Machine Learning Recommendations for Crops Based on Rain Data

Akash Kumar Bhagat, Assistant Professor

Department of Computer Science Eng., Arka Jain University, Jamshedpur, Jharkhand, India

Email Id-akash.b@arkajainuniversity.ac.in

ABSTRACT: ***India's economy is mostly based on agricultural yield growth and related agro-industry goods, since it is an agricultural nation. Agriculture is an important and necessary sector, but making even a modest profit in this field is difficult, if not impossible in certain instances. This is due to the huge impact of environmental variables on agriculture, such as water, weather, and so on. Rainwater, which is extremely unpredictable in India, has a significant impact on agriculture. India is now making tremendous progress in terms of technological development. As a consequence, technology will benefit agriculture by increasing crop production, resulting in higher yields for farmers. The loss can certainly be reduced if we concentrate more on crop selection. Crop selection is influenced by a variety of variables, including physical ones such as soil and season, economic considerations such as market price, human characteristics such as experience, crop knowledge, crop profiles, and the availability of resources like as machinery and manpower. By selecting the appropriate crop type with a high yield rate, you may make a significant profit. In this research, we use the machine learning algorithm Nave Bayes to suggest the crop with the highest production rate, taking into account environmental, physical, and economic aspects. The main goal of this project is to provide a method by which farmers of all levels, from novice to expert, may maximize their profits from agriculture while also simplifying their farming practices.***

KEYWORDS: Agriculture, Algorithm, Machine Learning, Naive Bayes, Polynomial Regression.

1. INTRODUCTION

Polynomial regression method sources are utilized to build a software that would collect rain data for south Karnataka in this study. Our polynomial regression method is trained and tested using the rain data we gathered. The rain data predicted by this method is output as a consequence of utilizing it. This information is input into the Naive Bayes algorithm, which categorizes the anticipated rain and determines which crop group it belongs to. On the basis of ultimate production, the best match of rain and crop category is determined [1]. As the world's climate changes as a result of factors such as global warming, conventional rain forecasting is becoming more difficult. And, since rain is the primary supply of water for much of Karnataka's agriculture, it has become essential to use contemporary methods and technology in order to achieve adequate agricultural yields. Our research focuses on using machine learning methods to forecast rainfall and suggest which crop would provide the highest yield. We selected southern Karnataka as the location for this project [2]. This area is divided into 15 districts, as indicated in Table 1.

Table 1: Districts in South Karnataka. The region chosen for this project is southern Karnataka region

There are 15 districts in South Karnataka:		
Bengaluru Urban	Mandya	Tumakuru
Bengaluru Rural	Kodagu	Chitradurga
Chikkaballapur	Hassan	Davanagere
Kolar	Chikmagalur	Chamarajanagar
Mysuru	Shivamogga	Ramanagara

This region's average data is what we're dealing with. The goal of this research is to use machine learning to get the highest agricultural production possible. This project will be completed in two phases. To begin, we gathered data and forecasted rainfall for this area. Second, for this research, we gather crop data (ragi, maize, and paddy) [3]. Agriculture has historically been the backbone of Karnataka's economy. With the introduction of secondary and tertiary sectors, the agricultural sector was mostly neglected, and as a consequence, its contribution to the

state's economy has gradually decreased over the past several decades. On the other hand, population expansion has exacerbated the agricultural sector's increasing demands, not just in terms of food security but also in terms of employment generation. In addition, despite the crop's adaptability for a drier environment, timely rainfall had a significant impact on production levels[4].

An intelligent gadget aimed at assisting Indian farmers in making informed choices about which crop to plant based on the sowing season, his farm's geographic location, soil qualities, and climatic variables such as temperature and rainfall. Agriculture is very important in the Indian economy. The majority of Indians are directly or indirectly reliant on agriculture for their survival. As a consequence, agriculture plays an undeniably important role in the area[5]. When it comes to deciding which crop to sow in a particular season, the vast majority of Indian farmers depend on their intuition. They take pleasure in just following traditional agricultural methods and conventions, completely ignoring the reality that crop production is largely reliant on current weather and soil conditions. A single farmer, on the other hand, cannot be expected to take into account all of the many factors that affect crop development before choosing which one to plant. A single farmer's erroneous or unwise decision may have negative repercussions for both himself and the agricultural industry of the area. A combination of Big Data Analytics and Machine Learning may be used to solve this problem[6].

Machine learning techniques to study the heuristic prediction of rainfall. Agriculture was, as we all recall, the nation's and economy's backbone. While a regular rain cycle is necessary for successful agriculture, too much or too little rain may be hazardous, resulting in crop damage. This study looks at rainfall rates from past years based on various agricultural seasons like Rabi, Kharif, and Zaid, and predicts rainfall for future seasons. In this research, the linear regression method is used to quantify different kinds of data in metrics in order to get a better knowledge of agriculture in India. We chose a genuine dataset that includes the previous year's rainfall rate, split out by season[7]. The findings of this application help farmers make the optimal decision for harvesting a certain crop according to the seasons. Agriculture model that helps farmers decide which crop to plant depending on field circumstances. To provide crop suggestions to farmers, the model, which focuses mainly on agriculture in the Telangana area, utilizes a Nave Bayes classifier. It also suggests which crops are most suited to a certain environment. Agriculture is India's main occupation. More than 70% of the workforce is employed in agriculture and associated sectors[8]. Agriculture must embrace cutting-edge technology and resources in order to feed the world's increasing population. Using big data analytics, IoT, and deep learning algorithms, crop production may be increased by a factor of 10. Because big data allows for reliable data collection, retrieval, and interpretation, its use in agriculture may benefit both farmers and the country's economic development[9].

A technique for forecasting crop production and price from a farmer's property. We predict based on numerous variables that affect agricultural production, such as rainfall, temperature, market prices, land area, and historical crop yield, using a sliding window non-linear regression technique. The research is being carried out for a number of districts in Tamil nadu, India. To react to the present socioeconomic circumstances that many farmers are experiencing, our framework would suggest the best crop selections for a farmer. Agriculture provides a livelihood for almost half of India's people, yet it only accounts for 14% of the country's Gross Domestic Product (GDP). One reason for this may be farmers' lack of crop preparation. Currently, there is no system in place to advise farmers on which crops to plant[10].

1.1 System Structure:

The goal of using Naive Bayes machine learning algorithms is to provide crop recommendations. Unlike the traditional method of planting crops based on the farmer's intuition or the season, this initiative will depend on data and computations to provide exact outcomes. Figure 1 shows the block diagram of system diagram.

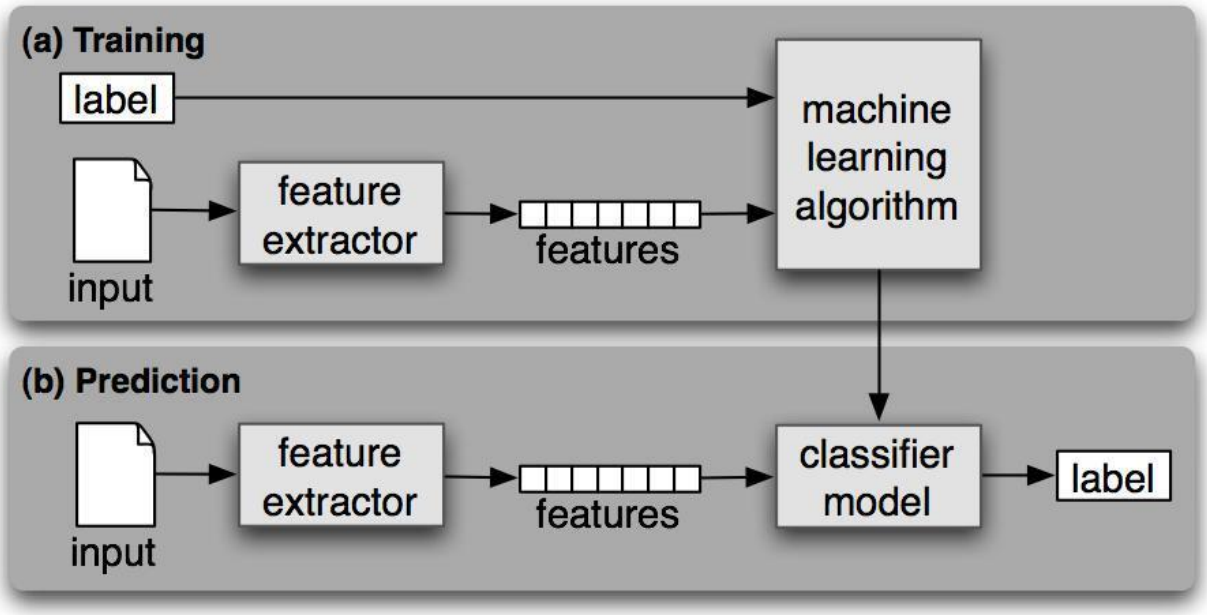

Figure 1: The block diagram of System Architecture. Naive Bayes machine learning algorithms are implemented to recommend crops for plantation.

1.1.1 Input:

Here, the input is a data collection including rainfall data from Karnataka, India.

1.1.2 Feature Extractor:

The feature extractor learns from the coding of the original data set to generate new ones, allowing it to detect important characteristics in the data.

1.1.3 Algorithms for Machine Learning:

We are utilizing polynomial regression to forecast rain data over the last 15 years, and the second Nave Bayes method to suggest crops based on rain and crop data.

1.1.4 Data for Training:

The software uses the training data to identify patterns in the data. The cross-validation data is utilized to verify that the method used to train the machine is more accurate and efficient.

1.1.5 Model of the Classifier:

We categorized the train and test data using the aforementioned machine learning method, the Nave Bayes algorithm, and used Gaussian NB to suggest crops.

1.1.6 Labels:

Labels are groups of samples that have had one or more labels applied to them. It is a label for data that defines certain features, traits, categories, or included items as shown in Figure 2.

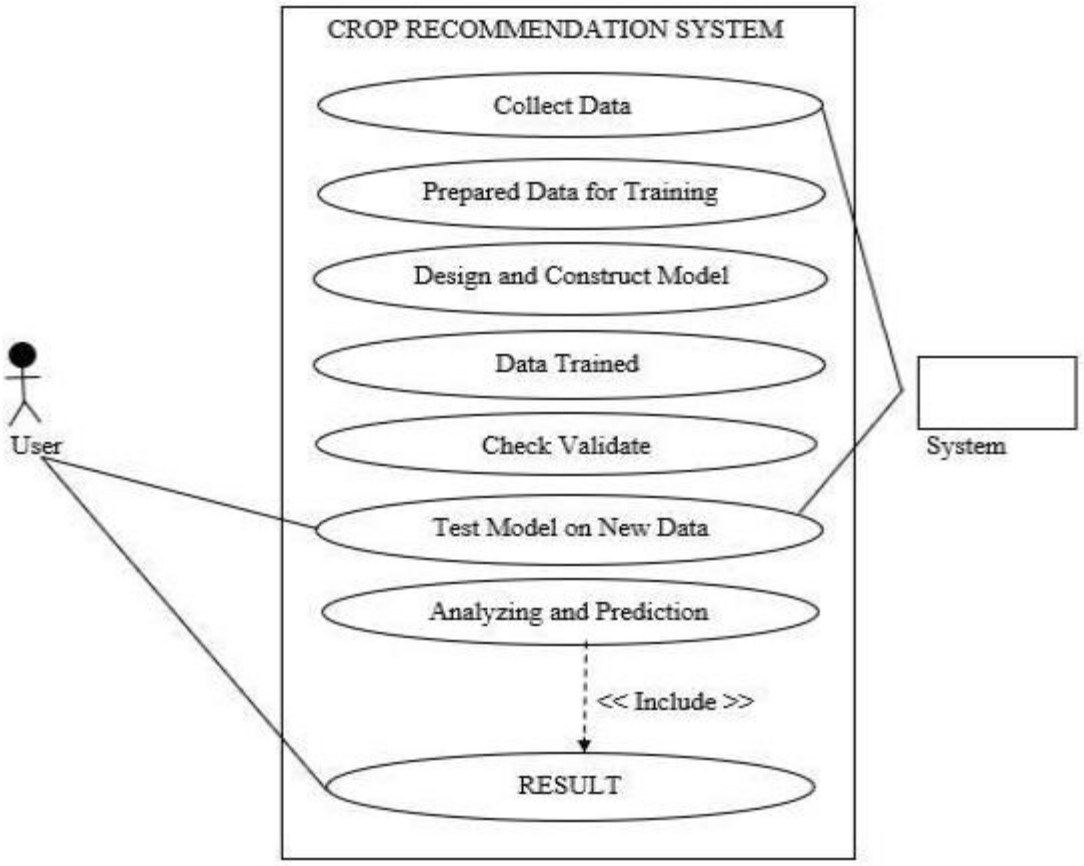

Figure 2: The layout of Use Case Diagram. Use case diagrams model the functionality of a system using actors and use cases.

1.2 Data Collection:

We start this project by collecting the required dataset for rainfall and crops for south Karnataka region. The rainfall data is collected from an official Government site of India. Contributor for this data are Ministry of Earth Sciences, India Meteorological Department (IMD). The data collected is fed into the polynomial regression algorithm. To test the accuracy of the data, data is predicted for month of January of year 2015, 2016, and 2017. Data for year 2015, 2016, 2017already exist. We compare the result of algorithm and the real data. Among the 3-algorithm used, i.e. Linear Regression, Polynomial Regression, and Random Forest, the result for Polynomial exceeds other two algorithms with 92% accuracy. Table 2 below explains the comparison. Table 3 shows the details of hardware requirements. Table 4 shows the details of software requirements.

Table 2: Comparison of Algorithms for Predicting Rain. The Result For Polynomial Exceeds Other Two Algorithms With 92% Accuracy.

January/Year	Actual Data	Linear Regression	Polynomial Regression	Random Forest	Accuracy
2015	1.7	2.8	2.4	1.5	88%
2016	3.6	2.16	3.3	3.1	92%
2017	5	2.9	1.9	4.6	92%

After the prediction of rain, project moves forward to the crops section. Data for crops is collected from an official government website for Karnataka state. The data collected for crops and the result achieved from the Polynomial Regression is fed to the Gaussian Naïve Bayes algorithms for the recommendation of crops. For the easy experience of the user, a front-end is created for this system using Python. Here user despite of having no knowledge of programming can easily make use of this recommendation system.

1.3 Hardware and Software Tools:

1.3.1 Hardware requirements:

Table 3: Hardware requirements refer to the hardware parts which are required to run the project on a machine. Following System is flexible to run this project.

Disk Space	50 GB
Memory (RAM)	512 MB<=
Desktop, Laptop	Any desktop / Laptop system with above Configuration or higher level.

1.3.2 Software requirements:

Table 4: Software Requirements. The following are the features of the machine with which the project has been built and deployed.

Operating system	Windows 7,10 OS
Platform	Windows 64-bit operating System
Programming Language/Tools	Python
Software	Anaconda or PyCharm
IDE	Python 3, Anaconda Jupyter Notebook
Interface	Tkinter
Database	SQLite or MySQLi

2. DISCUSSION

The author has discussed about that, Because India is an agricultural country, its economy is largely dependent on agricultural yield growth and associated agro-industry products. Agriculture is a vital and essential industry, yet earning even a little profit in this area is tough, if not impossible in certain cases. This is owing to the significant influence of environmental variables such as water, weather, and other factors on agriculture. In India, rainwater, which is notoriously unreliable, has a major effect on agriculture. In terms of technical advancement, India has made great strides recently. As a result, technology will help agriculture by boosting crop output and therefore raising farmer yields. If we focus more on crop selection, we can definitely minimize the loss. Crop selection is affected by a number of factors, including physical factors like soil and season, economic factors like market price, human factors like experience, crop expertise, crop profiles, and the availability of resources like equipment and people. You may earn a big profit by picking the right crop with a high yield rate. We utilize the machine learning algorithm Nave Bayes in this study to recommend the crop with the greatest production rate while taking into consideration environmental, physical, and economic factors. The

primary aim of this project is to provide a way for farmers of all levels, from beginner to expert, to increase their agricultural earnings while simplifying their farming operations.

3. CONCLUSION

The author discussed about the recommending the crop for the maximum yield on the basis of rainfall as the parameter is achieved by using Gaussian Naive Bayes of which result is show in the results and simulation section of this report. Data sources have been collected from official government source of Karnataka. This data is fed to polynomial regression algorithm to predict the rainfall. 92% of accuracy is achieved using polynomial regression. For the maximization of profit for farmers marketing can be integrated to the program by considering demand and supply as new parameters. Considering the market demand new recommendations of crops can be generated. This can prevent huge financial losses for the farmers. We have conducted this project on the basis of past data and further analyze those data to predict the rainfall. But the difference is technology we had that time and the technology we have now makes a significant impact. So along with this project, latest technology can be integrated so that the result predicted can be exceeded in the real time scenario. Estimation of crop yield can be an extension work for this project. Rainfall Data prediction can be used to plan the other measures to fulfill the required amount of water. For instance, here is plenty of cultivatable land but little rainfall is predicted then reservoir can be built so that crops can be grown even when rainfall is minimal.

REFERENCES

[1] Z. Doshi, S. Nadkarni, R. Agrawal, and N. Shah, "AgroConsultant: Intelligent Crop Recommendation System Using Machine Learning Algorithms," in *Proceedings - 2018 4th International Conference on Computing, Communication Control and Automation, ICCUBEA 2018*, 2018, doi: 10.1109/ICCUBEA.2018.8697349.

[2] R. H. L. Ip, L. M. Ang, K. P. Seng, J. C. Broster, and J. E. Pratley, "Big data and machine learning for crop protection," *Comput. Electron. Agric.*, 2018, doi: 10.1016/j.compag.2018.06.008.

[3] A. Gonzalez-Sanchez, J. Frausto-Solis, and W. Ojeda-Bustamante, "Predictive ability of machine learning methods for massive crop yield prediction," *Spanish J. Agric. Res.*, 2014, doi: 10.5424/sjar/2014122-4439.

[4] K. G. Liakos, P. Busato, D. Moshou, S. Pearson, and D. Bochtis, "Machine learning in agriculture: A review," *Sensors (Switzerland)*. 2018, doi: 10.3390/s18082674.

[5] S. Mishra, D. Mishra, and G. H. Santra, "Applications of machine learning techniques in agricultural crop production: A review paper," *Indian J. Sci. Technol.*, 2016, doi: 10.17485/ijst/2016/v9i38/95032.

[6] A. Chlingaryan, S. Sukkarieh, and B. Whelan, "Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review," *Computers and Electronics in Agriculture*. 2018, doi: 10.1016/j.compag.2018.05.012.

[7] D. Heckmann, U. Schlüter, and A. P. M. Weber, "Machine Learning Techniques for Predicting Crop Photosynthetic Capacity from Leaf Reflectance Spectra," *Mol. Plant*, 2017, doi: 10.1016/j.molp.2017.04.009.

[8] A. Crane-Droesch, "Machine learning methods for crop yield prediction and climate change impact assessment in agriculture," *Environ. Res. Lett.*, 2018, doi: 10.1088/1748-9326/aae159.

[9] J. M. Peña, P. A. Gutiérrez, C. Hervás-Martínez, J. Six, R. E. Plant, and F. López-Granados, "Object-based image classification of summer crops with machine learning methods," *Remote Sens.*, 2014, doi: 10.3390/rs6065019.

[10] A. Elvanidi, N. Katsoulas, and C. Kittas, "Automation for water and nitrogen deficit stress detection in soilless tomato crops based on spectral indices," *Horticulturae*, 2018, doi: 10.3390/horticulturae4040047.