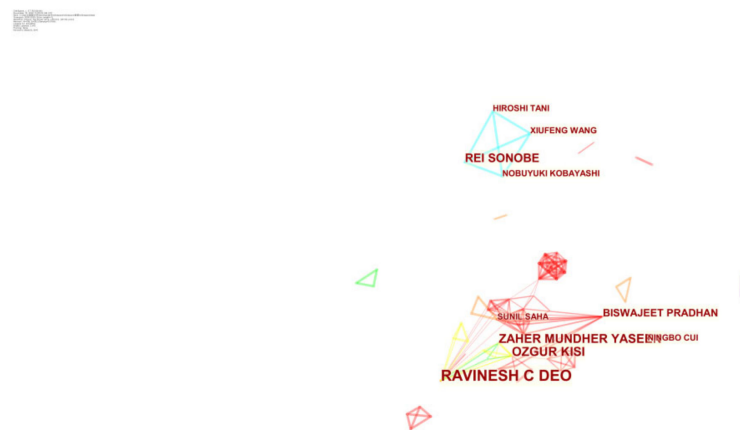
Figure 2. Productive authors.

Ravinesh Deo, a professor at the University of Southern Queensland, is first, due to his highly cited articles; many of his articles were searched. His main research fields include machine learning, algorithms, artificial intelligence, and agriculture, especially machine learning approaches. His great contributions to agriculture are reflected in several of his articles. For the article "*Soil moisture forecasting by a hybrid machine learning technique: ELM integrated with ensemble empirical mode decomposition*" [28], he used a hybrid machine learning algorithm to predict soil moisture, which is crucial to the agricultural system. Moreover, a multistage probabilistic machine learning model was designed to build an accurate rainfall prediction system, improving the constant drought risk of agricultural production [29].

Z.M. Yaseen, who is from the National University of Malaysia, is especially good at advanced data analysis and machine learning approaches. Due to his high enthusiasm for machine learning, he has contributed greatly to the agricultural applications of machine learning approaches. Published in Agriculture Water Management, his work "*Reference evapotranspiration prediction using hybridized fuzzy model with firefly algorithm: Regional case study in Burkina Faso*" has had a great influence on pushing machine learning approaches into agriculture [30].

The articles of Ozgur Kisi, from Llia State University, have been cited over 18,000 times. Moreover, he is one of the main participants of the influential masterpiece above, LeCun et al. [20]. He has achieved several great works by using machine learning approaches to solve evapotranspiration problems. His article "*Modelling reference evapotranspiration using a new wavelet conjunction heuristic method: wavelet extreme learning machine vs wavelet neural networks*" published in *Agricultural and Forest Meteorology* provided a unique thought to measure the best model [31].

The research fields of Rei Sonobe, a scholar at Hokkaido University, Japan, are mainly focused on crops and machine learning. There were also three members detected on his team: Xiufeng Wang, Hiroshi Tani, and Nobuyuki Kobayashi. He combines remote sensing, agricultural management, and machine learning. In his article "*Random forest classification of crop type using multi-temporal TerraSAR-X dual-polarimetric data*", he compared two algorithms and achieved high accuracy with classifying various crops [32]. He also published several similar studies using different machine learning methods [33]. Similarly, Biswajeet Pradhan, from the University of Technology Sydney, is also keen on remote sensing and machine learning in land cover [34], water pollution [35], soil erosion [36], and natural hazards [37]. Sunil Saha from the University of Gour Banga is also an important participant in this field.

Ningbo Cui, a scholar from Sichuan University, has especially mastered fruit quality [38], crop water resources [39], and machine learning. His work presents a mix of these topics that could lead to further research on applying machine learning.

6. Country and Institution Distribution

The visualization results when the threshold is set as 119 are shown in Figure 3. There are two aspects that need to be addressed. One aspect is the betweenness centrality, which reflects the importance of nodes in the visualization network. The sizes of the circles reflect the centrality in Figure 3. The other aspect is the lines between two countries, representing the collaborative relationships of these countries.

Figure 3. Country distribution.

Among the 1950 articles from 2000 to 2020, the United States ranks first, contributing the most to this field. Since the primary development of machine learning methods in agriculture began in the United States, the country has achieved early regulation and standards. Beginning in 2003, and over the 17 years thereafter, the United States published 467 papers. A large contribution was from China, whose number of publications has reached 274, although its starting time was much later than those of the other main contributing countries. Australia, which ranks third in this field, accounts for merely 10% of the total publications. The following countries are Spain (144), England (133), Iran (131), and Germany (119). More details are shown in Table 2. Overall, the United States is still the pioneer, while China is second. The other countries are much lower than these two. The research centers and leaders are still in the United States and China.

Table 2. Distribution of countries.

Country	Frequency	Country	Frequency
USA	467	Iran	131
China	274	Germany	119
Australia	187	Italy	102
Spain	144	France	100
England	133	India	100

Figure 4 shows the institutional distribution with a threshold of 20, which could help recognize the important institutions and learn about the correlation of different institutions.

The ranks and specific institution names are listed in Table 3. The results show that America, China, and Canada have the top three positions. Although America has the most institutions, most publications are from the Chinese Academy of Sciences, China.
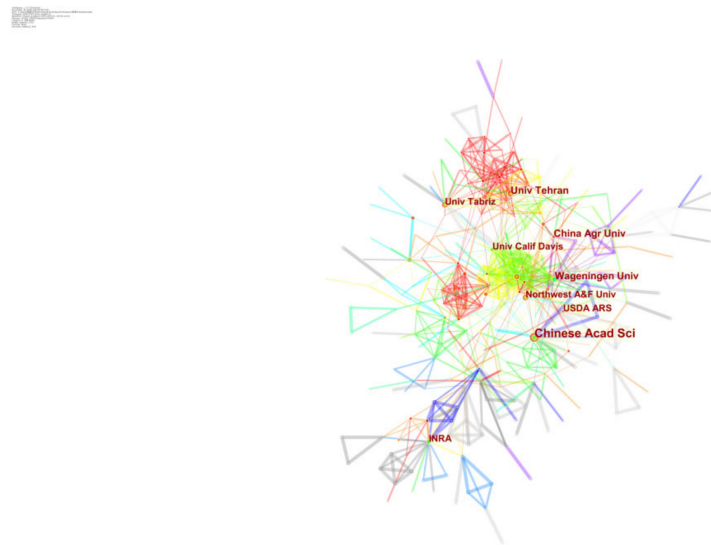
Figure 4. Institution distribution.

Table 3. Top 15 institutions.

Abbreviated Institution Name	Full Name of Institutions	Number of Publications	Affiliated Country
Chinese Acad Sci	Chinese Academy of Sciences	50	China
Univ Tehran	University of Tehran	27	Iran
Wageningen Univ	Wageningen University and Research	24	The Netherlands
China Agr Univ	China Agricultural University	23	China
INRA	Inland Northwest Research Alliance	22	France
Univ Tabriz	University of Tabriz	22	Iran
Univ Calif Davis	University of California, Davis	21	America
USDA ARS	United States Department of Agriculture Research Institute	20	America
Northwest A&F Univ	Northwest Agriculture and Forest University	20	China
Univ Wisconsin	University of Wisconsin	19	America
Univ Southern Queensland	University of Southern Queensland	17	Australia
SCIC	Saskatchewan Crop Insurance Corporation	16	Canada
Agr & Agri Food Canada	Agriculture and Agri-Food Canada	16	Australia
Ton Duc Thang Univ	Ton Duc Thang University	15	Vietnam
McGill Univ	McGill University	15	Canada
Univ Nebraska	University of Nebraska	14	America
CSIRO	Commonwealth Scientific and Industrial Research Organization	14	Australia

7. Cited Journals

To systematically learn about the publication status, the cited journal networks are shown in Figure 5. The most cited journal is *Agriculture Ecosystems & Environment*, which has been cited 508 times. The articles identified from *Science* and *Nature* accounted for 463 citations and 426 citations, respectively. The ranks of the different influential journals are listed in Table 4.
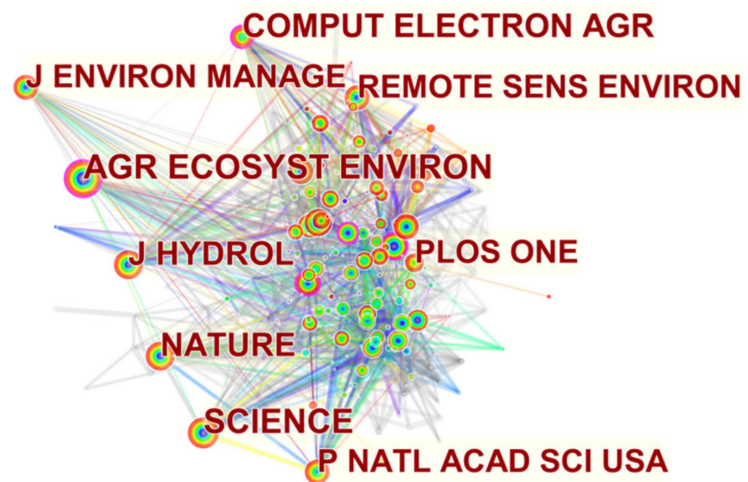
Figure 5. Cited journal networks.

Table 4. Top 15 cited journals.

| Abbreviated Journal Name	Full Journal Name	Frequency
AGR ECOSYST ENVIRON	Agriculture Ecosystems & Environment	508
SCIENCE	Science	463
NATURE	Nature	426
J HYDROL	Journal of Hydrology	413
COMPUT ELECTRON AGR	Computers and Electronics in Agriculture	391
P NATL ACAD SCI USA	Proceedings of the National Academy	354
PLOS ONE	Plos One	350
REMOTE SENS ENVIRON	Remote Sensing of Environment	348
J ENVIRON MANAGE	Journal of Environmental Management	310
SCI TOTAL ENVIRON	Science of The Total Environment	305
INT J REMOTE SENS	International Journal of Remote Sensing	302
AGR SYST	Agricultural System	301
AGR WATER MANAGE	Agricultural Water Management	300
ECOL MODEL	Ecological Modelling	284
WATER RESOUR RES	Water Resources Research	281

8. Popular Topics and Emerging Trends

In this section, popular topics and emerging trends are analyzed through the keyword visualization results. Figure 6 and Table 5 indicate the keywords with high frequencies, including "management", "agriculture", "prediction", "classification", and some machine learning methods. To help better understand the visualization results, Table 4 gives the top 25 most frequent keywords that appeared in the 1950 articles and their 88,984 references, and each keyword has appeared more than 60 times. These keywords could represent the research trends and popular topics to a large extent.

It is evident that "agriculture" and "management" take the first and second positions, respectively. The next several keywords, ordered by frequency, are "models", "artificial neural networks", "systems", "neural networks", "classification", and "prediction", which represent machine learning approaches and some of their applications in agriculture. Next, many keywords related to specific agriculture, such as "land use", "water", "precision agriculture", "yield", "forest" and "vegetation", appear in the high-frequency list. As Table 4 shows, the random forest, one of the main approaches in machine learning, although it started at a later time, still attracted much attention. In addition, it is easy to find that neural network approaches are widely used in agricultural fields. The application of machine learning mainly focuses on precision agriculture, prediction, classification,

yield management, land use, water, forest, and vegetation. Moreover, as a key point of building smart agriculture systems, remote sensing is an inevitable trend for agricultural development.
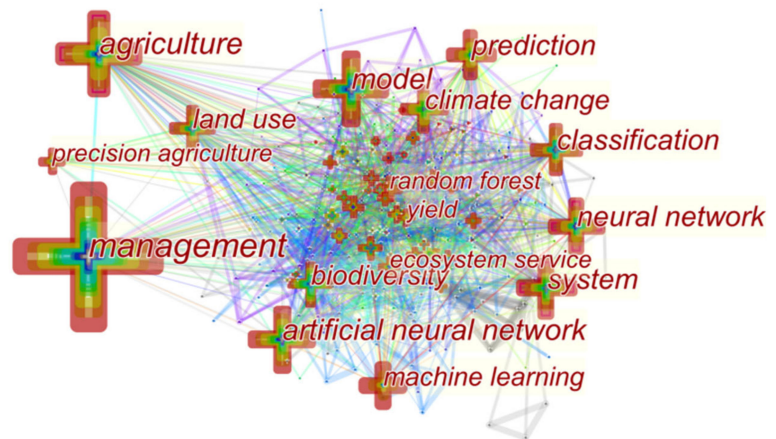
Figure 6. Hot topics and emerging trends.

Table 5. Top 25 research keywords.

Keyword	Year	Frequency
Management	2000	367
Agriculture	2000	224
Model	2004	190
Artificial neural network	2000	185
System	2007	160
Neural network	2000	156
Classification	2000	148
Prediction	2006	140
Biodiversity	2003	135
Climate change	2004	135
Impact	2005	128
Land use	2001	126
Conservation	2003	117
Machine learning	2013	104
Water	2009	85
Random forest	2014	72
Ecosystem service	2012	69
Precision agriculture	2009	68
Yield	2008	66
Forest	2005	65
Vegetation	2008	65
Remote sensing	2008	64
Performance	2008	63
Diversity	2005	60
Nitrogen	2005	60

These keyword networks could also be of great importance to detect the revolutionary trends of this field. To determine the developmental path in the agricultural applications of machine learning, the timeline of burst keywords from 2000 to 2020 is shown in Figure 7, and the burst keywords are classified in six ways.
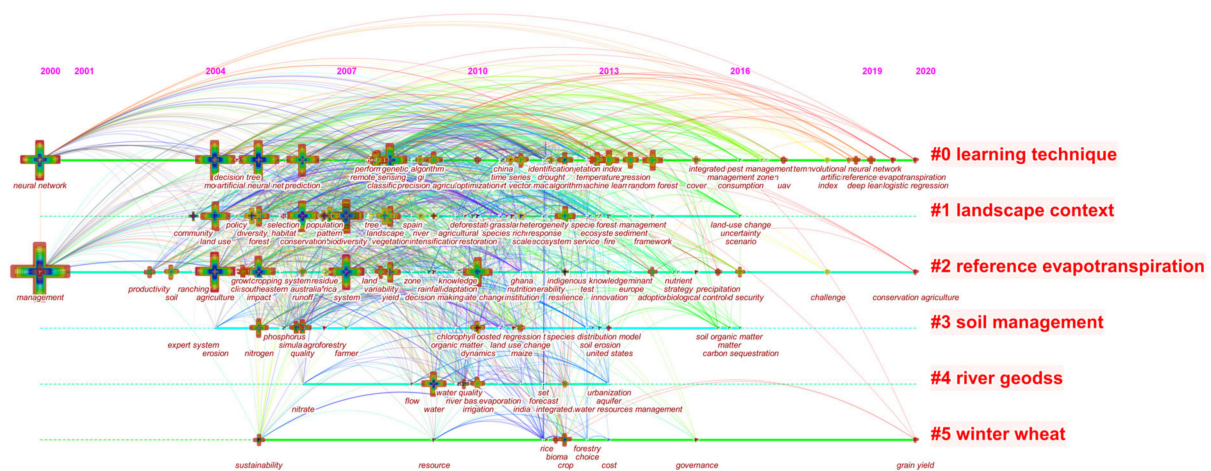
Figure 7. Burst trends of research keywords.

The "learning technique" (#0), initially starting with the neural network in 2000, is the key point of this field. As shown in Figure 7, all the following keywords have a strong relationship with machine learning approaches. The specific machine learning approaches include decision trees, random forests, neural networks, genetic algorithms, deep learning, and logistic regressions. In the last few years, various machine learning techniques have achieved rapid progress and have more application functions, such as pest monitoring and unmanned agriculture vehicles (UAVs).

The "landscape context" (#1), a fundamental component of agricultural management, is inevitable in the further development of agricultural management. This topic mainly discusses land use and the sustainable development of basic agricultural resources, such as land use, communities, forests, conservation, rivers, vegetation, and ecosystems. In 2004–2013, the landscape context was a hot research trend for scholars.

"Reference evapotranspiration" (#2) and the "river GEODSS" (river geographic decision support system) (#4) are two other research hot topics that have a strong relationship to water resource management. Rainfall prediction, nutrition ability, productivity, and soil are mainly included. The burst period of reference evapotranspiration is from 2004 to 2013, while that of river GEODSS is from 2009 to 2010. The primary application of several machine learning approaches is used to evaluate and judge the effect of simulated results [27]. Moreover, it is also a significant part of sustainable development, especially in water resource management.

"Soil management" (#3) is the fourth identified direction of the 1950 articles and their references. It is clear that the major contents of this field are erosion, nitrogen, quality, and organic matter. There were two burst periods of soil management research: 2005–2006 and 2010–2011. Later, the focus of soil management changed more to environmental problems.

Finally, a noticeable directional trend is "winter wheat" (#5). Although different machine learning approaches have their advantages and disadvantages [40], it is unquestionable that the development of machine learning approaches has pushed this direction into a new trend. Additionally, sustainability, resources, crops, rice, and biomass are the main research topics in this direction. Recently, grain yields have been especially popular, which may motivate some new ideas later on.

9. Category Network Visualization

The categories of the 1950 works are visualized to show the relationships of different subjects in this field. As shown in Figure 8, for machine learning approaches applied in agricultural management, there were seven main subjects: multidisciplinary agriculture, agriculture, environmental science, computer science, water ecology, ecology, and engineering. The results indicate that the method of one subject could be applied to another subject, which may bring new research trends and even create new revolutions.
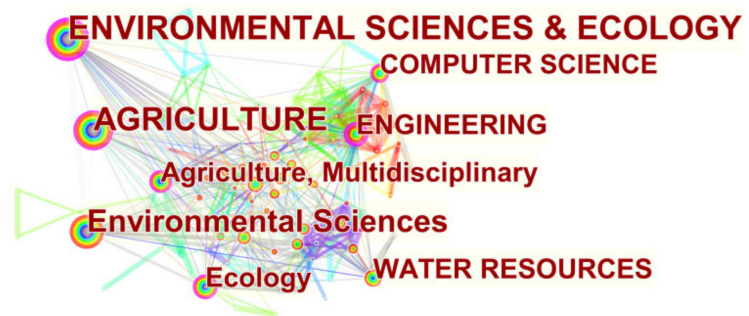
Figure 8. Research categories.

10. Conclusions

A systematic visualization of machine learning approaches and applications in agriculture management is given in this paper. By using the reliable and powerful tool CiteSpace, the author reviewed the current status in six ways. The research material, totaling 1950 articles and their 88,984 references, were all from the Web of Science. The top N strategy was selected as the selection criterion; furthermore, the time span was set from 2000 to 2020. The outcomes are shown in the corresponding part of this article. Moreover, the results are included in the mind map shown in Figure 9.
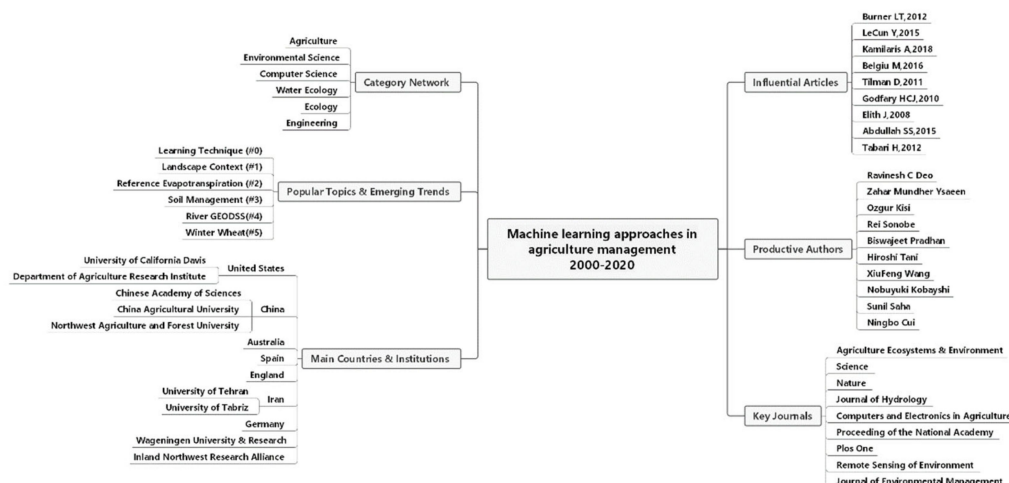

Figure 9. Mind map of this paper.

According to the visualization results, major features can be extracted. The reference analysis shows that the majority of high-value articles focus on specific machine learning technology, water resource management, remote sensing, and sustainable agriculture. According to the productive author analysis, the main research figures are divided into two groups; one is led by Rei Sonobe, while the other is led by Ravinesh Deo, whose influence is relatively bigger. As for countries, United States, China, and Australia are the three top research centers of relative research. When it comes to institutions, the Chinese Academy of Sciences has the most publications; there is no American or Australian institution among the top five. The journal *Agriculture Ecosystems & Environment* has become the most popular journal among all the research materials, followed by *Science, Nature, Journal of*

Hydrology, and *Computers and Electronics in Agriculture.* The six hot topics detected are "learning technique" (#0), "landscape context" (#1), "reference evapotranspiration" (#2), "soil management" (#3), "river GEODSS" (river geographic decision support system) (#4), and "winter wheat" (#5), and it is noticeable that "learning technique" is the most popular direction in recent years. According to the category analysis, agricultural management with machine learning is mainly linked to agriculture, ecology, computer science, and so on.

As can be drawn from the results, it is noticeable that machine learning is widely used in sustainable agricultural development, agricultural remote sensing technology, and water resource management; their application will be even more widespread in the future. Moreover, the fast development of machine learning methods (neural networks, deep learning, random forests, logistic regression, etc.) in recent years will energize the synthetic application in smart and sustainable agriculture. What is more, China and the United States will remain the two most important contributors in this field for a long time. The Chinese Academy of Sciences, the University of Tehran, Wageningen University and Research, and China Agricultural University are the leaders of the relevant research with huge developmental potential. Additionally, the application of machine learning in sustainable agriculture would provide more space for precision agriculture, yield prediction, crop grading, and pest and disease management in the future.

The application of machine learning approaches could directly affect the process of smart agriculture and promote sustainable agriculture development. With the development of machine learning, increasingly more agriculture problems could be better solved. With all the research material considered, the authors provided an integrated mind map of this area and a corresponding overview. In addition, this research could update the latest research trend of this field. Further research can be conducted by extending more databases or integrating more aspects of material analysis.

Author Contributions: Conceptualization, J.Z. and J.L.; Methodology, J.Z. and Y.C.; Software, J.Z.; Validation, X.F. and J.L.; Formal Analysis, J.Z., Z.S. and J.L.; Investigation J.Z.; Data Curation, J.L. and J.Z.; Writing—Original Draft Preparation, J.Z. and J.L.; Writing—Review and Editing, J.Z., Y.C. and X.F.; Supervision, X.F. and Z.S. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable; this research did not involve humans or animals.

Data Availability Statement: The data used in this article all came from the Web of Science.

Conflicts of Interest: The authors declare no conflict of interest.

References

1.　Liakos, K.G.; Busato, P.; Moshou, D.; Pearson, S.; Bochtis, D. Machine Learning in Agriculture: A Review. *Sensors* **2018**, *18*, 2674. [CrossRef] [PubMed]

2.　Ayaz, M.; Ammad-Uddin, M.; Sharif, Z.; Mansour, A.; Aggoune, E.H.M. Internet-of-Things (IoT)-based smart agriculture: Toward making the fields talk. *IEEE Access* **2019**, *7*, 129551–129583. [CrossRef]

3.　Akhter, R.; Sofi, S.A. Precision Agriculture using IoT Data Analytics and Machine Learning. *J. King Saud Univ. Comput. Inf. Sci.* **2021**. [CrossRef]

4.　Priya, R.; Ramesh, D. ML based sustainable precision agriculture: A future generation perspective. *Sustain. Comput. Inform. Syst.* **2020**, *28*, 100439.

5.　Khanal, S.; Klopfenstein, A.; Kushal, K.C.; Ramarao, V.; Fulton, J.; Douridas, N.; Shearer, S.A. Assessing the impact of agricultural field traffic on corn grain yield using remote sensing and machine learning. *Soil Tillage Res.* **2021**, *208*, 104880. [CrossRef]

6. Cunha, R.L.; Silva, B.; Netto, M.A. A scalable machine learning system for pre-season agriculture yield forecast. In Proceedings of the 2018 IEEE 14th International Conference on e-Science (e-Science), Amsterdam, The Netherlands, 29 October–1 November 2018; pp. 423–430.
7. Duarte-Carvajalino, J.M.; Alzate, D.F.; Ramirez, A.A.; Santa-Sepulveda, J.D.; Fajardo-Rojas, A.E.; Soto-Suárez, M. Evaluating late blight severity in potato crops using unmanned aerial vehicles and machine learning algorithms. *Remote Sens.* **2018**, *10*, 1513. [CrossRef]
8. Hu, X.; Sun, L.; Zhou, Y.; Ruan, J. Review of operational management in intelligent agriculture based on the Internet of Things. *Front. Eng. Manag.* **2020**, *7*, 309–322. [CrossRef]
9. Sharma, R.; Kamble, S.S.; Gunasekaran, A.; Kumar, V.; Kumar, A. A Systematic Literature Review on Machine Learning Applications for Sustainable Agriculture Supply Chain Performance. *Comput. Oper. Res.* **2020**, *119*, 104926. [CrossRef]
10. Jordan, M.I.; Mitchell, T.M. Machine learning: Trends, perspectives, and prospects. *Science* **2015**, *349*, 255–260. [CrossRef]
11. Mazzia, V.; Comba, L.; Khaliq, A.; Chiaberge, M.; Gay, P. UAV and Machine Learning Based Refinement of a Satellite-Driven Vegetation Index for Precision Agriculture. *Sensors* **2020**, *20*, 2530. [CrossRef] [PubMed]
12. Wuest, T.; Weimer, D.; Irgens, C.; Thoben, K.-D. Machine learning in manufacturing: Advantages, challenges, and applications. *Prod. Manuf. Res.* **2016**, *4*, 23–45. [CrossRef]
13. Lee, Y.C.; Chen, C.; Tsai, X.T. Visualizing the knowledge domain of nanoparticle drug delivery technologies: A scientometric review. *Appl. Sci.* **2016**, *6*, 11. [CrossRef]
14. Chen, C. CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature. *J. Am. Soc. Inf. Sci. Technol.* **2006**, *57*, 359–377. [CrossRef]
15. Synnestvedt, M.B.; Chen, C.; Holmes, J.H. CiteSpace II: Visualization and knowledge discovery in bibliographic databases. *AMIA Annu. Symp. Proc. AMIA Symp.* **2005**, *2005*, 724–728.
16. Chen, C.; Hu, Z.; Liu, S.; Tseng, H. Emerging trends in regenerative medicine: A scientometric analysis in CiteSpace. *Expert Opin. Biol. Ther.* **2012**, *12*, 593–608. [CrossRef]
17. Chen, Y.; Chen, C.M.; Liu, Z.Y.; Hu, Z.G.; Wang, X.W. The methodology function of Cite Space mapping knowledge domains. *Stud. Sci. Sci.* **2015**, *33*, 242–253.
18. Berner, L.T.; Beck, P.S.A.; Loranty, M.M.; Alexander, H.D.; Mack, M.C.; Goetz, S.J. Cajander larch (Larix cajanderi) biomass distribution, fire regime and post-fire recovery in northeastern Siberia. *Biogeosciences* **2012**, *9*, 3943–3959. [CrossRef]
19. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [CrossRef]
20. Kamilaris, A.; Prenafeta-Boldú, F.X. Deep learning in agriculture: A survey. *Comput. Electron. Agric.* **2018**, *147*, 70–90. [CrossRef]
21. Godfray, H.C.J.; Beddington, J.R.; Crute, I.R.; Haddad, L.; Lawrence, D.; Muir, J.F.; Pretty, J.; Robinson, S.; Thomas, S.M.; Toulmin, C. Food Security: The Challenge of Feeding 9 Billion People. *Science* **2010**, *327*, 812–818. [CrossRef]
22. Tilman, D.; Balzer, C.; Hill, J.; Befort, B.L. Global food demand and the sustainable intensification of agriculture. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 20260–20264. [CrossRef]
23. Ruan, J.; Jiang, H.; Zhu, C.; Hu, X.; Shi, Y.; Liu, T.; Rao, W.; Chan, F.T.S. Agriculture IoT: Emerging Trends, Cooperation Networks, and Outlook. *IEEE Wirel. Commun.* **2019**, *26*, 56–63. [CrossRef]
24. Belgiu, M.; Drăguţ, L. Random forest in remote sensing: A review of applications and future directions. *ISPRS J. Photogramm. Remote. Sens.* **2016**, *114*, 24–31. [CrossRef]
25. Elith, J.; Leathwick, J.R.; Hastie, T. A working guide to boosted regression trees. *J. Anim. Ecol.* **2008**, *77*, 802–813. [CrossRef]
26. Tabari, H.; Kisi, O.; Ezani, A.; Talaee, P.H. SVM, ANFIS, regression and climate based models for reference evapo-transpiration modeling using limited climatic data in a semi-arid highland environment. *J. Hydrol.* **2012**, *444*, 78–89. [CrossRef]
27. Abdullah, S.S.; Malek, M.; Abdullah, N.S.; Kisi, O.; Yap, K.S. Extreme Learning Machines: A new approach for prediction of reference evapotranspiration. *J. Hydrol.* **2015**, *527*, 184–195. [CrossRef]
28. Prasad, R.; Deo, R.C.; Li, Y.; Maraseni, T. Soil moisture forecasting by a hybrid machine learning technique: ELM integrated with ensemble empirical mode decomposition. *Geoderma* **2018**, *330*, 136–161. [CrossRef]
29. Ali, M.; Deo, R.C.; Downs, N.J.; Maraseni, T. Multi-stage hybridized online sequential extreme learning machine integrated with Markov Chain Monte Carlo copula-Bat algorithm for rainfall forecasting. *Atmos. Res.* **2018**, *213*, 450–464. [CrossRef]
30. Tao, H.; Diop, L.; Bodian, A.; Djaman, K.; Ndiaye, P.M.; Yaseen, Z.M. Reference evapotranspiration prediction using hybridized fuzzy model with firefly algorithm: Regional case study in Burkina Faso. *Agric. Water Manag.* **2018**, *208*, 140–151. [CrossRef]
31. Kisi, O.; Alizamir, M. Modelling reference evapotranspiration using a new wavelet conjunction heuristic method: Wavelet extreme learning machine vs wavelet neural networks. *Agric. For. Meteorol.* **2018**, *263*, 41–48. [CrossRef]
32. Sonobe, R.; Tani, H.; Wang, X.; Kobayashi, N.; Shimamura, H. Parameter tuning in the support vector machine and random forest and their performances in cross-and same-year crop classification using TerraSAR-X. *Int. J. Remote Sens.* **2014**, *35*, 7898–7909. [CrossRef]
33. Sonobe, R.; Yamaya, Y.; Tani, H.; Wang, X.; Kobayashi, N.; Mochizuki, K.-I. Assessing the suitability of data from Sentinel-1A and 2A for crop classification. *GISci. Remote Sens.* **2017**, *54*, 918–938. [CrossRef]
34. Biro, K.; Pradhan, B.; Buchroithner, M.; Makeschin, F. Land Use/Land Cover Change Analysis and Its Impact on Soil Properties in the Northern Part of Gadarif Region, Sudan. *Land Degrad. Dev.* **2011**, *24*, 90–102. [CrossRef]
35. Neshat, A.; Pradhan, B.; Pirasteh, S.; Shafri, H.Z.M. Estimating groundwater vulnerability to pollution using a modified DRASTIC model in the Kerman agricultural area, Iran. *Environ. Earth Sci.* **2013**, *71*, 3119–3131. [CrossRef]

36. Pradhan, B.; Chaudhari, A.; Adinarayana, J.; Buchroithner, M.F. Soil erosion assessment and its correlation with landslide events using remote sensing data and GIS: A case study at Penang Island, Malaysia. *Environ. Monit. Assess.* **2011**, *184*, 715–727. [CrossRef]

37. Pradhan, B.; Lee, S. Landslide susceptibility assessment and factor effect analysis: Backpropagation artificial neural networks and their comparison with frequency ratio and bivariate logistic regression modelling. *Environ. Model. Softw.* **2010**, *25*, 747–759. [CrossRef]

38. Cui, N.; Du, T.; Kang, S.; Li, F.; Zhang, J.; Wang, M.; Li, Z. Regulated deficit irrigation improved fruit quality and water use efficiency of pear-jujube trees. *Agric. Water Manag.* **2008**, *95*, 489–497. [CrossRef]

39. Tang, D.; Feng, Y.; Gong, D.; Hao, W.; Cui, N. Evaluation of artificial intelligence models for actual crop evapo-transpiration modeling in mulched and non-mulched maize croplands. *Comput. Electron. Agric.* **2018**, *152*, 375–384. [CrossRef]

40. Bhavsar, H.; Ganatra, A. A comparative study of training algorithms for supervised machine learning. *Int. J. Soft Comput. Eng. (IJSCE)* **2012**, *2*, 2231–2307.