MULTI FACTOR ANALYSIS TO PREDICT BEST CROP USING RAINFALL, LOCATION AND SOIL TYPE USING MACHINE LEARNING AND BIG DATA

S.Aiswarya

Department of Computer Science

UG Student

St.peter's College of Engineering and Technology

Chennai, India

T.S.Aswitha

Department of Computer Science

UG Student

St.peter's College of Engineering and Technology

Chennai, India

Ms.J.Wisely Joe

Department of Computer Science

Assistant Professor

St.peter's College of Engineering and Technology

Chennai, India

Abstract—Machine learning presents good numerical and logical compounds which is used to analysis of large complementary, different and unorganized data file on the Big Data scale and it gently gains in obtaining the agriculture. The paper initiates its fundamental abstractions and detailed view of machine learning applications and estimating how machine learning communicate with big data computers to clarify the fundamental research of the biotech in the plant science. The application is to recognize the types of soil, water source of that land, whether that land is based on rain water or bore water and to suggest what type of crop is suitable for that soil. So through this application we can provide solutions for the people to know about the agriculture. We can estimate the type of crop which is suitable for that particular soil, weather condition, temperature and so on. So far, we are using machine learning with the set of datasets to identify the crop for the suitable soil.

Key Words: Crop Prediction, Support Vector Machine, Decision Tree Algorithm.

I. INTRODUCTION

Data mining is an essential for modern research area or the world for analysing processing, examining, extracting and measuring many number of datasets. It is used to get the relation between them, classifications, and collections, etc. Calculated factors between different attributes are mainly used to predict and the results are in certain patterns. Particularly, the above strategy can be used in agriculture field to find the relationship between the crops and soil andcan make important predictions which can be used for farmers to do a profitable business.We can add scientific factors and possibilities into consideration for analysing the relationship between soil and crops. In agriculture, always the decisions are made based on the history of the farming and nothing related to scientific factors.The decisions of farmers do not based on the soil quality or suitability and that mostly leads the loss or less yield from crops.

In data mining, enormous amount of previous data from every field can be collected and stored. The algorithms or technologies can give the common factors between any two attributes in a fraction of time. Also the strategies can be used to analyse the results and conclude a decision which will be a favourable one. We can use this analytical potential of algorithms for finding valid results, analysis reports, accurate predictions to help the farmers. Predefined parameters were already given by government agriculture research centers. Many algorithms are available to find the correlation and association among the data present in huge datasets .They are decision tree algorithms, Naive based algorithms, classification rule algorithms, Neural Network based algorithms, Supports Vector Machine and Genetic Algorithms etc., We can use different functions which can jointly work with existing algorithms and can be used for crop

prediction.Also we can use numerical weather prediction algorithms too with our crop prediction approach.

II. OBJECTIVE

The objective of the work is to predict the crop for cultivation based on sand, weather condition and water resources.

III. LITERATURE SURVEY

B.T Christensen's [1] stated that the physical fraction was used for the study of organic matter in soil. The fraction of this method overlaid darning levels of constructional and practical complication. As per their size and density of soil it practised and given importance to both organic and inorganic soils. It is recommended that structural and functional complexity where the primary level of complexity has been isolated fully from dispersed soil.

Shengchen Fang, and Hsiao-Dong Chiang [2] stated that the forecasting models included Gaussian process. The novel composite covariance function is used in exact wind forecast which is introduced.. It required both high precision and high trueness. Thus the implemented compound covariance function was formed on the investigation of joint effects among NWP. Further implementation of the proposed forecasting model was checked by 2012 global energy forecasting competitions.

Andrade José R., and Ricardo J. Bessa [3] declared that the past 2 decades, Renewable energy has to forecast and proceed to the evolution for a further physical and statistical algorithms. This work gives the additional information about the forecasting framework to find out whether the data in a grid predicting the numerical weather is used by wind energy and solar energy.

Zhao Yongning, Lin Ye, Pierre Pinson, Yong Tang, and Peng Lu [4] implemented that the huge total of wind farm have provocation and possibility in the evolution of wind power forecasting .This is pros of interconnecting in between one and hundreds of structural distributed wind farms. The outcome obtained to view execution of actual SC-VAR and other techniques. The evolution of indicator, sparsity, controllability, precise and efficiency.

Pengwei Du, HailongHui, and Ning Lu their [5] stated that the study represents of two enhancement was a short-term for wind forecast algorithm. The operation processed within few seconds to few hours which was developed to verify the performance of this method. The historic data which is used to demonstrate the regulation of requirements can meet the ERCOT reliability which gives addititional information generation of variations.

González-AparicioA.ZuckerA.Zucker [6] derived that the stochastic nature of power system is to increase the growth to share the generating electricity to different variable of renewable sources.This step is to understand the forecast unreliability. It supplies a newdetermination of forecast errors in the form of time period between the day of the closure. This methodology is evolved using time series from 2010-2013. Using clustering techniques it is used to decrease the span of uncertainty as well as regressive models. This system of action show it's suitable to develop an intra-day bidding strategies. This method could help the wind power into retail based on correct scenarios.

NischalRajbhandari ; WeifengLi ; Pengwei Du ; Sandip Sharma ; Bill Blevins [7] stated that in this analysis the information of the divine error for wind traces a fresh techniques to decide NSRS ERCOT. There are more services which are condems to help ERCOT. Non-Spin Reserve Service(NSRS) set provided by generator which can come online within 30 minutes.

Mohrlen ,et al., and their [8] stated that the increase of renewable energy source in Europe was based on consumer paid financial incentives. A direct sales approach was introduced to collaborate the ensemble forecast. The essential way to join RES successfully to the system along with higher authenticity, risk detection and detailed forecasting.

YaoZhang, ,et al., [9] Implemented that spatro temporal correlations which are captured using offsite predictions among geographical distributed find farms which is one of the solutions for better forecast precision of wind power generation. The real large scale forecasting complexity was created by multiple quantite regression.

Ezzat Ahmed Aziz,et al.,[10] stated that the large number of spatiotemporal details was piled up in wind farms. Their vast usage in separate process was caused by controlling wind farm.

IV.PROPOSED SYSTEM
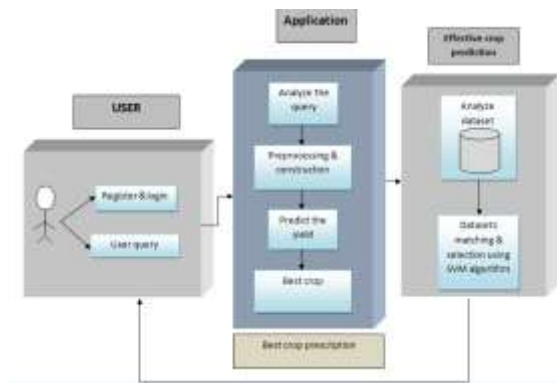
Application to identify the types of soil, water source of that land whether that land is based on rain or bore water and suggest which type of crop is suitable for that soil. So this application helps the people to know about the agriculture. This work predicts the type of crop which one is suitable for that particular soil, weather condition, temperature and so on. Using the user's location and their budget which soil is suitable for which crop. It is mainly used for non- farmers who is interested and focused in doing agriculture. This project uses machine learning with the set of data and to identify the crop for the suitable soil.

V.ADVANTAGES

- People can easily know about the crops
- People gets benefits through this application
- Educated people can easily understand about soil and crops

We can increase the income by cultivating absolute crop on the land based on the weather condition of particular location

VI.RESULTS AND DISCUSSION

Fig:1. Registration Form

Fig:2.Login/Sign in

Fig:3.Data Base

Fig:4.Data Base

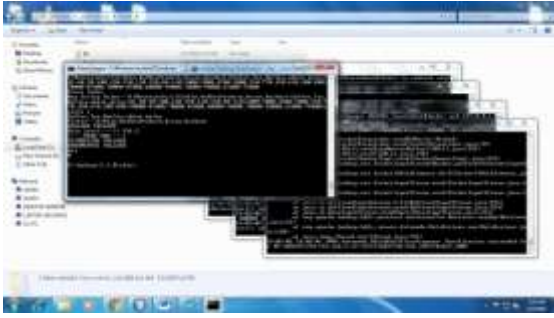**Fig:5. Command Prompt**

Fig:6 Data Base

Fig:7. Result

VII.CONCLUSION

Thus the paperconcludes that using machine learning we implement a system to predict the crop and yield for that crop. Through this app, the farmers and the normal people can get more advantages.

REFERENCES

[1] B. T. Christensen, "Physical fractionation of soil and structural and functional complexity in organic matter turnover", 2001.

[2] Shengchen Fang, and Hsiao-Dong Chiang, "A high-accuracy wind power forecasting model," IEEE Transactions on Power Systems, vol.32, pp.1589-1590, 2017.

[3] Andrade José R., and Ricardo J. Bessa, "Improving renewable energy forecasting with a grid of numerical weather predictions," IEEE Transactions on Sustainable Energy, vol. 4, pp.1571-1580, 2017.

[4] Zhao Yongning, Lin Ye, Pierre Pinson, Yong Tang, and Peng Lu. "Correlation-constrained and sparsity-controlled vector autoregressive model for spatio-temporal wind power forecasting," IEEE Transactions on Power Systems, 2018.

[5] Pengwei Du, Ross Baldick, and Aidan Tuohy, Integration of Large Scale Renewable Energy into Bulk Power Systems: From Planning to Operation, Springer, 2017.

 [6] González-Aparicio, I., and A. Zucker, "Impact of wind power uncertainty forecasting on the market integration of wind energy in Spain," Applied Energy, vol.159, pp.334-349, 2015.

[7] NischalRajbhandari, Weifeng Li, Pengwei Du, Sandip Sharma, and Bill Blevins, "Analysis of net-load forecasterror and new methodology to determine Non-Spin Reserve Service requirement," IEEE Power and Energy Society General Meeting (PESGM), 2016, 1-5.

[8] Möhrlen, Corinna, and Jess U. Jørgensen, The Role of Ensemble Forecasting in Integrating Renewables into Power Systems: From Theory to Real-Time Applications, In Integration of Large-Scale Renewable Energy into Bulk Power Systems, Springer, Cham, 2017.

 [9] Yao Zhang, Jianxue Wang, "A distributed approach for wind power probabilistic forecasting considering spatio-temporal correlation without direct access to off-site information," IEEE Transactions on Power Systems, 2018.

[10] Ezzat Ahmed Aziz, Mikyoung Jun, and Yu Ding, "Spatio-temporal asymmetry of local wind fields and its impact on short-term wind forecasting," IEEE Transactions on Sustainable Energy, 2018.