ISSN: 2664-9640

Research On Agricultural Information Collection And Management System Based On Big Data

Kaijie Li , Lifei Sun

Liaoning University of science and Technology, China

Abstract

This design provides agricultural information collection and management system based on big data, which includes the following steps: collecting environmental data and current growth characteristics data of the same kind of crops in multiple regions;The environmental data and the current growth characteristic data are preprocessed;The correlation coefficient between the growth data and the current environment is obtained;Several prediction models are constructed and trained by using environmental data with correlation coefficient greater than the preset value and current growth characteristic data;According to the trained prediction models, the crop yield is estimated.This design can well predict the final yield of crops,Therefore, the final yield of the comprehensive prediction of crops will not have a large error with the actual final yield.Accordingly, this design also provides agricultural information collection and management system based on big data.

Keywords

Big data, agriculture, information, collection system.

1. Introduction

Data information is the basis and basis of decision-making. Agricultural data information is all kinds of material data and energy data involved in agricultural production or agricultural economic activities.Agricultural data information is the basis of agricultural decision-making, including agricultural production data, soil quality data, meteorological and hydrological observation data and other agricultural spatial data information.

For crops in different regions of the same longitude and latitude zone, the different planting time and regional environmental data information will lead to different yields.Through the collection and analysis of agricultural big data information, such as the collection and analysis of crop soil information, environmental information and crop information,The yield of crops can be estimated comprehensively.

In the existing technology, the crop yield can be estimated only by collecting the relevant data and information of crops in the same region to build a single crop yield prediction model. However, due to the limitations of crop related data information in the same region, it can not fully reflect the physiological and biochemical characteristics of crops,There is a big error between the predicted crop yield and the final actual yield, which needs to be improved.

2. Design content

Based on this, in order to solve the problem that the existing technology only collects the relevant data of crops in the same region to build a single crop yield prediction model to estimate the agricultural yield, and the estimated crop yield often has a large error with the final actual yield, this design provides the research of agricultural information collection and management system based on big data,The specific technical scheme is as follows:

The research of agricultural information collection and management system based on big data includes big data acquisition module, preprocessing module, correlation coefficient acquisition module, model construction module and prediction module.

The big data acquisition module is used to collect the environmental data and current growth characteristic data of the same kind of crops in multiple regions, and the preprocessing module is used to preprocess the environmental data and the current growth characteristic data.

The correlation coefficient acquisition module is used to obtain the correlation coefficient between the environmental data and the current growth characteristic data, and the model building module is used to build multiple prediction models, and train multiple prediction models by using the environmental data with correlation coefficient greater than the preset coefficient value and the current growth characteristic data respectively.

The prediction module is used to predict crop yield according to the trained prediction models.

The correlation coefficient represents the degree of correlation between environmental data and current crop growth characteristics data.When the correlation coefficient is greater than the preset value, it indicates that there is a strong correlation between the environmental data and the current growth characteristic data.

The current crop growth model can be used to predict the current crop growth data through the data collection module of multiple strong correlation,Then, the final yield of crops is comprehensively estimated by using the crop yields obtained from multiple forecasts.

To sum up, the agricultural information collection system can comprehensively predict the final yield of crops by constructing multiple prediction models and training multiple prediction models with strong correlation of environmental data and current growth characteristic data, so that the final yield of crops predicted comprehensively will not produce large error with the actual yield,It overcomes the problem that the existing technology only collects the relevant data information of crops in the same region to construct a single crop yield prediction model to estimate the crop yield, and the estimated crop yield often has a large error with the final actual yield.

The correlation coefficient acquisition module includes the first fitting curve acquisition module and the first phase relationship value acquisition module.

The first fitting curve acquisition module is used to obtain the first fitting curve by fitting the spatial environmental parameters of different regions of the same crop with the current growth characteristic data of corresponding crops.

The first phase relationship value acquisition module is used to obtain the first phase relationship value between the first fitting curve corresponding to crops in any region and the first fitting curve corresponding to crops in other regions.

Among them, the correlation coefficient includes the first phase relationship value. The model building module is used to train multiple prediction models by using the spatial environment parameters whose first phase relationship value is greater than the preset coefficient value and the corresponding current growth characteristic data of crops.

The correlation coefficient acquisition module also includes the second fitting curve acquisition module and the second phase relationship value acquisition module.

The second fitting curve acquisition module is used to fit the underground soil parameters of different regions of the same crop with the current growth characteristic data of corresponding crops to obtain the second fitting curve.

The second phase relationship value acquisition module is used to obtain the second phase relationship values between the second fitting curve corresponding to crops in any region and the second fitting curve corresponding to crops in other regions.

Among them, the correlation coefficient also includes the second phase relationship value. The model building module is also used to train multiple prediction models by using the underground soil parameters whose second phase relationship value is greater than the preset coefficient value and the corresponding current growth characteristic data of crops.

The current growth characteristic data is the variation of a certain part of the crop in a preset time period. The spatial environmental parameters include the average value of air temperature and humidity, the average value of light intensity, the average value of oxygen content and the average value of carbon dioxide content.

The underground soil parameters include the average value of soil light sensitivity, the average value of soil temperature and humidity, and the average value of soil pH.

Agricultural information collection and management system based on big data

The agricultural information collection and management system based on big data includes the following steps:

S1, environmental data and current growth characteristic data of the same kind of crops in multiple regions were collected.

S2, the environmental data and the current growth characteristic data were preprocessed.

S4, the correlation coefficient between environmental data and current growth characteristic data is obtained.

S4, construct multiple prediction models and train multiple prediction models by using environmental data with correlation coefficient greater than the preset coefficient value and current growth characteristic data.

S5, according to a number of trained prediction models to predict crop yield.

The specific method to obtain the correlation coefficient between environmental data and current growth characteristic data includes the following steps:

The first kind of crop growth curve was fitted to the same type of spatial characteristics of crop growth;

The first correlation coefficient values between the first fitting curve corresponding to crops in any region and the first fitting curve corresponding to crops in other regions are obtained respectively;

The second fitting curve was obtained by fitting the underground soil parameters of different regions of the same crop with the current growth characteristics data of corresponding crops;

The second correlation coefficient values between the second fitting curve corresponding to crops in any region and the second fitting curve corresponding to crops in other regions are obtained respectively;

The correlation coefficient includes the first phase relation value and the second phase relation value.

The specific method of training the prediction model by using the environmental data with correlation coefficient greater than the preset coefficient value and the current growth characteristic data includes the following steps:

The spatial environmental parameters with the first phase relationship value larger than the preset coefficient value and the corresponding current growth characteristic data of crops were used to train multiple prediction models;

Several prediction models were trained by using the underground soil parameters whose second phase relationship value was greater than the preset coefficient value and the corresponding crop growth characteristic data.

The current growth characteristic data is the variation of a certain plant part in a preset time period. The spatial environmental parameters include the average value of air temperature and humidity, the average value of light intensity, the average value of oxygen content and the

average value of carbon dioxide content. The underground soil parameters include the average value of soil light sensitivity, the average value of soil temperature and humidity, and the average value of soil pH.

The invention relates to a computer-readable storage medium. The computer-readable storage medium stores the computer program. When the computer program is executed, the agricultural information collection method based on big data is realized.

3. Specific implementation mode

The research of agricultural information collection and management system based on big data in an implementation example includes big data acquisition module, preprocessing module, correlation coefficient acquisition module, model building module and prediction module.

The big data acquisition module is used to collect the environmental data and current growth characteristic data of the same kind of crops in multiple regions, and the preprocessing module is used to preprocess the environmental data and the current growth characteristic data.

Specifically, the big data acquisition module can be used for various sensors installed in multiple regions to collect the environmental data and current growth characteristic data of the same kind of crops, such as image sensor, temperature and humidity sensor, light intensity sensor, etc.

The preprocessing methods include filtering and noise reduction to eliminate the invalid data in the environment data and the current growth feature data.

The correlation coefficient acquisition module is used to obtain the correlation coefficient between the environmental data and the current growth characteristic data, and the model building module is used to build multiple prediction models, and train multiple prediction models by using the environmental data with correlation coefficient greater than the preset coefficient value and the current growth characteristic data respectively.

The correlation coefficient represents the degree of correlation between environmental data and current crop growth characteristics data.When the correlation coefficient is greater than the preset value, it indicates that there is a strong correlation between the environmental data and the current growth characteristic data.

But not limited to the average value of air temperature and air temperature in a certain period, but not limited to the average value of air temperature and air temperature in a certain period of time,The underground soil parameters include but are not limited to the average value of soil light sensitivity, the average value of soil temperature and humidity, and the average value of soil pH value.

The above average value can be the average value of parameters in a certain preset time period, such as day, night or continuous n hours.

A plant part includes but is not limited to the roots, stems, branches, flowers and fruits of crops. Correspondingly, the variation can be elongation, shrinkage or change rate.

By analyzing the correlation between the environmental data and the changes of a certain part of the crop in the preset time period, and training the prediction model by using the strong correlation environmental data and the current growth characteristic data, the growth trend of crops can be well predicted, and then the crop yield can be well predicted.

The prediction module is used to predict crop yield according to the trained prediction models.

The current crop growth model can be used to predict the current crop growth data through the data collection module of multiple strong correlation,Then, the final yield of crops is comprehensively estimated by using the crop yields obtained from multiple forecasts.

To sum up, the agricultural information collection system can comprehensively predict the final yield of crops by constructing multiple prediction models and training multiple prediction

models with strong correlation of environmental data and current growth characteristic data, so that the final yield of crops predicted comprehensively will not produce large error with the actual yield,It overcomes the problem that the existing technology only collects the relevant data information of crops in the same region to construct a single crop yield prediction model to estimate the crop yield, and the estimated crop yield often has a large error with the final actual yield.

In one embodiment, the correlation coefficient acquisition module includes the first fitting curve acquisition module, the first phase relationship value acquisition module, the second fitting curve acquisition module and the second phase relationship value acquisition module.

The first fitting curve acquisition module is used to fit the spatial environmental parameters of different regions of the same crop with the current growth characteristic data of corresponding crops to obtain the first fitting curve,The first phase relationship value acquisition module is used to obtain the first phase relationship value between the first fitting curve corresponding to crops in any region and the first fitting curve corresponding to crops in other regions.

The second fitting curve acquisition module is used to fit the underground soil parameters of different regions of the same crop with the current growth characteristic data of corresponding crops to obtain the second fitting curve,The second phase relationship value acquisition module is used to obtain the second phase relationship values between the second fitting curve corresponding to crops in any region and the second fitting curve corresponding to crops in other regions.

Among them, the first fitting curve represents the functional curve between the spatial environmental parameters and the current growth characteristic data in each growth cycle of crops, and the correlation coefficient includes the first phase relationship value and the second phase relationship value,The model building module is used to train multiple prediction models by using the spatial environmental parameters whose first phase relationship value is greater than the preset coefficient value and the corresponding current growth characteristic data of crops.

The model building module is also used to train multiple prediction models by using the underground soil parameters whose second phase relationship value is greater than the preset coefficient value and the corresponding current growth characteristic data of crops.

In the actual process, due to the differences of crop sowing time and spatial environmental parameters in different regions of the same longitude and latitude zone, there are certain differences in the growth trend and yield of crops, and the correlation degree between the first fitting curves of different regions may be quite different,There can be strong correlation, weak correlation or no correlation.

The model building module is used to train multiple prediction models by using the spatial environmental parameters whose first phase relationship value is greater than the preset coefficient value and the corresponding current growth characteristic data of crops, which can make full use of the spatial environment parameters with strong correlation and the current growth characteristic data to train the prediction models,Make full use of the collected spatial environmental parameters of the same kind of crops in multiple regions and the current growth characteristic data, and give full play to the advantages of big data, so as to better predict the crop yield through the prediction model.

The model building module is used to train multiple prediction models by using the current growth characteristic data of the empty underground soil and corresponding crops whose second phase relationship value is greater than the preset coefficient value. It can make full use of the underground soil parameters with strong correlation and the current growth characteristic data to train the prediction models,Make full use of the collected underground soil parameters and current growth characteristic data of the same kind of crops in multiple

regions, and give full play to the advantages of big data, so as to better predict the crop yield through the prediction model.

Through the construction of multiple prediction models, the differences of crop sowing time and spatial environmental parameters in different regions of the same longitude and latitude zone are fully considered, so that the multiple prediction models can be targeted to predict the crops in different regions,To avoid the problem that single prediction model can not meet the demand of crop yield prediction in all regions, and the problem that crop yield in different regions can be predicted by a single prediction model is prone to have a large error with the actual final yield.

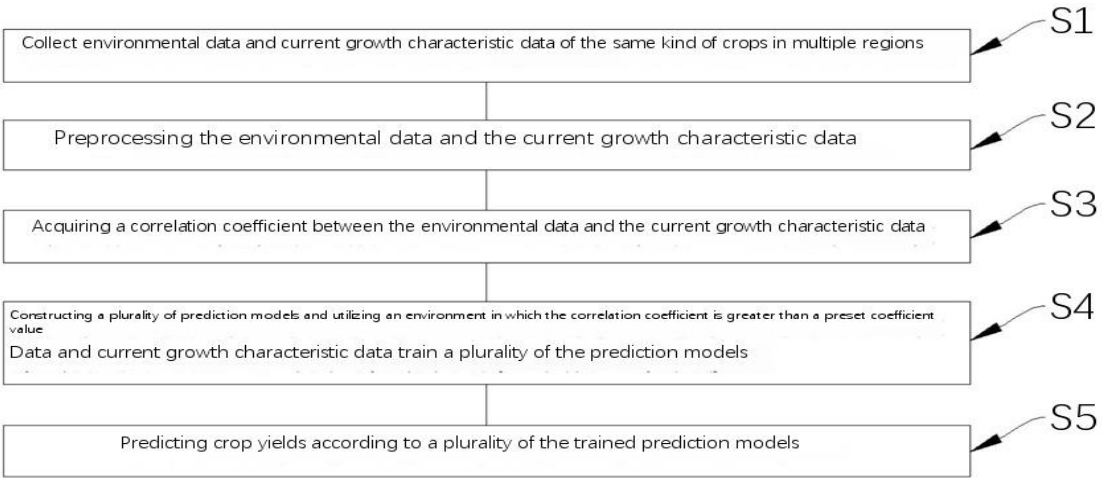

Figure 1 overall flow diagram of agricultural information collection and management system based on big data

In one embodiment, as shown in Figure 1, the agricultural information collection and management system based on big data includes the following steps:

S1, environmental data and current growth characteristic data of the same kind of crops in multiple regions were collected.

S2, the environmental data and the current growth characteristic data were preprocessed.

S3, the correlation coefficient between environmental data and current growth characteristic data is obtained.

S4, construct multiple prediction models and train multiple prediction models by using environmental data with correlation coefficient greater than the preset coefficient value and current growth characteristic data.

S5, according to a number of trained prediction models to predict crop yield.

For example, the data acquisition of different types of data in the same area can be achieved by installing various types of sensors in the same area.

The environmental data include spatial environmental parameters and underground soil parameters. The current growth characteristic data is the variation of a certain plant part in a preset time period. The spatial environmental parameters include the average air temperature and humidity, the average light intensity, the average oxygen content and the average carbon dioxide content,The underground soil parameters include the average value of soil light sensitivity, the average value of soil temperature and humidity, and the average value of soil pH.

The multiple prediction models can be one-to-one corresponding to each region, and at least two prediction models can be set in each area, so that the prediction models can be trained by using the spatial environment parameters, underground soil parameters and current growth characteristics data.Among them, the prediction model includes but is not limited to neural network.

Since neural network is a conventional technical means, it will not be repeated here.

The correlation coefficient represents the degree of correlation between environmental data and current crop growth characteristics data.When the correlation coefficient is greater than the preset value, it indicates that there is a strong correlation between the environmental data and the current growth characteristic data.

By analyzing the correlation between the environmental data and the changes of a certain part of the crop in the preset time period, and training the prediction model by using the strong correlation environmental data and the current growth characteristic data, the growth trend of crops can be well predicted, and then the crop yield can be well predicted.

By constructing multiple prediction models and training multiple prediction models with strong correlation of environmental data and current growth characteristic data, the final yield of crops can be comprehensively estimated, so that the final yield of crops can not produce large error with the actual final yield,It overcomes the problem that the existing technology only collects the relevant data information of crops in the same region to construct a single crop yield prediction model to estimate the crop yield, and the estimated crop yield often has a large error with the final actual yield.

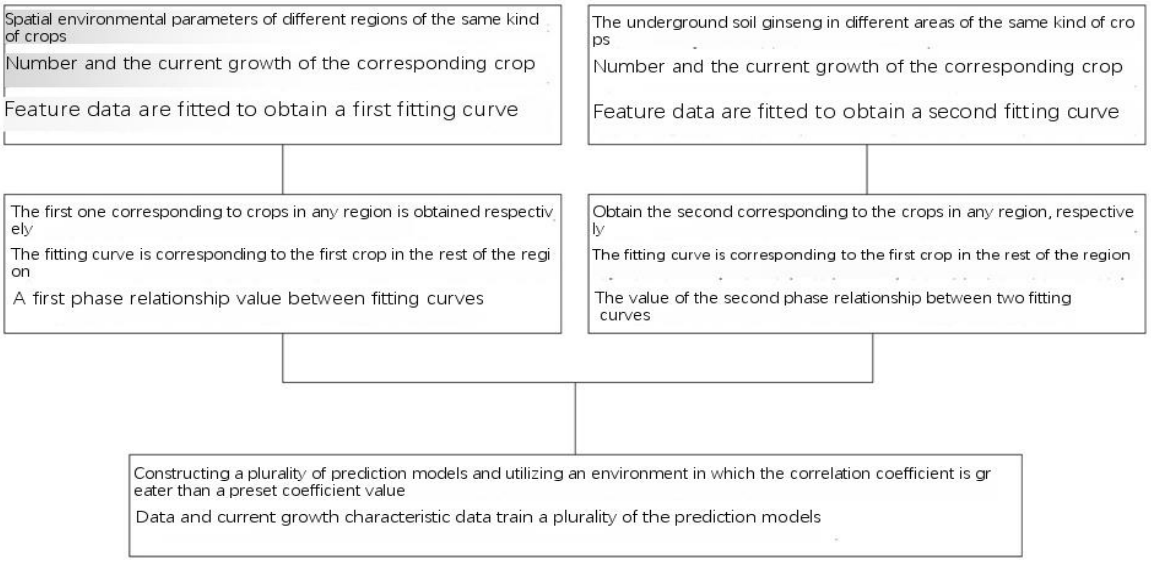

Fig. 2 flow chart of specific method for obtaining correlation coefficient between environmental data and current growth characteristic data

In one embodiment, as shown in Fig. 2, the specific method for obtaining the correlation coefficient between the environmental data and the current growth characteristic data includes the following steps:

The first fitting curve is obtained by fitting the spatial environmental parameters of different regions of the same crop with the current growth characteristics of the corresponding crops.

S31, the first correlation coefficient values between the first fitting curve corresponding to crops in any region and the first fitting curve corresponding to crops in other regions are obtained respectively.

S32, the second fitting curve was obtained by fitting the underground soil parameters of different regions of the same kind of crops with the current growth characteristic data of corresponding crops.

S33, the second correlation coefficient values between the second fitting curve corresponding to crops in any region and the second fitting curve corresponding to crops in other regions are obtained respectively.

The correlation coefficient includes the first phase relation value and the second phase relation value.

The spatial environmental parameters with the first phase relationship value greater than the preset coefficient value and the corresponding current growth characteristic data of crops were used to train multiple prediction models, and the empty underground soil with the second phase relationship value greater than the preset coefficient value and the current growth characteristic data of corresponding crops were used to train multiple prediction models,We can make full use of the spatial environment parameters with strong correlation, the current growth characteristic data, the underground soil parameters and the current growth characteristic data to train the prediction model, and make full use of the spatial environment parameters, underground soil parameters and current growth characteristic data of the same kind of crops in multiple regions, and give full play to the advantages of big data,In order to better predict the crop yield through the prediction model.

In one embodiment, the specific method for training the prediction model using environmental data with correlation coefficient greater than the preset coefficient value and current growth characteristic data includes the following steps:

S40, using the spatial environmental parameters whose first phase relationship value is greater than the preset coefficient value and the corresponding crop current growth characteristic data to train multiple prediction models.

S41, the underground soil parameters whose second phase relationship value is greater than the preset coefficient value and the corresponding crop current growth characteristic data are used to train multiple prediction models.

In one embodiment, a computer-readable storage medium, which stores computer programs, is a method of agricultural information collection based on big data when the computer program is executed.

References

[1] Design of intelligent fertilization system based on data mining technology [J]Liu Jingcheng Agricultural Mechanization Research2022(09).
[2] Design of agricultural intelligent management platform based on data mining [J]Bi ran, Liu Xinsu, Zhang JianxiAgricultural Mechanization Research2022(09).
[3] Research and application progress of agricultural information processing technology [J]Ou Fanfan, Zhang ChaoqunBulletin of Chinese agronomy2021(20).
[4] Application analysis of data mining in agricultural production [J]Peng ZhihuaRural economy and science and technology2021(03).
[5] Problems and solutions in developing electronic archives of seed production and management under the background of big data [J]Li Xinyi, Zhang Yi, Wu Jiayu, Weng YijieSeed technology2021(01).
[6] Data collection of fermentation process of edible fungus culture material based on data mining [J]Sun YahongChinese edible fungi2020(08).