International Journal of Advances in Engineering and Management (IJAEM)
Volume 2, Issue 3, pp: 62-66 www.ijaem.net ISSN: 2395-5252

Rain Prediction Using Polynomial Regression for the Field of Agriculture Prediction for Karnatakka

Asst. Professor Shashikala, Bikash Kumar Shah, Sudeep Thapa Chettri

Computer Science and Engineering Department, Jain University Ramanagara District, Karnataka562112, India
Corresponding Author: Asst. Professor Shashikala

Date of Submission: 15-07-2020 Date of Acceptance: 31-07-2020

ABSTRACT— In Today's era global warming is affecting all over the world which majorly effect on mankind and cause the expedite the change in climate. Due to this air and oceans are warming, sea level isrising and flooding and drought etc.India being an agricultural country, its economy predominantly depends on agriculture yield growth and allied agro-industry products. In India, agriculture is largely influenced by rainwater which is highly unpredictable. Also, the global warming adding challenges due to the unexpected drought, unpredictable rain the agriculture here is being a challenge as the traditional knowledge of crop seasoning is unable to produce enough yield. India now is rapidly progressing towards technical development. Thus, technology will prove to be beneficial to agriculture which will increase crop productivity resulting in better yields to the farmer. In this paper considering the environmental, physical and economic factors we are using Naïve Bayes to predict the rainfall which is to be used in recommendation of crops so that maximum yield can be produced.

Keywords—Machine learning application,**Polynomial Regression, Rain prediction, Naïve Bayes, Classifier,Crop recommendation, Supervised**learning

I. INTRODUCTION

In this paper we show that we have used Polynomial Regression algorithm source to code a program that will take rain data for south Karnataka. The collected rain data is used to train and test our polynomial regression algorithm. The result of using this algorithm outputs the predicted data of rain. This output is fed to the Naïve Bayes algorithm to categories the predicted rain and check which category of crops it matches with. The best match of the rain and crop category is out final output. As the world environment is changing due to factor like global warming, it is turning out quite difficult to predict the rain in a traditional way. And as rain is the main source of water for most of the agriculture done in Karnataka, it has become necessary for the use of modern ways and technologies for having the proper yield of crop. This paper focuses on the way to predict rainfall using machine learning algorithm and recommend crop that matches the amount of rainfall.

II. LITERATURE SURVEY

The paper [1] states the requirements and planning needed for developing a software model for rainfall prediction. The paper [2] makes a study of comparative study of classification algorithms and their performance in predicting the crops for the high efficiency in yield. The paper [3] uses the naïve Bayes algorithm for Crop Prediction on the Region Belts of India. For the testing of software, the data set of Rain is collected from Indian Meteorological Department website. [4] states that their work is an attempt to predict crop yield and price that a farmer can obtain from his land, by analysing patterns in past data. [5]This paper proposed a method named Crop Selection Method (CSM) to solve crop selection problem, and maximize net yield rate of crop over season and subsequently achieves maximum economic growth of the country. The proposed method may improve net yield rate of crops.[6] Recommends Naïve Bayes algorithm for the recommendation purpose. It explains the advantages on how with less range of data the algorithm can be trained and tested with reliability.

III. METHODOLOGY

3.1 Dataset Collection

Dataset for this project is collected from Indian Meteorological Department Website. Collected Dataset is of South Interior Karnataka. The data is from 2001 to 2019.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 2, Issue 3, pp: 62-66 www.ijaem.net ISSN: 2395-5252

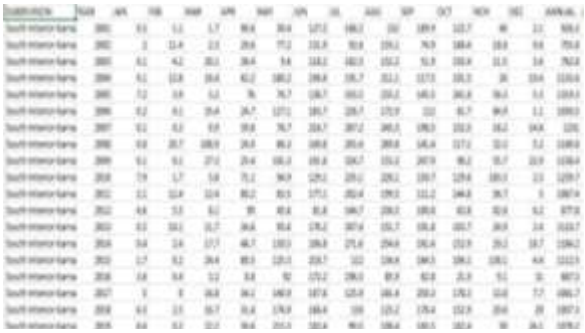


Fig: Data Set for South Interior Karnataka Region

Fig: Data Flow Diagram

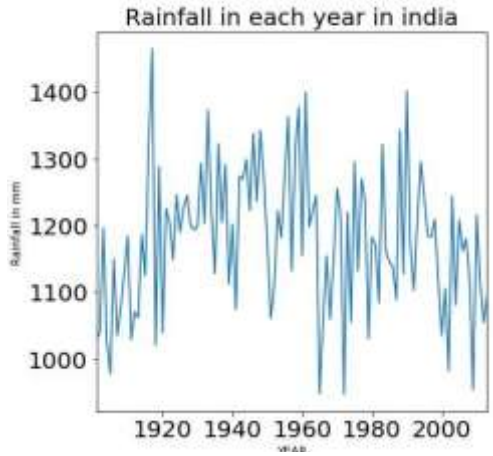


Fig: graph showing rainfall in India.

3.2 Learners Used in the Model:

POLYNIMIAL REGRESSION:

This function fits a polynomial regression model to powers of a single predictor by the method of linear least squares. Interpolation and calculation of areas under the curve are also given.

If a polynomial model is appropriate for your study then you may use this function to fit a k order/degree polynomial to your data:

$$\hat{Y} = b_0 + b_1 x_1 + b_2 x_2^2 + \ldots + b_k x_k^k$$

where Y caret is the predicted outcome value for the polynomial model with regression coefficients b1 to k for each degree and Y intercept b0. The model is simply a general linear regression model with k predictors raised to the power of i where i=1 to k. A second order (k=2) polynomial forms a quadratic expression (parabolic curve), a third order (k=3) polynomial forms a cubic expression and a fourth order (k=4) polynomial forms a quartic expression. See Klein Baum et al. (1998) and Armitage and Berry (1994) for more information.

Some general principles:

the fitted model is more reliable when it is built on large numbers of observations.

do not extrapolate beyond the limits of observed values.

choose values for the predictor (x) that are not too large as they will cause overflow with higher degree polynomials; scale x down if necessary.

do not draw false confidence from low P values, use these to support your model only if the plot looks reasonable.

NAÏVE BAYES:

Naive Bayesclassifier is a simple probabilistic classifier which works based on applying Bayes' theorem (from Bayesianstatistics) with strong naive independence assumptions. Naive Bayes is a technique for constructing classifier models which assign class labels to problem instances which are represented as vectors of feature values, where the class labels are drawn from some finite set. It is not just a single algorithm for training such classifiers, but a family of algorithms based on a common principle. All naive Bayes classifiers assumes that the value of a particular feature is independent of the value of any other feature, given the class variable.

These Learners predict the class label for each of the training data set. The class label that is predicted by the majority of the models is voted through the majority voting technique and the class label of the training data set is decided. From the ensembled models the rules are generated.

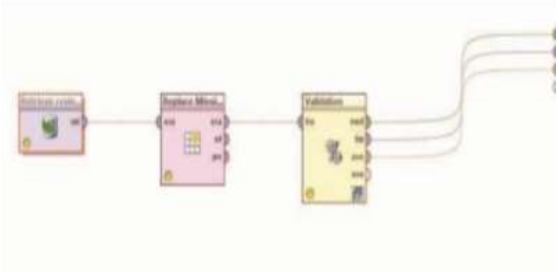


Fig: Illustrates the entire process work flow.

It shows three operators namely retrieve, replace missing values, Validation. The retrieve operator retrieves the dataset that is being uploaded in the tool. The replace the missing values operator

International Journal of Advances in Engineering and Management (IJAEM)
Volume 2, Issue 3, pp: 62-66 www.ijaem.net ISSN: 2395-5252

replaces missing values if any. Replacement can be done by four methods namely minimum, maximum, average and zero. In order to estimate the statistical performance of a learning operator a cross-validation is performed by the validation operator.

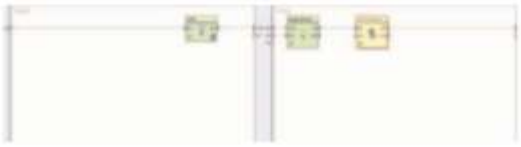


Fig: Illustrates the sub-process of cross validation operator.

The training process consists of the voting operator which is the technique that we propose for better results. On the testing sub process lies the apply model and performance operators which evaluate the correctness of the model.

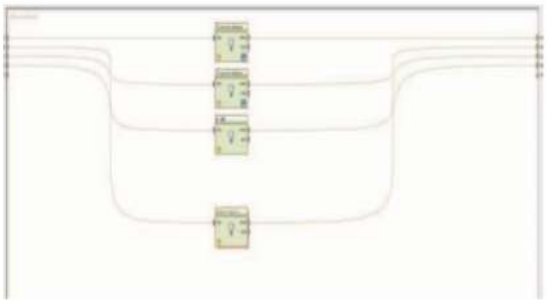


Fig: Illustrates the base learners which lie under the vote operator.

It consists of four machine learners namely Naïve Bayes. The operators corresponding to each learner is positioned. The operator performs the classification correspondingly. The tree to rules operator is used to induce rules directly from the CHAID and random tree.

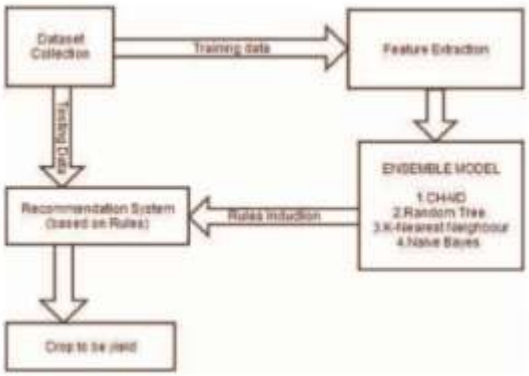


Fig: depicts the overall methodology of proposed system.

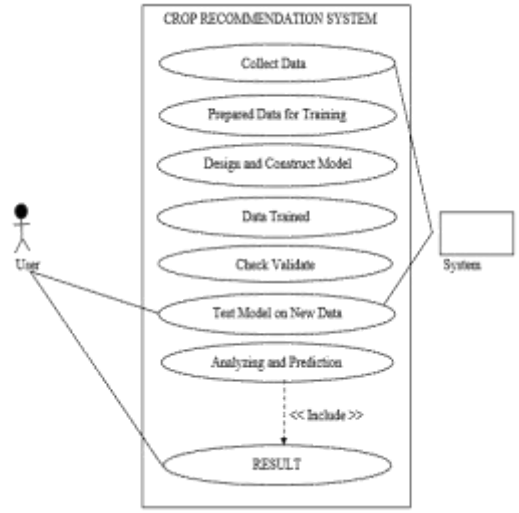


Fig: Use Case Diagram

IV. INNOVATION AND CONTRIBUTION TO THE FIELD

This project directly contributes to the agricultural field. The main objective of this project is maximizing the yield for the crops. India is a country where 50% of the workforce is directly working in the field of agriculture but its contribution to the GDP is only about 17% (Economic Survey 2018/2019). Due to the lack of proper yield of crops farmers face financial crisis every year. So, in process of maximizing the yield, we also contribute to the improved economy of farmers.

V. PROPOSED SYSTEM

In this paper we are implementing Polynomial Regression to predict the rain. The data is collected from Indian Meteorology. We chose polynomial regression for our data by testing 3 algorithms out of which polynomial regression stands 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	3.9	4.6	92%

Fig: Comparison of Algorithm for predicting rain.

For the recommendation of crops we are using Gaussian Naïve Bayes Machine Learning algorithms to recommend crops for plantation. Unlike the ancestral way where crops to be planted are based on farmer's intuition or the season, in this project we have planned to rely on data and calculations for precise results. Here we are taking rain as our parameter. Among all the parameters the reason why we chose rain is because water is the life of plants and

International Journal of Advances in Engineering and Management (IJAEM)
Volume 2, Issue 3, pp: 62-66 www.ijaem.net ISSN: 2395-5252

rain being the most effective means of watering is important to agriculture. Plants need varying amounts of rainfall to survive. For example, certain cacti require small amounts of water, while tropical plants may need up to hundreds of inches of rain per year to survive.

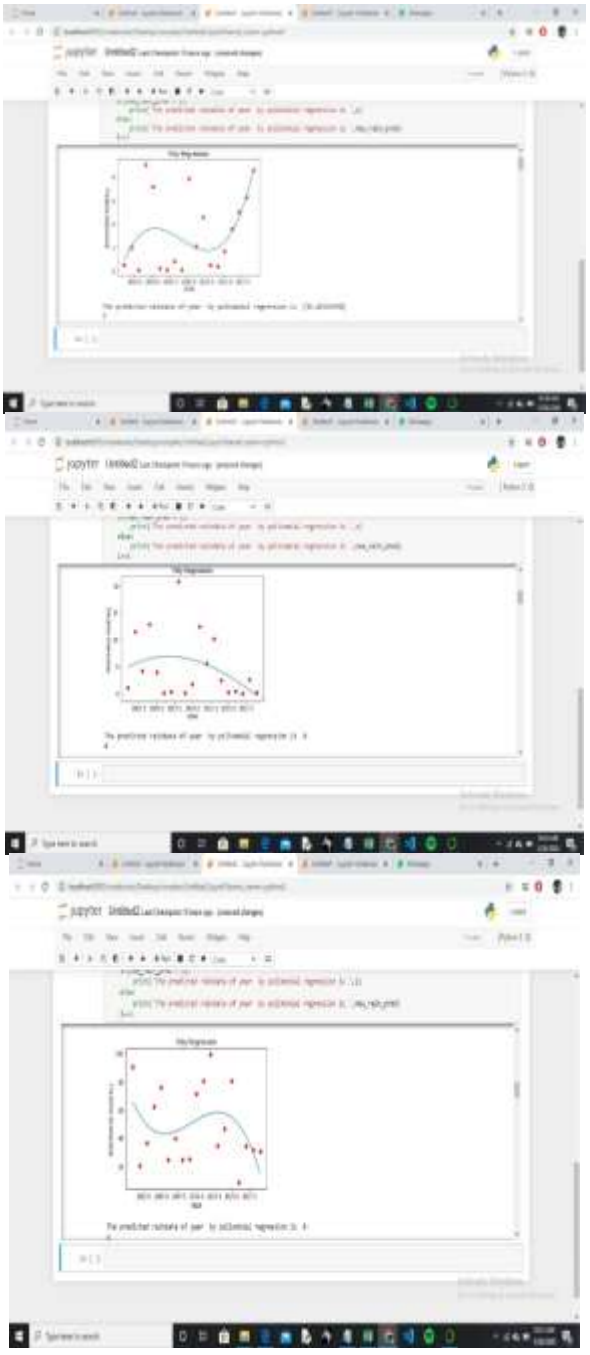
VI. INNOVATION AND CONTRIBUTION TO THE FIELD

This project directly contributes to the agricultural field. The main objective of this project is maximizing the yield for the crops. India is a country where 50% of the workforce is directly working in the field of agriculture but its contribution to the GDP is only about 17% (Economic Survey 2018/2019). Due to the lack of proper yield of crops farmers face financial crisis every year. So, in process of maximizing the yield, we also contribute to the improved economy of farmers.

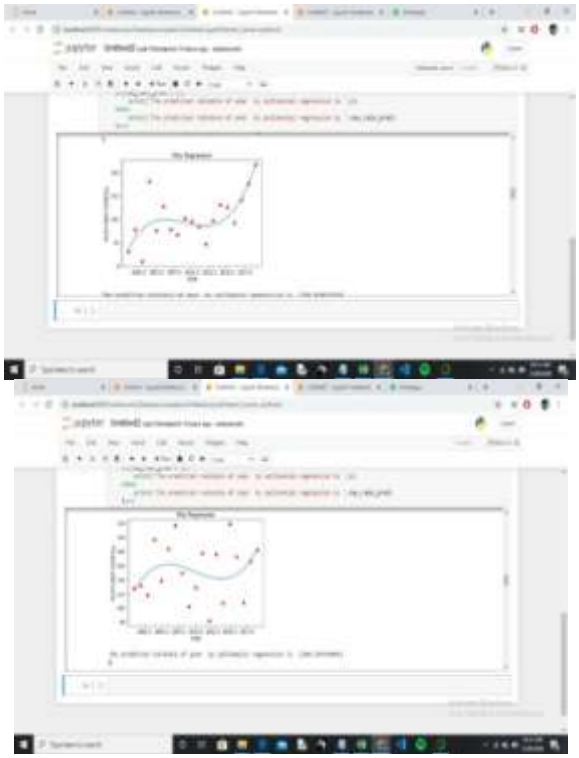
VII. RESULTS AND DISCUSSIONS

Climate change is always a major issue for whole world and making any prediction on that is now days pretty difficult and unpredictable. Climate change is due to the current global warming trend is human expansion. Due to this air and oceans are warming, sea level is rising and flooding and drought etc. One of the serious consequences due to this climate change is on Rainfall. Rainfall prediction now days is an arduous task which is taking into the consideration of most of the major world-wide authorities.

Below are the results of the predicted rainfall data for year 2020.



International Journal of Advances in Engineering and Management (IJAEM)
Volume 2, Issue 3, pp: 62-66 www.ijaem.net ISSN: 2395-5252



REFERENCES

[1]. Chandrasegar Thirumalai; K Sri Harsha; M Lakshmi Deepak; K Chaitanya Krishna 22 February 2018 IEEE - "Heuristic prediction of rainfall using machine learning techniques"

[2]. Zeel Doshi; Subhash Nadkarni; Rashi Agrawal; Neepa Shah 25 April 2019 IEEE - "AgroConsultant: Intelligent Crop Recommendation System Using Machine Learning Algorithms"

[3]. Rashmi Priya; Dharavath Ramesh; Ekaansh Khosla Department of Computer Science and Engineering, Indian Institute of Technology (ISM), Dhanbad, Jharkhand, 826004, India 03 December 2018 IEEE - "Crop Prediction on the Region Belts of India: A Naïve Bayes MapReduce Precision Agricultural Model"

[4]. S. Kanaga Subh Raja, R. Rishi, E. Sundaresan, V. Srijit, "Demand based Crop Recommender System for Farmers" in IEEE Technological Innovations in ICT for Agriculture and Rural Development, Chennai, India, pp. 194-199, 2017.

[5]. Rakesh Kumar, M.P. Singh, Prabhat Kumar and J.P. Singh (2015), '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).

[6]. Ahmad Ashari, Iman Paryud, A Min Tjoa, of 2013 'Performance Comparison between Naïve Bayes, Decision Tree and k-Nearest Neighbor in Searching Alternative Design in an Energy Simulation', International Journal of Advanced Computer Science and Applications.