Available Online at www.ijcsmc.com

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X
IMPACT FACTOR: 6.199

IJCSMC, Vol. 9, Issue. 4, April 2020, pg.11 – 14

CROPS RECOMMENDATION BASED ON RAIN DATA USING MACHINE LEARNING

Sis Maharjan; Ramsharan Diyali; Asst. Professor Shashikala
Computer Science and Engineering Department, Jain University, Ramanagara District, Karnataka 562112, India
alexsis1001@gmail.com; babudiyali986@gmail.com; shashi.hk85@gmail.com

Abstract*— 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. 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 recommend the crop which gives the high yield rate. Main objective of this project is to explore a way by which farmers from beginner to experienced level get maximum profit from agriculture and in the meantime can ease their way of agriculture.

***Keywords*— Machine learning application, Naïve Bayes, Rain Data, Crop recommendation, supervised learning**

I. INTRODUCTION

In this paper we show that we have used Naïve Bayes algorithm source to code a program that will take rain data for a specific region that we provide. Along with the rain data we also provide few suitable crops and the water it requires to grow. Our work here is to verify whether using technology in the most dominant economic field of India i.e. agriculture improves the current status of agriculture. This survey paper will show the way we plan to collect data and process our project.

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.

III.METHODOLOGY

3.1 Dataset Collection

Dataset for this project is collected from Indian Meteorological Department Website. Collected Dataset is of Karnataka Region which is further divided into 3 sub-regions. They are Coastal Karnataka, North Interior

Karnataka, South Interior Karnataka. We have rain data of each month of above-mentioned sub-regions. The data is from 1901 to 2017.

	SUBDIVISION	YEAR	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL	JF	MAM	JJAS	OND
2	Coastal Karnataka	1901	1.8	0.6	10.7	52.4	81.6	960.9	991.2	606.4	108	120.5	104.9	17.8	3056.9	2.4	144.7	2666.6	243.3
3	Coastal Karnataka	1902	3.2	0.3	4.9	10.2	54.6	698.4	1401.6	454.2	708.4	180.4	50.8	132.2	3699.2	3.5	69.7	3262.6	363.5
4	Coastal Karnataka	1903	0.7	0	0	4.1	202.8	536.5	1405.5	593.8	304.4	185	79.3	5.3	3317.4	0.7	206.9	2840.1	269.6
5	Coastal Karnataka	1904	2.4	0	4.8	23.7	93.2	1108.2	1070	465.6	245.3	127.2	0.7	0	3141.1	2.4	121.7	2889.1	127.9
6	Coastal Karnataka	1905	0	0.2	0	6.4	83.1	767.3	777.3	586.9	172.9	222.2	36.1	0	2652.3	0.2	89.4	2304.3	258.4

Fig: Data Set for Coastal Karnataka Region

230	North Interior Karnataka	2012	28.5	6.2	0.4	35.4	19.5	60	114.5	105.5	79.2	85.2	46.5	2.9	583.8	34.7	55.3	359.3	134.5
231	North Interior Karnataka	2013	1.2	6.1	3	25.4	47.4	99.4	160.7	73.9	201	101	4.2	0.1	723.2	7.3	75.7	534.9	105.2
232	North Interior Karnataka	2014	0	6.1	29.2	26.4	93	50.4	136.8	205.2	90.2	80.3	25	14.1	756.8	6.1	148.7	482.5	119.5
233	North Interior Karnataka	2015	2.4	0	27.5	50.8	45.3	89.6	38.5	78.4	150.8	61.2	5.7	1.7	551.9	2.4	123.6	357.2	68.6
234	North Interior Karnataka	2016	0.3	0.4	6.1	11	50	136.8	162.6	70	155	28.7	3.1	1	625	0.7	67.2	524.3	32.8
235	North Interior Karnataka	2017	0	0	10.7	9.9	44.8	131.3	76.8	105.8	206.6	147.9	4.9	0.9	739.6	0	65.5	520.5	153.7

Fig: Data Set for Coastal Karnataka Region

348	South Interior Karnataka	2013	0.5	10.1	11.7	34.6	95.6	176.2	307.4	151.7	191.8	103.7	24.9	2.4	1110.7	10.6	142	827.1	131
349	South Interior Karnataka	2014	0.4	2.4	17.7	46.7	130.5	106.8	271.6	254.6	161.6	152.9	20.2	18.7	1184.2	2.8	195	794.5	191.8
350	South Interior Karnataka	2015	1.7	0.2	24.4	80.5	125.3	218.7	112	136.6	164.5	106.1	138.1	4.4	1112.5	1.9	230.2	631.8	248.6
351	South Interior Karnataka	2016	3.6	0.4	3.2	8.8	92	172.2	196.5	85.9	62.8	21.9	9.1	31	687.3	4	103.9	517.4	61.9
352	South Interior Karnataka	2017	5	0	16.8	34.1	140.9	107.6	125.9	181.4	250.3	178.3	13.8	7.7	1061.7	5	191.8	665.2	199.8

Fig: Data Set for Coastal Karnataka Region

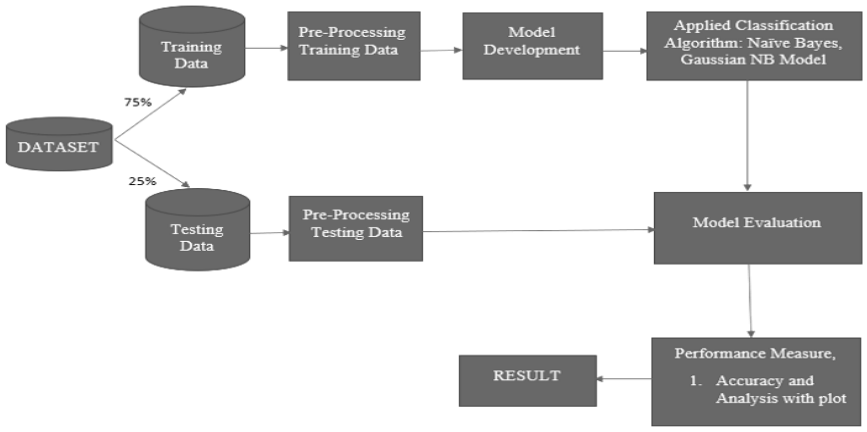

Fig: Data Flow Diagram

3.2 Learners Used in the Model:

NAÏVE BAYES:

Naive Bayes classifier is a simple probabilistic classifier which works based on applying Bayes' theorem (from Bayesian statistics) 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.

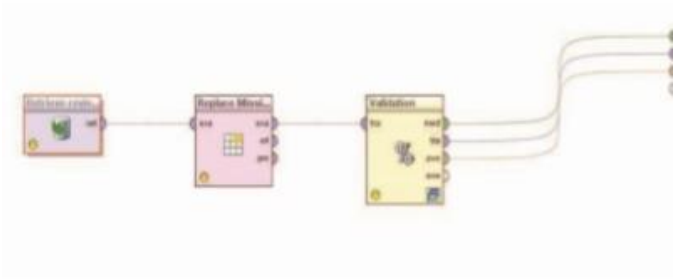

Fig: Illustrates the entire process work flow.

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 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.

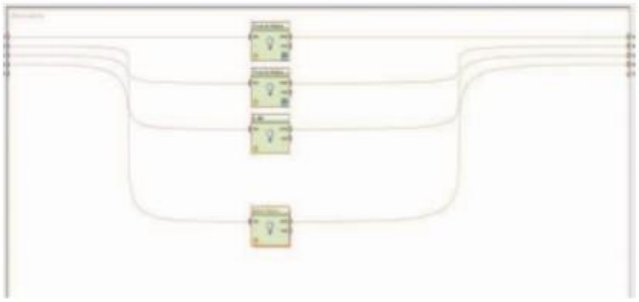

Fig: Illustrates the base learners which lie under the vote operator.

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.

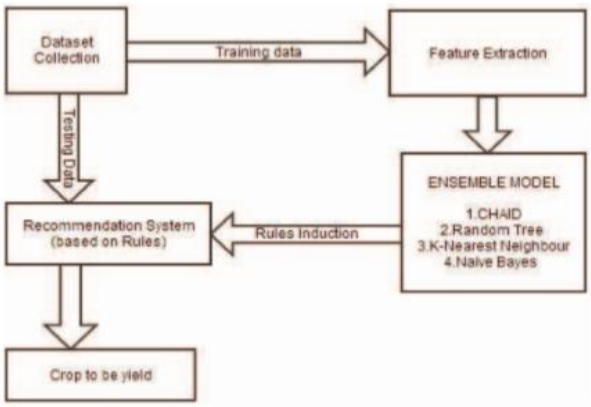

Fig: depicts the overall methodology of proposed system.

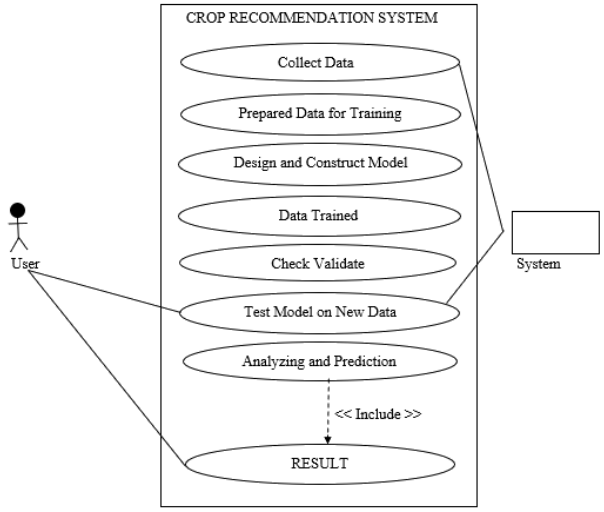

Fig: Use Case Diagram

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 project we are implementing Naive 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 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.

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".

