IoT based agricultural cultivation yield system using Machine Learning

Archit Nanavare[1], Bhushan Tiware[2], Radhika Mundhada[3], Sanskar More[4], Dhananjay Gaikwad[5]

[1245]UG Scholar, [3]Assistant Professor

[12345]Department of Computer Engineering

[12345]Alamuri Ratnamala Institute of Engineering and Technology, Shahapur, India

Abstract - The majority of India's agricultural products have been severely impacted by the effects of global warming. when compared to their production during the previous 20 years. Policymakers and farmers will be able to estimate crop yields early in the harvest by using efficient marketing and storage measures. Farmers will be able to make the appropriate decisions thanks to this technology because it will enable them to measure the yield of their crops prior to cultivation. An easy-to-use graphic software can be used to implement this strategy, after which the machine learning algorithm can be deployed.

An easy-to-use graphic software can be used to implement this strategy, after which the machine learning algorithm can be deployed. The farmer is allowed access to the results. However, there are a number of protocols or methods for using data analytics to predict crop yields, and with the help of all those algorithms, we can forecast agricultural productivity. For the working, it makes use of IoT and machine learning. There are no suitable ideas or technologies to deal with the circumstances we face, despite research into challenges and topics including weather, soil quality, etc. In nations like India, there are numerous different sorts of rising economic growth, including in the agriculture sector. The procedure is very helpful for predicting production.

Keywords: Agricultural, IoT, Machine Learning, Crops Yield

I. INTRODUCTION

The research involves making crop recommendations to farmers depending on the soil, weather, and water levels. The Indian economy is heavily reliant on the agricultural sector. Households in rural areas depend on agriculture to a greater than 70% extent. Given that it accounts for 17% of the country's GDP overall and employs more than 60% of the workforce, agriculture is a significant sector of the Indian economy. Over the past few decades, Indian agriculture has experienced significant expansion. We have also put in place a system of tracking for the same reason, which will help us keep an eye on the crops' quality and benefit the farmers' individual output yields. Traditional datasets and machine learning algorithms were utilized to get the data. Various machine learning methods are used to improve the production yield.

Usually, a farmer's intuition and other immaterial factors, like the desire to make quick money, ignorance of market demand, exaggeration of a soil's suitability for a particular crop, and so on, impact his choice of what crop to produce. It is vital to develop a system that might provide Indian farmers with predicted information so they could make informed crop decisions. This necessitates the usage of IoT for smart farming, which is necessary. The use of IoT in agriculture may transform humankind and the entire planet. Analytics of sensor data promotes openness in the agricultural industry by providing farmers with invaluable information about the effectiveness of their fields, greenhouses as well, etc. Using its highly accurate algorithms, machine learning-powered farming is a novel idea that is currently being developed. This cutting-edge initiative encourages sustainable improvement in productivity for everyone involved in agriculture, with the goal of increasing product quantity and quality.

We have created a user interface that is straightforward and simple to grasp and will aid farmers in increasing their output. Since we added the capability to monitor crops, the system is dynamic and will operate in real time as well. A section for government-provided beneficiaries and programmers created for farmers that will aid them in their yield and output financially is also included in the section on production yield. The machine uses both unsupervised and supervised machine learning methods to produce excellent results that are wholly accuracy-based. As a result, the machine will help to lessen the issues that arise when using farmers and stop them from attempting suicide. It will serve as a conduit for delivering the farmers' green records needed to encourage high output and, as a result, maximise income, which will subsequently lower the rate of suicide and lessen his issues.

II. LITERATURE SURVEY

Crop yields are likely to decrease, with the largest decreases anticipated in numerous developing economies, including Southeast Asia (-5%) and India (-5%), according to the Climate Change Agricultural Yield Assumptions report. The range of factors, such as infrastructural and marketing difficulties, improper harvest timing, unexpectedly harsh climatic conditions, and the inability to foresee acceptable crops for agriculture in such settings, may help to explain at least some of the difference in on-farm losses between locations. Below is a comparison that shows:

Rushika Ghadge makes use of supervised and unsupervised learning algorithms like the Back Propagation Network and the Kohonen Self Organizing Map. Learning networks are trained to categories the dataset into organic, inorganic, and real estate in order to forecast the kind of soil. The most accurate outcome is given to the user after it analyzes the accuracy attained using various network learning approaches. The system will evaluate the soil quality, forecast the crop yield in accordance, and make fertiliser recommendations as necessary based on the soil quality.

The Modified Support Vector Regression technique, a well-known machine learning algorithm, and four modules are used in the Reference Paper to identify real-time sampling of soil parameters. The modules comprise the Agri user interface (AUI), the Agri cloud, the sensor data analysis in real time, and the sensor connected to IoT device. The first module is a mobile Internet of Things gadget (Node MCU) with environmental sensors like pH and soil moisture sensors. Storage is included in the agri cloud module. Processing of various crops and tiny plants recommended using a modified support vector machines method is part of the real-time data analysis module. A basic web interface is the agri-user interface. Thus, with the aid of soil attributes and the Modified Support Vector Machine algorithm, farmers will be able to determine the kinds of crops and small plants that can be cultivated on farms.

Using the ARIMA model, it forecasts temperature, moisture, and pH for crop prediction. The model estimates the probable value of that specific parameter one month from now using the values from the database as input. The K Means technique is then used to classify the projected values based on pH value, resulting in k clusters of plants with related pH values. Top N eligible crops are predicted using the KNN algorithm and displayed to the user.

Based on the values obtained in real time, the Machine Learning Algorithm (KNN) determines the parameter to advise the crop that is optimal to grow in the specific field. For the purpose of crop prediction, a standard dataset that contains the minimal needs for a specific crop is kept up to date. The field where the readings are required for calculation receives the sensors. The readings are transmitted in real time to the cloud server using the DHT11, MQ2, Soil Moisture Sensor, and Light Intensity Sensor.

Uses the training set of data and machine learning's supervised and reinforcement learning models to assess the quality of the crop factor based on previously determined weather conditions and soil characteristics. If any unfavorable conditions are anticipated in advance, alternative and cautious steps are taken to protect the crops being grown and agricultural land. For the purpose of improving production overall, which will also be expected as part of the new farming revolution, specific procedures are being performed to forecast the ideal time for sowing, reaping, and harvesting.

Title	Authors	Advantages	Disadvantages	Result
Demand Based Crop Recommender System For Farmers	S. Kanaga Suba Raja, R. Rishi, E. Sundaresan, and V. Srijith	It increases the crop rate by 3%.	Additional sensors can be employed to obtain a comprehensive overview of the health of the crops, produce, and soil conditions. As batteries are required for the sensors in every node, battery energy usage may be higher.	Farmers will profit from using this method and see an increase in crop yield.
RSF: A recommendation system for farmers	M. J. Mokarrama and M. S. Arefin	It elevates the output rate by 7%.	The combination of IoT and an expert system is a great technique to increase crop rate.	In this research, recommendations for crop actions are produced at the upazil level using the algorithm used and data sets gathered.
Crop recommendation system for precision agriculture	S. Pudumalar, E. Ramanujam, R. H. Rajashree, C. Kavya, T. Kiruthika, and J. Nisha,	The combination of IoT and an expert system is a great technique to increase crop rate.	The panel of experts does provide recommendations for pesticides and herbicides based on the disease, but their paper also makes use of a digital camera or other image-capture tool. Additionally, battery usage may be high.	This method helped many people grow the proper crop at an appropriate time and was more profitable as well.
IoT based Weather and Location Aware Recommender System	C. Shampa and M. Aseem	It improves the precision of weather data.	Since Li-Fi is highly sophisticated, it takes time to set it up in developing nations, but once it is, it can be a godsend for the industry, and that idea can be put into practice. Additionally, using cattle sensors is dangerous since the animals could fumble with them and harm the sensors. Since each animal requires a different sensor, it is barely financially feasible.	It demonstrates how the suggested system is considerably improved by the meteorological background.
Affordable field environmental monitoring and plant growth measurement system for smart agriculture	T. Okayasu et al	It improves the accuracy of IOT networks.	The biggest drawback is that playhouses cannot be utilised on broad fields because doing so raises the expense. When we take into account a big area, the cost of installing drip irrigation could also go substantially.	Both precision and coverage are increased by including trust into the RS.

Research on optimizing recommend system for agriculture information personalization based on user clustering	Y. Zhao and S. H. Bai	It increases safety.	Security issues can be encountered.	It tries to offer defence against IOT hackers to address security issues.
A cross-language personalized recommendation model in digital libraries,	Y. Lai and J. Zeng,	It improved the service quality of the website	The transmission distance is short, and if connectivity is poor, it may lead to issues.	The framework significantly improves the mechanism for the customisation of agricultural information.
Design of fertilization recommendation knowledge base and application	Z. Ren and X. Lu	It enhances precession.	Failures in the ubimote sensor board, which is used to collect and transmit data to the farmers, might prevent the system as a whole from functioning.	The following theory was supported by similar users and service resource value.

III. PROPOSED SYSTEM

The interface asks the user to log in before processing the precise spot of the field along with the kind of soil that is suitable for farming. The history data of a certain area, or field, is collected using location as an input. Utilizing public websites or non-government applications such APIs for meteorological and temperature data as well as local rainfall amounts, historical data is gathered. Putting the IoT gadget on the field allows for the collection of real-time data. The DHT 22 humidity and temperature sensor is coupled with an ESP8266 Wi-Fi module and is connected to an Arduino UNO. The real-time data is gathered hourly and saved on the Thing Speak Cloud platform. Both current and past data are gathered. In order to forecast the rainfall, temperature, and humidity for the time when the farmer is scheduled to cultivate the crop, the VAR (Vector auto regression) models is used to the obtained data.

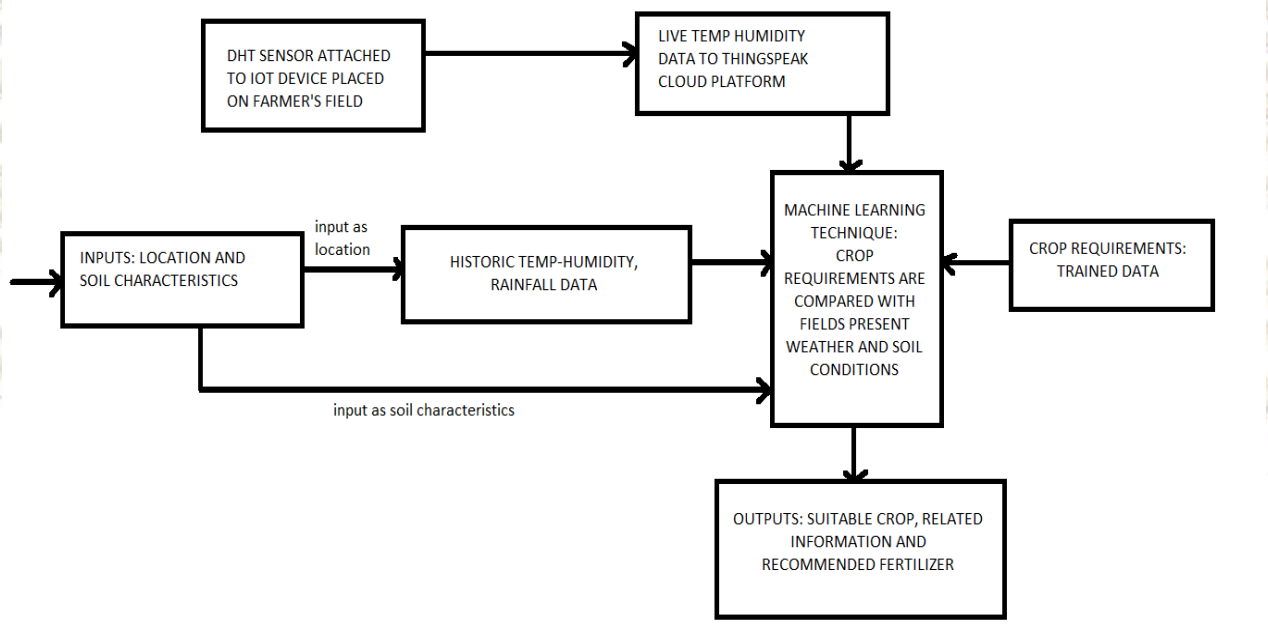

Fig.1 System Architecture Flow

Now, three distinct machine learning (ML) algorithms—Decision Tree, KNN, and Support Vector Machine—are given these projected weather variables—temperature, humidity, and precipitation—along with the soil characteristic input by the farmer. The resultant combination and the predetermined data set—i.e., the actual crop requirements kept in the crop data store—are then contrasted. After comparing the level of accuracy attained by several machine learning algorithms, the user is finally presented with the most accurate outcome, or the most suitable crop.

Temperature	Humidity	pH	Rainfall	Crop
45.6974671	11.19524988	7.341612619	36.522087	5
31.76167796	57.2743884	6.198411713	82.10434980	2
2917510907	79.21985892	6.658683087	131.0207088	14
24.30736241	79.26861738	7.014083944	164.2897011	8
23.17124881	52.a7841162	6.760184468	153.1201644	9
23.89271875	61.7a779413	6.658605362	52.55730112	7
29.70143197	95.68754365	6.078807239	215.1968037	16
24.388717	62.50453082	6.711341147	47.28052494	7
24.32119187	55.84027641	4.958920312	202.281286	4
20.27514608	23.2353804	5.877347515	139.7521843	15

Fig.1.2 Training Dataset

Farmers who use the website receive the best crop as a result. Additionally, the consumer is given complete details on the plant and the ideal fertilizer. Decision Tree is used to predict the crop since it has the greatest degree of precision. The farmer enters the soil pH, location, and anticipated month of planting on the "Smart Crop Prediction" website. Decision Tree is used to predict the crop since it has the highest accuracy.

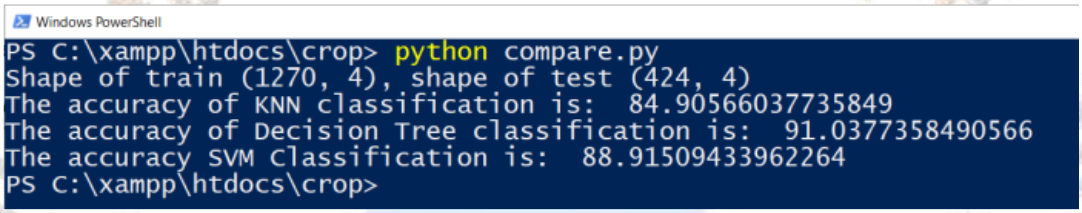

Fig.1.3 Comparison of the accuracy of the three algorithms

Agriculture is the main driver of economic growth in a previously developed nation. As a nation's population grows, so does its reliance on agriculture, which in turn affects the nation's following economic growth. In this instance, the rate of yields of crops significantly influences the economic growth of the nation. Therefore, crop yield rates must be increased. Some biological measures (such as crop quality of seeds, crop hybridization, and powerful insecticides) and a few chemical-based ones (such as fertilizer, urea, and potash use) are utilised to address this issue. To improve the web production rate of the crop throughout the season, a crop sequencing technology is required alongside to those options. For the seasonal realisation of a net crop yield rate, one of the present systems we have described is the Crop Selection Method (CSM). To demonstrate how it enables farmers to produce a higher yield, we used the CSM as an example.

Hardware: Digital Temperature and Humidity Sensor: The DHT22 sensor is recommended for real-time temperature and humidity monitoring. This sensor has been shown to be more accurate and precise. It measures the humidity in the air using a thermistor and a capacitive humidity sensor, and it outputs a digital signal to an Arduino Uno port pin on the data pin. The DHT22 has a temperature range of -40 to 80 degrees Celsius and a humidity range of 0 to 100% RH.

It is important to conduct a thorough risk analysis of all threats that can reasonably threaten the relationship regardless of the type of risk the objectives of business recognition potential perils that are in a position to arise either within or outside of the partnership should be assessed despite the fact the exact setup of anticipated disasters or there after effects district unit postponed to outline and after a breakdown, the overall likelihood of a failure occurring should be determined. Things to look at in a hurry include the probability of a specific breakdown, which should be limited to represent but not be limited to field characteristic examination of the world's closeness to vital power sources and airports, the level of workplace accessibility, and the track record of local service organisations in providing ongoing types of he Potential gaps may even be assigned to common, specific, or human threats. Models consist of:

Characteristic Threats: inner flooding, outer flooding, interior hearth, outside chimney, seismic movement, high breezes, snow and ice storms, emission, cyclone, typhoon, pandemic, torrent , hurricane.

Specialized Threats: power disappointment/variance, warming, ventilation or air con disappointment, glitch or disappointment of hardware, disappointment of framework code, disappointment of use code, broadcast communications disappointment, gas spills, interchanges disappointment, atomic aftermath.

Human Threats: robbery, bomb dangers, theft, blackmail, thievery, defacing, psychological warfare, common problem, synthetic spill, damage, blast, war, natural pollution, radiation tainting, perilous waste, vehicle crash, airdrome nearness, strike (Internal/External), PC wrongdoing.
Instead of attempting to determine what the true chances of each disaster are, all regions and workplaces should be covered by the risk investigation. To start, an overall relative strategy plan of high, medium, and low is used to determine the likelihood that the danger would materialise. A risk analysis form that can be found here in pdf format will function with the strategy although the abilities or division can change depending on the sort of association the planning for additionally needs to confirm the impact of such a possible danger on various offices or capacities inside the association every anticipated risk should be defined and given due consideration in the strategy, and each division's potential effects on the organisation should be thoroughly examined, even though the chief framework itself poses the single greatest risk. It's not the only important issue; in fact, even in the first programmed organizations, some offices won't be managed or programmed in fully programmed divisions in the slightest degree. A pc information program bundle hangs on

disks or supporting documents for the data section, and vital records remain outside of the structure as legal records. The effect is assessed as zero (no effect) a pause between tasks Taking a pause from activities for up to eight hours has one obvious benefit. 2 incidents of instrumentation and office damage duties suspended for eight to 48 hours 3 significant damage to the offices or possibly the instrumentation tasks will be interrupted every 48 hours. All base camps and prospective PC emphasis points should be relocated. It's also crucial to continually apply evaluations to all potential threats.

Following are run of the mill suspicions which can be utilized all through the peril evaluation measure:

1. Despite the fact that influence assessments for any office may fluctuate between one and three depending on the specific circumstances, the applied appraisals should reflect the expected, evident, or predicted impact on each space.

2. Each potential risk needs to be considered as being restricted to the power being evaluated.

3. No consequences should be anticipated, notwithstanding the possibility that one prospective threat could develop into an uncommonly likely threat (a typhoon, for example, could give rise to cyclones).

4. If the danger's effects don't necessitate developing to a nursing substitute site or sites, the effect should be graded no higher than a "2."

5. The force should be in charge of assessing threats. A balanced reason rating system is used to estimate the expected risks.

Functional requirement

In making by mental demonstrations a purposeful premium describes a perform of an item group or its section a perform is depict as a lot of information sources the lead and yields see together programming conscious prerequisites is besides calculations particular nuances data the board and cooperation and elective express rationale that format what a structure need to make a few bucks activity necessities depicting the entirety of the cases anyplace the system uses the deliberate prerequisites region unit found getting utilized cases deliberate prerequisites area unit maintained by non-utilitarian necessities similarly referenced as quality conditions that power necessities on the orchestrate or execution like execution prerequisites security or dependability by and enormous purposeful prerequisites domain unit conveyed inside the sort structure should be constrained to attempt to however non-utilitarian prerequisites district unit system are the found for discipline purposeful necessities is elucidated inside the structure style they found for discipline non-utilitarian necessities is clarified inside the system organize as represented in prerequisites arranging purposeful prerequisites show explicit delayed consequences of a system this may be separated from non-utilitarian prerequisites that affirm by and monstrous characteristics like cost and light-mindedness purposeful necessities drive the gear mastermind of a framework, though non-useful necessities drive the specialized plan of a framework.

IV. CONCLUSIONS

The usefulness of data mining techniques for predicting agricultural yields is based on input factors from the climate in the suggested model. The developed website is intuitive, and every one of the additional grains and regions selected for the analysis should have reliability of prediction above 75%, suggesting higher predictive performance. By combining the Internet of Information and artificial intelligence, we have suggested a novel strategy for smart agriculture. Utilizing both current and historical data improves the accuracy of the outcome. Additionally, comparing several ML algorithms improves the system's accuracy. The farmers' challenges will be lessened by this approach, which will also improve the amount and calibre of their labour.

V. FUTURE SCOPE

The following features can be added to the system to make it even better: Utilization of moisture sensors for soil, environment sensors, and pH sensors to improve crop prediction accuracy. While recommending a crop, the market requirements for the area and the crops grown by nearby farmers can both be taken into account. The current study demonstrated the usefulness of data mining methods for forecasting agricultural yields based on climatic input variables. The developed site is intuitive, and all additional grains and regions selected for the analysis should have reliability of prediction above 75%, suggesting higher predictive performance.

Some other important points are as followed- Building this precise application in the nearby Language, in order that it might be extra comfortable for farmers. Crop illnesses detection and prevention A generalized prediction model for numerous plants via way of means of considering distinct parameters like humidity and solar radiations may be developed. Giving records approximately micro nutrients. The information processing from the climate may be qualified sufficient to support that information. Mainly if any complicated comparative of the information is needed the similarity of the weather may be indulged in the information processing area to manipulate the computation time.

VI. REFERENCES

[1] S.Veenadhari, Dr Bharat Misra, Dr CD Singh.2019."Machine learning approach for forecasting crop yield based on climatic parameters." International Conference on Computer Communication and Informatics (ICCCI), 2014.
[2] P.Priya, U.Muthaiah M.Balamurugan.Predicting yield of the crop using machine learning algorithm. International Journal of Engineering Science Research Technology (IJESRT), 2018.
[3] D Ramesh, B Vishnu Vardhan. Analysis of Crop Yield Prediction Using Data Mining Techniques. IJRET: International Journal of Research in Engineering and Technology.(IJERT),2015.
[4] F. Baheri, F. Davardoost and V. Ahmadzadeh, "Data mining with learning decision tree and Bayesian network for data replication in Data Grid", Proceedings of the Third International Conference on Contemporary Issues in Computer and Information Sciences, pp. 51- 51, 2012.
[5] Abdullah, A., Brobst, S, Pervaiz.I., Umer M.,and A.Nisar.2004. Learning dynamics of pesticide abuse through data mining. Proceedings of Australian Workshop on Data mining and Web Intelligence, New Zealand, January, 2004.
[6] Rushika Ghadge, Juilee Kulkarni, Pooja More, Sachee Nene, Priya R L, "Prediction of Crop Yield using Machine Learning," International Research Journal of Engineering and Technology (IRJET) Feb 2018, pp. 2237-2239
[7] Radhika, Narendiran, "Kind of Crops and Small Plants Prediction using IoT with Machine Learning," International Journal of Computer & Mathematical Sciences April 2018, pp. 93-97
[8] Shridhar Mhaiskar, Chinmay Patil, Piyush Wadhai, Aniket Patil, Vaishali Deshmukh, "A Survey on Predicting Suitable Crops for Cultivation Using IoT," International Journal of Innovative Research in Computer and Communication Engineering January 2017, pp. 318- 323
[9] Quinlan,J.R.(1985b). Decision trees and multi-valued attributes. In J.E. Hayes & D.Michie (Eds.), Machine intelligence 11. Oxford University Press (in press).
[10] T Raghav Kumar, Bhagavatula Aiswarya, Aashish Suresh, Drishti Jain, Natesh Balaji, Varshini Sankaran, "Smart Management of Crop Cultivation using IOT and Machine Learning," International Research Journal of Engineering and Technology (IRJET) Nov 2018, pp. 845- 850
[11] Akash Raj N, Balaji Srinivasan, Deepit Abhishek D, Sarath Jeyavanth J, Vinith Kannan A, "IoT based Agro Automation System using Machine Learning Algorithms", International Journal of Innovative Research in Science, Engineering and Technology November 2016, pp. 19938-19342
[12] Food and Agricultural Organization. (2018). The state of Agricultural Commodity Markets. 19-22.
[13] Food and Agricultural Organization. (2019). The state of Food and Agriculture. 23-27.
[14] Rajagopalan B. Lall U (1999) A k- nearest neighbor daily precipitation and other weather variables. WatResResearch35(10) :3089 – 3101.
[15] Shahin,M.A.,Tollner,E.W.,McClendon,R.W. Arabnia, H.R.,(2002). Apple classification based on surface bruises using image processing and neural networks. Trans. ASAE 45, 1619–1627.
[16] Tripathi S, Srinivas VV, Nanjudiah RS (2006).Down scaling of precipitation for climate change scenarios: a support vector machine approach, J. Hydrology 330-337.
[17] Verheyen, K., Adrianens, M. Hermy and S.Deckers(2001). High resolution continuous soil classification using morphological soil profile descriptions. Geoderma, 101:31-48.
[18] Zelu Zia (2009). An Expert System Based on Spatial Data Mining used Decision Tree for Agriculture Land Grading. Second International Conference on Intelligent Computation Technology and Automation.Oct10-11, China\