

SMART FARMING USING MACHINE LEARNING AND DATA ANALYTICS

Ms.P.Thulasi ,Mca.,Assistant Professor, Pg Department Of Computer Applications Marudhar
Kesari Jain College For Women, Vaniyambadi pragasamthulasi@gmail.com

ABSRTACT:

‘Smart farming using machine learning and data analytics’ is a Proof of concepts intended for betterment of farmers and agriculture in India. Mentioned in the report are the exact goals to be fulfilled by this chapter – namely, recommendation of the top n crops based on soil properties and environment factors. These take into consideration all the nutrients and micronutrients (around 12 of them) with reference to the soil health card. From these top n crops recommended, the farmer will also be able to know the economically beneficial crop where the model implements price prediction module for the crops. also a year round plan will be provided to farmer to maximize his cultivation and earnings through CSM. Besides generic information regarding government schemes or soil testing labs will be provided. The researcher along with these functionalities can access certain datasets that he wants to explore. The researcher can also study the impact of various algorithms on varies data. An add on ‘what if’ scenario will help not only the farmer but also the researcher to play around with various factors and see the variations in results and proceed with decision making accordingly.

Keywords:

Micronutrients: Iron,zinc,chlorin,magnesium are used by field crops in very small quantities.

Farmer: Farmers are involved in planting,cultivating,harvesting duties, live stock of farm from performing agricultural.

Introduction

‘Smart farming using machine learning and data analytics’ is a paper intended studying for betterment of farmers and agriculture in India. Mentioned in the report are the exact goals to be fulfilled by this project – namely, recommendation of the top n crops based on soil properties and environment factors. These take into consideration all the nutrients and micronutrients (around 12 of them) with reference to the soil health card. From these top n crops recommended, the farmer will also be able to know the economically beneficial crop where the model implements price prediction module for the crops. also a year round plan will be provided to farmer to maximize his cultivation and earnings through CSM. Besides generic information regarding government schemes or soil testing labs will be provided. The researcher along with these functionalities can access certain datasets that he wants to explore. The researcher can also study the impact of various algorithms on varies data. An add on ‘what if’ scenario will help not only the farmer but also the researcher to play around with various factors and see the variations in results and proceed with decision making accordingly.



1.1 Methodologies of Problem Solving

Due to the unpredictable environmental factors or price patterns, farmers are unable to decide which crop to cultivate so as to maximize profit. Even if the crop suits his land a lot of it goes waste if there is no market demand. Thus a model was needed to guide the farmer considering his land and environment and price factors. Collection of information and data: Data required in the project was to train the module for prediction. Creating other alternative solutions From user's point of view, he has been given the flexibility to choose amongst the various solutions provided by app – single or alternative crop, or a year round plan. From developing point more than 4-5 algorithms have been explored, considered, implemented to provide alternative solutions. Evaluating the consequence of different solutions: User might choose a solution as per his necessity. Algorithms were evaluated for accuracy and mean errors to improve accuracy and reduce errors. Algorithm with highest accuracy used for base model. Researcher can refer to data needs.

1.2 Background and Domain

India being an agro-based country, its economy mainly depends on agriculture produce growth and related agro-industry products. Indian agriculture largely depends on the rainfall which is highly unpredictable. Farming yield is also largely dependent on diverse and multiple soil parameters, like Nitrogen, Phosphorus, Potassium, Crop cycle, Soil moisture, soil pH, temperature and climatic aspects like temperature, rainfall, etc. Nowadays India is very rapidly progressing towards technical innovation and development. Thus, technology will be beneficial as well as blessing to agriculture which will enhance crop productivity which will result in better yields to the farming.

2. Agriculture

Different types of soils have various mineral contents and every crop require multiple mineral components for its better growth. Each and every soil has some specific characteristics and is suitable to grow only a particular number of crops. about the type of soil he has so that he could cultivate better crops. Another important issue that is farmers have to face is the uncertainty in the market demand and prices from the time of sow to the time of harvest.

2.1 Machine Learning

Machine Learning is a section of Computer Science where new developments evolve at recent times, and also helps in automating the evaluation and processing done by the mankind, thus reducing the burden on the manual human power. It is a type of Artificial Intelligence that provides devices with the ability to learn without being separately programmed. It gives focus on the development of computer programs that can change when exposed to new data. Finding out the suitable crops based on the soil's appearance becomes tedious for novice farmers. There also exists a need to prevent the agricultural decay.

2.2 Data Analytics

It is a method of analyzing, cleaning, transforming & creating models of data with the aim of finding meaningful and useful information, concluding that information and supporting decision making. Data analytics projects and initiatives support businesses increase revenues, help in improving operational and functional efficiency, responds quickly to emerging market trends and gain a competitive edge over rivals, optimizes marketing campaigns and customer service efforts.

3. Software Requirements

3.1 Analysis Models

SDLC Software development life cycle for the proposed chapter went through following phases: Requirements Gathering: The requirements for the chapter started right from the basics to see if the problem is really worth solving. Thus various agriculture colleges were visited to gather the requirements. The requirements mostly included precision in the prediction as, though existing models can be seen, most of them fail to provide accurate results. The technical requirements of platform,

technologies, etc. The results from the existing models failed due to the uncertainty in environmental conditions and lack of parameters considered for prediction. Thus it was decided to consider parameters that would yield a better result. Dataset was generated from the knowledge set by study and analysis. A lot of papers through literature survey also helped to analyses the gaps that could be worked upon for a better model. Plan and Design: Thus the software design was generated that included the software architecture, Key deliverables of the app were decided and usage of algorithms, technologies, platforms, language, database, etc. were also decided. The layout of the app and its interface were proposed. A plan was thus sketched. Build: Actual coding required was done. This included machine learning algorithms execution, front end codes – for static and dynamic web pages, machine learning codes API generation, etc. between devices. It also enables better utilization of the grid by detecting the device's efficiency at a particular time. The generated web app was tested and validated for various situational inputs. This included mostly usability testing, security testing, maintainability testing, scalability testing, login testing, database testing, compatibility testing, etc. Deploy: Presently the app is deployed as client server architecture with clients residing in the same network as that of the server. Further deployment on cloud for access over the internet is planned. Maintain: Check for bugs and tackling them is planned when will be needed. The system would be maintained for its availability, security, usability, scalability, and any updates based on users demand would be tried to fulfilled.

4. System Design

4.1 System Architecture Diagram

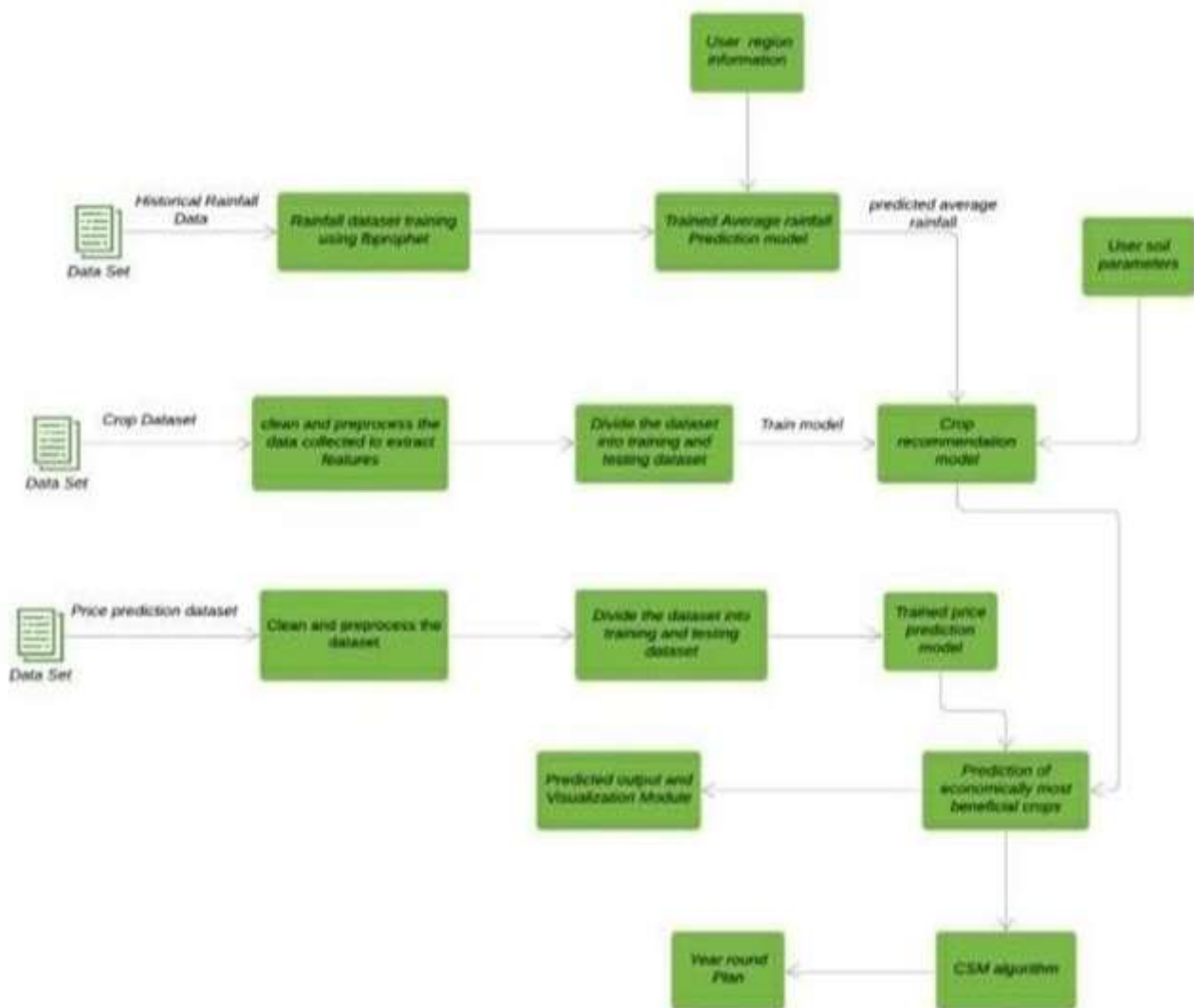


Figure 4.1 System Architecture

5. Datasets

5.1 Soil nutrient and climatic factors for crop dataset

Many researchers nowadays are working on crop recommendation and crop prediction problem despite of that finding data related to this was the biggest hurdle we faced during our project. The major focus of our project was to include climatic conditions as well as all the major soil parameters for crop recommendation because no model yet includes them all together for multiple crops. Hence, no readymade dataset was available to us. We made use of the available knowledge sets to build our final dataset which includes all the necessary soil parameters as well as climatic factors.

The next important question to discuss is why we have chosen only these particular soil parameters. During our research work in the early phase of our project we came across the concept of soil testing. We came to know that lot of soil testing laboratories have been set up by government as well as by some private institutions which performs the analysis of soil and give details of that analysis in the form of SOIL HEALTH CARD to farmer

5.2 Soil Health Card

It is Indian Government's scheme promoted by the Department of Agriculture and Co-operation under the Ministry of Agriculture and Farmer's welfare. This scheme is being implemented by the Department of Agriculture of all the State and Union Territory Governments. A Soil Health Card gives farmer soil nutrient status of his holding and advices him on the usage of fertilizers and also the needed soil amendments that he should apply to maintain soil health. A printed report of Soil Health Card has the status of soil having 12 parameters, namely N, P, K (Macro- nutrients) Fe, Cu, S, Zn, Mn, Bo (Micro nutrients); and EC, pH, OC (Physical parameters). We wanted to include all the major factors that influence the crop suitable for particular soil type hence we had to consider all these 12 factors along with type of soil average rainfall, temperature of that region.

5.3 Market Price Prediction Dataset

For future price prediction, the dataset is collected from agmarknet.nic.in in which a government website is providing users with prices of crops. It has district wise price distribution having minimum price, maximum price and model price.

5.4 CSM Dataset

The third dataset that we are using is CSM_dataset for Crop Selection Method algorithm. This dataset contains the sowing period, harvesting period, growth period, and predicted yield rate of crop

	A	B	C	D	E
	crop_name	Sowing_period	Harvesting_period	Growth_day	Predicted_yield_rate
1	Rice	June	Sept	4	2000
2	Rice	July	Oct	4	2000
3	Soybean	June	Oct	5	1264
4	Soybean	July	Nov	5	1264
5	Sweet_potato	July	Oct	4	800
6	Sweet_potato	Aug	Nov	4	800
7	Arhar	July	Dec	6	1359
8	Arhar	Aug	Jan	6	1359
9	Castor_seed	July	Feb	8	1064
10	Castor_seed	Aug	Mar	8	1064
11	Wheat	Oct	Feb	5	788
12	Wheat	Nov	Mar	5	788
13	Potato	Oct	Dec	3	1650
14	Potato	Dec	Feb	3	1650
15	Ratoil	Oct	Dec	3	1472
16	Ratoil	Nov	Jan	3	1472
17	Sarso	Oct	Dec	3	1800
18	Sarso	Nov	Jan	3	1800
19	Linseed	Oct	Feb	5	1182
20	Linseed	Nov	Mar	5	1182

Figure 5.4 CSM Dataset

6. Algorithm

6.1 Machine Learning

Machine Learning is a section of Computer Science where new developments evolve at recent times, and also helps in automation of the evaluation and processing done by the mankind and reduces the

burden on the manual human power. It is a type of Artificial Intelligence that provides devices with the ability to learn without being separately programmed. It gives focus on the development of computer programs that can change when exposed to new data. Finding the suitable crops based on the soil's appearance becomes tedious for novice farmers. There also exists a need to prevent the agricultural decay.

6.2 Neural Networks

The Neural Networks Model (NNM) is used to consider the functionality of non linear data. The model shows the capability to learn as it follows comparable data processing structure. These techniques provide successful results by applying on many problems like classification, control and prediction like a biological brain. The model is dissimilar to both classification model as well as decision tree due to its factor of likely hood prediction. The neural network has many techniques which has merits and demerits. The research suggests neural network to be better than decision tree and regression analysis model used in churn prediction.

6.3 Decision Tree

The decision tree is one of the most prominent predictive models that are used for the churn prediction based on classification.

The decision tree has of two steps:

1. Tree building - The training data set is recursively partitioned according to values of the attributes. This process continues till there are no partitions left having identical values.
2. Tree pruning - Here some values may be removed from the data which is noisy data. The largest found out error rate branches are selected and further removed in pruning. To predict accuracy and reduce complexity of decision tree is thus called tree pruning.

6.4 Data Analytics

It is a method of analyzing, cleaning, transforming & creating models of data with the aim of finding meaningful and useful information, concluding that information and supporting decision making. Data analytics projects and initiatives support businesses increase revenues, help in improving operational and functional efficiency, responds quickly to emerging market trends and gain a competitive edge over rivals, optimizes marketing campaigns and customer service efforts.

7. Conclusion

This chapter is an attempt is made to bridge the gap between labor, finance, land etc. invested in agriculture and yet its lower contribution to GDP. The datasets needed for the project have been gathered. It includes datasets with fields of land properties, environmental factors as well as the future price consideration of crops.

7.1 Future enhancement

Interactions with various agricultural colleges, labs and experts have been done for the same. As mentioned, data analytics and machine learning are the main domain areas focused on technically.

8. References

- [1] Recommendation System for Crop Identification and Pest Control Technique in Agriculture [IEEE 2019] - Avinash Kumar, Sobhangi Sarkar and Chittaranjan Pradhan.
- [2] Crop Yield Prediction and Efficient use of Fertilizers [IEEE 2019] - S.Bhanumathi, M.Vineeth and N.Rohit.
- [3] Decision Making Support System for Prediction of Prices in Agricultural Commodity [IEEE 2019] - Aman Vohra¹, Nitin Pandey², S.K. Khatri³.