All in one: farmer assistant system using machine learning algorithm.

Recommendation System using K-nearest neighbor

[1]Shraddha DattatrayGangurde,[2]Snehal DilipPagare,[3]Prof. Pankaj Desai

[1]Student of BEIT,[2]Student of BEIT,[3]Assistant professor
[1]Information Technology,
[1]SNJB's KBJ Collage of Engineering, Nasik,India

Abstract

Being an agrarian economy country, India is handling agricultural activities since ages. Agricultural system is cyclic in nature. The strong Climatic changes due to many reasons like global warming cause difficulty to understand climatic conditions. So the farmers unable to understand which crop to select by which the production will improve. The common problem existing among the Indian farmers are they don't choose the right crop based on statistical information about crop. Due to this they face a serious setback in productivity. This problem of the farmers has been addressed through **Farmer Assistant System (FAS)**. FAS is a modern farming technique that uses research data of crop statistics, fertilizers, weather information, crop yield data collection and suggests the farmers the right crop based on their site-specific parameters. This reduces the wrong choice on a crop and increase in productivity. In this system, problem is solved by proposing a crop as well as fertilizer recommendation to farmers. The main features of the FAS includes information retrieval facilities for users from anywhere in the form of obtaining statistical information about crop. Another feature of FAS includes representation of crop data by using Google map for a particular region for better understanding of users.

***IndexTerms*: -**Agriculture, fertilizer recommendation, Google map.

1. Introduction

Agriculture is a prime occupation in India from ages and thus plays a vital role inan Indian economy. India possesses a tremendous potential to be a superpower in the field of agriculture. Agriculture promotes poverty upliftment and rural development. Today in India agriculture is being neglected which has led to losing hope of farmers in agriculture which has led to rise in the number of farmer suicides. There is no such universal system to assist farmers in agriculture. In India agricultural is carried out from ages and thus we have a rich collection of agricultural past data which can used for recommendation. Machine Learning techniques and algorithms can be used for recommending single crop and pattern of crops for crop rotation. However to obtain optimized and valid results system needs to be in continuous learning which can be done by including latest data sets in the system. So to assist and support the farmers we have proposed a System which will provide information about the crop grown by other farmers in their farm by using graphical representation and also suggest them which crop will be beneficial and suitable for the atmospheric environment of the farm and as per the weather the sensors will calculate the values of temperature, Humidity and moisture and provides those values to System.

The motivation behind this system is that the system will be one of the many online resources that new farmers can use to get a practical sense of what type of crops they can grow in their Farm area. Most of the Indian population is involved in agriculture hence the economy is largely dependent on agriculture.

The aim therefore is to develop a system which will be a complete guide for farmers. It is the idea to create the platform for all the farmers where farmers can post about their way of developing the crop which will be beneficial for other farmers. The system will also allow some expertise to express their knowledge about the crop. In India, atmospheric conditions are such that the multiple farmers grow the same crop at a certain point without having knowledge of crops grown by other farmers which will results in high production of crops which gives fewer prices to the farmers in the market.

The proposed system will provide the categorized information about crops belonging to the certain area. Farmers have no knowledge about local weather conditions as the previous systems provide the average of weather values such as humidity, temperature and moisture. Hence the system will provide local weather conditions to the farmers.

2. Literature survey

The requirements and planning needed for developing a software model for precision farming is the topic for discussion. It deeply studies the basics of precision farming. The author's start from the basics of precision farming and move towards developing a model that would support it. This paper describes a model that applies Precision Agriculture (PA) principles to

small, open farms at the individual farmer and crop level, to affect a degree of control over variability. The comprehensive objective of the model is to deliver direct advisory services to even the smallest farmer at the level of his/her smallest plot of crop, using the most accessible technologies such as SMS and email. This model has been designed for the scenario in Kerala State where the average holding size is much lower than most of India. Hence this model can be positioned elsewhere in India only with some modifications [1].

The importance of crop selection and the factors deciding the crop selection like production quantity, market price and government policies are discussed. This paper proposes a Crop Selection Method (CSM) which solves the crop selection problem and improves net yield rate of the crop. It suggests a series of crop to be selected over a season considering factors like weather, soil type, water density, crop type. The predicted value of influential parameters determines the accuracy of CSM. Hence there is a need to include a prediction method with improved accuracy and performance [2].

Some systems aim to solve the crucial problem of selecting the classifiers for the ensemble learning. A method to select a best classifier set from a pool of classifiers has been proposed. The proposal aims to achieve higher accuracy and performance [3].

Some systems propose various classification methods to classify the liver disease data set. The paper emphasizes the need for accuracy because it depends on the dataset and the learning algorithm. Classification algorithms such as Naïve Bayes, ANN, ZeroR and VFI were used to classify these diseases and compare the effectiveness, correction rate among them. The performance of the models where compared with accuracy and computational time. It was concluded that all the classifiers except naive Bayes showed improved predictive performance. Multilayer perceptron show the highest accuracy among the proposed algorithms [4].

In last few years many researches are going on the accurate recommendation of crops because this work helps the farmers to increase their productivity. This crop as well as fertilizer recommendation algorithm also helps to identify actual need of the market. Farmer assistant system aims to identify a small group of farmers which are having same crops in their farms.
Based on Many existing system have been studied which provides figures of environmental factors such as temperature, humidity, soil moisture but those figures were average of whole district not from the specific geographical region of user.

3. Data and sources of data

Data required for research consist of data of farmers as well as experts. Data about farmers includes meteorological data, soil data, crop data and farmer's data. Meteorological data is made up of daily maximum temperature, daily minimum temperature, daily average temperature provided by IOT sensors placed into the farms. Soil data consist of soil moisture level which gives water level of soil. Soil data can also be acquired from country soil and fertilization station. Crop data consist of crop species, crop varieties, planting region, planting area, yield per hectare and the annual production. Farmer's data consist of geographical location of that farmer, total land in hectare and identification information of the farmers such as Adhar card.

Expert's data includes educational qualification of that person, personal identification information, etc. All the personal information of farmers and experts are acquired at the time of sign up to the system.

4. System Architecture

The architecture shows the combine working of all the modules. Firstly the system will be used by users. Users are of two types, farmers and experts. Both the users are supposed to provide the inputs to the system. Experts will first sign up to the system. While signing up all the experts are supposed to provide the evidence of their expertise. System will ask them to provide their educational qualification to cross check their degrees. Whenever farmers need any help they will check the expert advices given by experts. Second type of user is farmer. Farmers are supposed to provide all information regarding their farms such as area in acres, number of crops they can grow, etc.

In system architecture, the outgoing arrow from farmer to system shows the flow of data in same direction. Once all the data gets collected from both the users, system will transfer this data to the cloud. In system architecture, bidirectional arrow in between system and cloud shows the data flow in both directions. System will send the data towards cloud storage where the data gets stored on cloud and different operations on data takes place. On cloud storage, machine learning algorithm is used to draw the statistics of data. Linear classification algorithm is used for the same purpose.Once the data is prepared for representation, this data is displayed to the user i.e.farmer. Farmers can also check the environmental measurements with the help of the sensors which are placed in the farm of farmers.

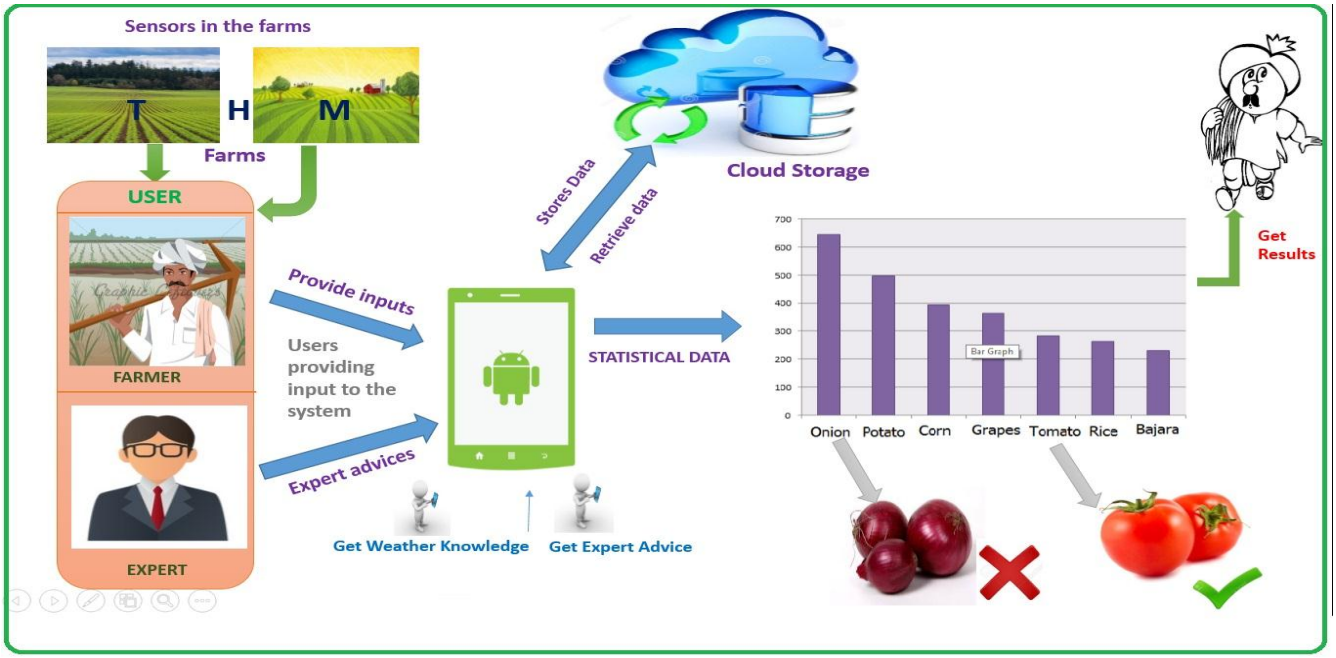

Figure 1: System architecture of farmer assistance system

4.1 Objectives

1) To provide the knowledge about type of crop along with its amount and process.
2) To avoid or restrict wrong figures of weather parameters such as temperature, humidity and moisture.
3) To provide the graphical representation of information with respect to the specific area of region.
4) To get the advices regarding to the crops from experts this will help all the farmers to get the exact solutions of their problems.
5) To provide the crop as well as fertilizer recommendations to farmers depending on their environmental factors and the type of a crop respectively.

4.2 System features

1) Collecting data from users
The system will collect data from users. There are two types of users Farmer and experts. Each user gives different input to the system. System will represent that data statistically to get the knowledge about quantity of crops in theparticular region.

2) Weather forecasting
System contains different types of Sensors such as Temperature, Humidity andMoisture level detection sensor to get the knowledge about environmental parameters of particular farm area by using these sensors

3) Crop recommendation using KNN algorithm
System uses K-nearest algorithm for the purpose of crop as well as fertilizer recommendation to the farmers. It makes use of collaborative recommendation system.

4) Fertilizer recommendation
As system is going to recommend the crops to the farmer, depending on thesecrops system can also recommend fertilizers to the farmers.

5) Expert advise
As System is including experts as users, it can allow experts to post their advicesregarding to the questions of farmers so that farmers can take benefits of expert advises.

6) Detailed process of farming

System can allows the farmers to post their process of farming of a particular crop.If that process is helping farmers to gain more profit then others can refer it.

5. Methodology

In this section, we provide a detailed description about proposed approaches to **Farmer Assistant System**. Given a set of training data *S* which consists of *l* normal examples and a small amount of *n*abnormal examples, the objective is to build a classifier using both normal and abnormal training data and the classifier is there after applied to classify unseen test data. However, normal example may behave like abnormal ones, even though the example itself may not be an abnormal. Such error factors might result in an imperfectly labeled training data, which makes the subsequent recommendation system grossly inaccurate. To deal with this problem, we put forward two likelihood models as follows.

Single likelihood model: In the model, we associate each input data with a likelihood value *(*x*i, m(*x*i))*, which indicates degree of membership of an example towards its own class label.
Bi-likelihood model: In the model, each sample is associate with bi-likelihood values, denoted as *(*x*i,mt(*x*i), mn(*x*i))*, in which *mt(*x*i)* and *mn(*x*i)* indicate the degree of an input data x*i* belonging to the positive class and negative class respectively.

5.1 Classification

This refers to a form of data analysis that can be used to produce models that brings out important data classes, classification predict categorical labels. A classifier is an abstract model that describes a set of predefined classes generated from a collection of labeled .Stated that there are different techniques for data classification which includes; decision tree classifier, Bayesian classifier, K-Nearest Neighbor classifier, rule base classifier etc. In our work, the K-Nearest Neighbor classification method was adopted.

5.2 K-nearest neighbor

K-NN classification is the simplest algorithm used to classify the users in particular category. So, recommendation system can provide the browsing option to user according to his interest. K-NN algorithm takes users previous logs as an input and find out which data is accessed mostly and from that data it predicts the user's interest. So, from that information it classifies the users in category as per his interest and needs.

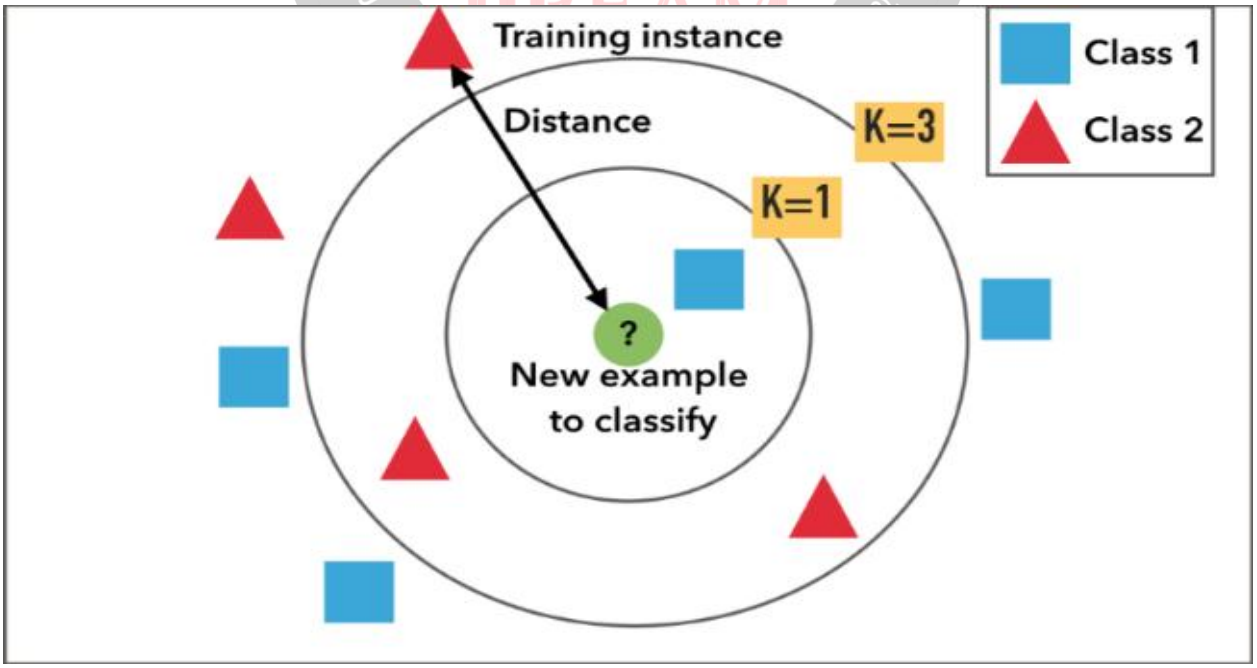

Figure 2: K-nearest neighbor algorithm

5.2.1 Algorithms steps

The algorithm (as described in [5] and [6]) can be summarized as:

1. A positive integer k is specified, along with a new sample
2. We select the k entries in our database which are closest to the new sample
3. We find the most common classification of these entries
4. This is the classification we give to the new sample

5.3 Algorithm for proposed system:

1. All the data farmers have given as input (geographical location information, crop information, automatically fetch information via sensors like temperature, humidity, etc.) at the time of signing in to the system will be considered as training data set D.
2. The K-nearest-neighbour (KNN) algorithm measures the distance between a query farmer q and a set of farmers in the data set D.
   We can compute the distance between query farmer and a set of farmers in the data set using some distance function d(x, y), where x, y are farmers composed of N features, such that $x = \{x_1, x_2, \ldots\ldots, x_N\}$, $y = \{y_1, y_2, \ldots\ldots, y_N\}$.
3. Now that we have established a measure in which to determine the distance between two farmers, we can simply pass through the data set, one farmer at a time and compare it to the query farmer.This is the classification we give to the new sample.
4. After comparison, we will select first K farmers from dataset which are nearest to the query farmer.
5. Store the output values of the first K nearest farmers to the query farmer q in vector V.
6. Analyze the vector V and find the maximum probability farmers having same information (temperature, humidity, geographical location, etc.). Now we can recommend the crops as well as fertilizers used by highest probability farmers to the query farmer.

6. Conclusion

The drawbacks of existing system are lack in improvement of single crop farmingSystem; it does not consider the mutual understanding between farmers and crop ratio, which results in suicide of farmers, decrease in quality, quantity and income of farmers. There is no single automated system for farmers which will solve all the problems in one platform. Also it was impossible for farmers to know the actual environmental factors.

The proposed system overcomes the above stated drawbacks. The system will provide the statistical analysis of crops. System is a stepping stone towards more automated resolution of farmer assistant system which will solve many problems of farmers.

7. References

[1] SatishBabu , ‘A Software Model for Precision Agriculture for Small and Marginal Farmers’, at the International Centre for Free and Open Source Software (ICFOSS) Trivandrum, India,2013.

[2] Rakesh Kumar, M.P. Singh, Prabhat Kumar and J.P. Singh , ’Crop Selection Method to Maximize Crop Yield Rate using Machine Learning Technique’, International Conference on Smart Technologies and Management for Computing, Communication, Controls, Energy and Materials (ICSTM),2015.

[3] LiyingYang , ‘Classifiers selection for ensemble learning based on accuracy and diversity’ Published by Elsevier Ltd. Selection and/or peer-review under responsibility of [CEIS],2011.

[4] A.T.M ShakilAhamed, NavidTanzeemMahmood, NazmulHossain, Mohammad TanzirKabir, Kallal Das, FaridurRahman, Rashedur M Rahman , ‘Applying Data Mining Techniques to Predict Annual Yield of Major Crops and Recommend Planting Different Crops in Different Districts in Bangladesh’ , (SNPD) IEEE/ACIS International Conference,2015.

[5] E. Mirkes, KNN and Potential Energy (Applet).University of Leicester. Available: http://www.math.le.ac.uk/people/ag153/homepage/KNN/KNN3.html, 2011.

[6] L. Kozma, k NearestNeighbours Algorithm. Helsinki University of Technology, Available:http://www.lkozma.net/knn2.pdf, 2008.

