CROP FORE CASTING USING MACHINE LEARNING

Suhaila M.P[1], S. Hemalatha[2]*

ABSTRACT

India is mainly an agricultural country. Agriculture is the process of using land to grow a variety of crops. It hasbeen cultivated in India for a long time and is known as the backbone of the Indian economy. The majority of the people in India are directly or indirectly dependent on agriculture fortheir livelihood.[3]this leads to reduce yield increase strain on agricultural resources, exacerbating the challenges faced by farmers worldwide. Crop forecasting depends on soil, geography and climatic characteristics, and ac- curate cropfore casting increases crop production.[1]Further more, this system can offer real-time monitoring and feedback, allowing farmers to make informed decisions throughout the crop cycle, ultimately optimizing yields and promoting environmentally responsible farming practices.

Keywords : crop prediction, machine learning, support vector machine, decision trees.

I. INTRODUCTION

For a country, one of the main parts of its development rotates around creating food potential. For ages, the creation of fundamental food crops has been associated with agribusiness[1]. Farming is the significant occupations in India. It is a tremendous financial area and assumes a significant part in the general improvement of the country. Prior crop expectations depended on the experience of ranchers in a specific region. They pick just the old or neighborhood or more popular yield nearby around their territory and they have close to zero familiarity with the substance of soil supplements like soil nitrogen, phosphorus and potassium. The current situation is that inadequate utilization of soil nutrients without crop rotation leads to reduced yields, soil contamination and damage to the top layer. Given the challenges posed by urbanization and globalization, preserving and enhancing land use and fertility in the agricultural sector becomes paramount[3]. As arable land availability remains limited, there is a pressing need to maximize productivity through accurate crop forecasting methods, which agricultural researchers are actively exploring to determine the most suitable crops for various regions worldwide.

Machine learning, a fast-growing approach, allows it to expand and make practical decisions to make all are as at the forefront of its applications. The main emphasis will beon quality cultivating rather than undesirable environmental factors[5].Different AI classifiers, for example, Calculated Relapse, Nave Bayes and Irregular Woodland to make exact expectations and to monitor conflicting patterns in temperature and precipitation[1]. The planned framework will suggest the suitable yield for the specific land. In light of soil boundaries like environment boundary, precipitation, temperature, stickiness and pH. They are gathered from Government web-site,indiastat.org and Meteorological Division. The framework takes the expected contribution from the cultivator or from sensors like temperature, moistness and pH. This multitude of information are material to AI prescient calculations, for example, Backing Vector Machine (SVM) and Choice Tree distinguish designs inside information and cycle them as indicated by input conditions. The framework prescribes the harvest to the rancher and how much supplements to be added to the gauge crop. This framework has a few different elements like surmised yield per q/section of land, seeds expected for development in kg/section of land and market cost of the harvest.

II. METHDOLOGY

The proposed system aims to forecast the optimal crop for specific land by analyzing climatic parameters including soil composition, temperature, humidity, pH levels, and rainfall patterns[10].

[1,2]Department of Computer Science,
Karpagam Academy of Higher Education, Coimbatore, Tamil Nadu, India
* Corresponding Author

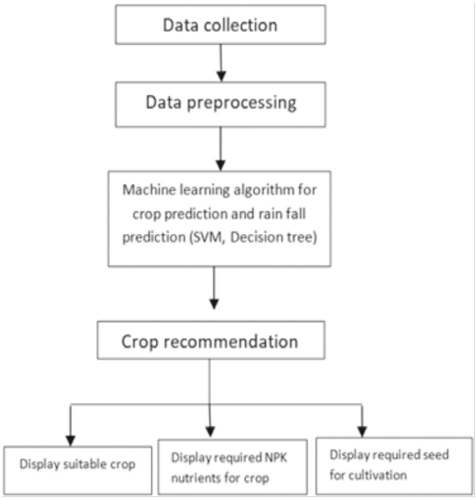
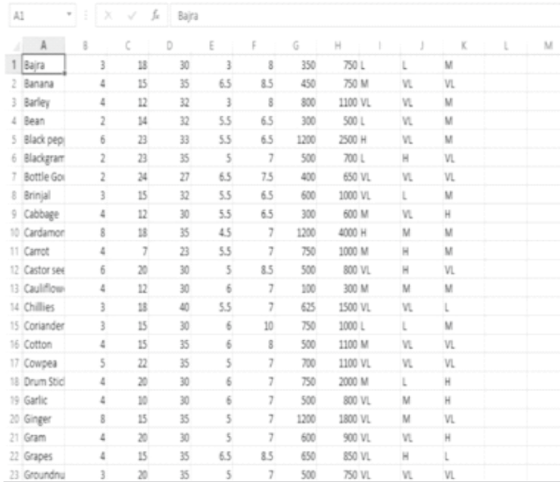


Fig.1. System Architecture

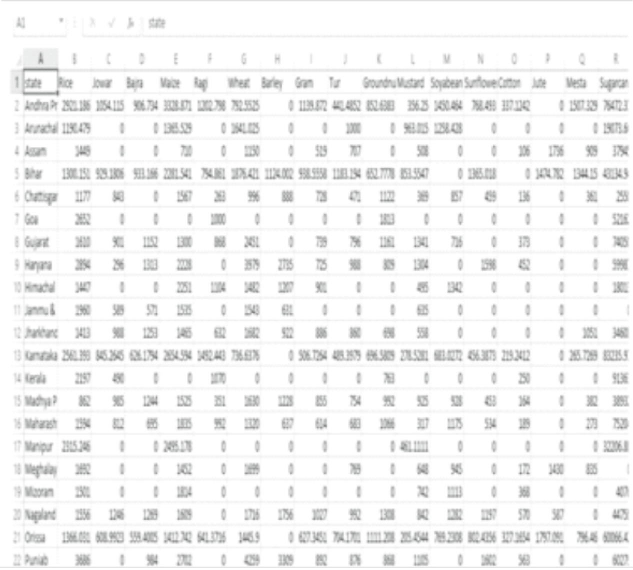


	A	B	C	D	E	F	G	H	I	J	K
1	Bajra	3	18	30	3	8	350	750	L	L	M
2	Banana	4	15	35	6.5	8.5	450	750	M	VL	VL
3	Barley	4	12	32	3	8	800	1100	VL	VL	M
4	Bean	2	14	32	5.5	6.5	300	500	L	VL	M
5	Black pep	6	23	33	5.5	6.5	1200	2500	H	VL	M
6	Blackgram	2	23	35	5	7	500	700	L	H	VL
7	Bottle Go	2	24	27	6.5	7.5	400	650	VL	VL	VL
8	Brinjal	3	15	32	5.5	6.5	600	1000	VL	L	M
9	Cabbage	4	12	30	5.5	6.5	300	600	M	VL	H
10	Cardamon	8	18	35	4.5	7	1200	4000	H	M	M
11	Carrot	4	7	23	5.5	7	750	1000	M	H	M
12	Castor see	6	20	30	5	8.5	500	800	VL	H	VL
13	Cauliflow	4	12	30	6	7	100	300	M	M	M
14	Chillies	3	18	40	5.5	7	625	1500	VL	VL	L
15	Coriander	3	15	30	6	10	750	1000	L	L	M
16	Cotton	4	15	35	6	8	500	1100	M	VL	VL
17	Cowpea	5	22	35	5	7	700	1100	VL	VL	VL
18	Drum Stic	4	20	30	6	7	750	2000	M	L	H
19	Garlic	4	10	30	6	7	500	800	VL	M	H
20	Ginger	8	15	35	5	7	1200	1800	VL	M	VL
21	Gram	4	20	30	5	7	600	900	VL	VL	H
22	Grapes	4	15	35	6.5	8.5	650	850	VL	H	L
23	Groundnu	3	20	35	5	7	500	750	VL	VL	VL

Fig.3. Temperature, Rainfall and Nutrients data set

A. Data Collection for Crop Prediction

Data collection is the most productive strategy for gathering and estimating information from different sources, for example, government websites, indiastat.org and climate sites[1].To get an estimated dataset for the planned framework. This dataset ought to contain the accompanying credits like Soil PH, Temperature, Humidity, Rain, Harvest information ,NPK values, those boundaries will be considered for crop estimating. For the yearly precipitation estimate, we gather the earlier year's precipitation information.



	state	Rice	Jowar	Bajra	Maize	Ragi	Wheat	Barley	Gram	Tur	Groundnu	Mustard	Soyabean	Sunflowe	Cotton	Jute	Mesta	Sugarcan
2	Andhra Pr	2921.186	1054.115	906.734	3328.871	1202.798	792.5525	0	1139.872	441.4852	852.6383	356.25	1450.464	768.493	337.1242	0	1507.329	76472.3
3	Arunachal	1190.479	0	0	1365.529	0	1641.025	0	0	1000	0	963.015	1258.428	0	0	0	0	19073.6
4	Assam	1449	0	0	710	0	1150	0	519	707	0	508	0	0	106	1736	909	3794
5	Bihar	1300.151	929.1806	933.166	2281.541	794.861	1876.421	1124.002	938.5558	1183.194	652.7778	853.5547	0	1365.018	0	1474.782	1344.15	43134.9
6	Chattisgar	1177	843	0	1567	263	996	888	728	471	1122	369	857	499	136	0	361	255
7	Goa	2652	0	0	0	1000	0	0	0	0	1813	0	0	0	0	0	0	5216
8	Gujarat	1610	901	1152	1300	868	2451	0	739	796	1161	1341	716	0	373	0	0	7405
9	Haryana	2894	296	1313	2228	0	3979	2735	725	988	809	1304	0	1598	452	0	0	5998
10	Himachal	1447	0	0	2251	1104	1482	1207	901	0	0	495	1342	0	0	0	0	1801
11	Jammu &	1960	589	571	1535	0	1543	631	0	0	0	635	0	0	0	0	0	
12	Jharkhanc	1413	988	1253	1465	632	1682	922	886	860	698	558	0	0	0	0	1051	3460
13	Karnataka	2561.393	845.2645	626.1754	2654.594	1492.443	736.6376	0	506.7264	489.3975	696.5809	278.5281	683.0272	456.3875	219.2412	0	265.7269	83235.9
14	Kerala	2197	490	0	0	1070	0	0	0	0	763	0	0	0	250	0	0	9136
15	Madhya P	862	985	1244	1525	351	1630	1228	855	754	992	925	928	453	164	0	382	3891
16	Maharash	1594	812	695	1835	992	1320	637	614	683	1066	317	1175	534	189	0	273	7520
17	Manipur	2315.246	0	0	2495.178	0	0	0	0	0	0	461.1111	0	0	0	0	0	32206.8
18	Meghalay	1692	0	0	1452	0	1699	0	0	769	0	648	945	0	172	1430	835	
19	Mizoram	1501	0	0	1814	0	0	0	0	0	0	742	1113	0	368	0	0	407
20	Nagaland	1556	1246	1269	1609	0	1716	1756	1027	992	1308	842	1282	1197	570	587	0	4475
21	Orissa	1366.031	608.9923	559.4005	1412.742	641.3716	1445.9	0	627.3451	704.1701	1111.208	205.4544	789.2308	802.4356	327.1654	1797.091	796.46	60066.4
22	Punjab	3686	0	984	2782	0	4259	3309	892	876	868	1105	0	1602	563	0	0	6027

Fig.2. Crop yield dataset

State	N	P	K			
A& N	VL	VL	L	ppm(parts per million)		
A P	L	VH	M	Nitrogen	Phosphorus	Potassium
Karnatak	H	M	M	VL: 10	VL: > 5	VL: >100
Assam	M	L	VL	L:10-20	L:5 -10	L: 100-150
Bihar	VL	VH	H	M:20-30	M:10-20	M:150-250
Goa	M	VL	M	H:30-40	H: 20-30	H: 250-300
Gujarat	VL	VH	H	VH:40+	VH: 30+	VH: 300+

Fig.4. Soil Nutrients distribution as per state

B. Data preprocessing

In the wake of gathering datasets from various sources. The dataset should be preprocessed before the model can be prepared[5]. Data preprocessing is expected for effective representation of data. It tends to be finished in a few stages, beginning with reading the collected dataset and proceeding with the data cleaning process. In data cleaning, datasets contain a few undesirable credits, those credits are cropping conjecture isn't thought of. In this manner, we really want to dispose of superfluous properties and datasets that contain a few missing qualities or dispose of these missing qualities or fill in pointless non-values for better exactness [6]. Then, at that point, characterize the motivation behind a model. After clearing the information, the dataset will be separated into preparation and test.

	A	B	C	D	E	F	G	H	I	J	K
1	State_N	District_	Crop_Ye	Season	Crop	Area	Production	Rainfall	Temper.	Humidit	Windspeed
2	Kerala	ALAPPU	1997	Whole Y	Arecanu	2253	1518	271	24.54	79.64	1.88
3	Kerala	ALAPPU	1999	Whole Y	Arecanu	2308	1043	242.9	23.97	80.66	2.12
4	Kerala	ALAPPU	2004	Whole Y	Arecanu	2376	1006	240.5	24.28	79.87	2.05
5	Kerala	ALAPPU	2007	Whole Y	Arecanu	1696	687	290.8	24.35	79.08	1.97
6	Kerala	ALAPPU	2008	Whole Y	Arecanu	1577	955	210.4	23.98	81.34	1.87
7	Kerala	ALAPPU	2011	Whole Y	Arecanu	1615.4	659.29	252.9	24.06	80.86	1.99
8	Kerala	ERNAKL	1998	Whole Y	Arecanu	3604	1941	262.6	24.78	79.9	2.15
9	Kerala	ERNAKL	2003	Whole Y	Arecanu	5275	3813	199.6	24.48	80.6	1.89
10	Kerala	ERNAKL	2007	Whole Y	Arecanu	5207	6395	290.8	24.35	79.08	1.97
11	Kerala	ERNAKL	2010	Whole Y	Arecanu	4549.9	4889.9	261	24.54	80.84	1.99
12	Kerala	ERNAKL	2014	Whole Y	Arecanu	4133	4533	253.9	24.66	79.45	1.93
13	Kerala	IDUKKI	2005	Whole Y	Arecanu	4009	4669	252.6	24.34	82.23	2.03

Fig.5. Data Preprocessed

C. Machine Learning Algorithm for Prediction

[2]Machine learning finds extensive application in agriculture, particularly in analyzing large datasets to identify significant classifications and patterns. The primary objective of machine learning is to extract insights from data and transform it into a coherent structure for practical applications. This study focuses on analyzing crop yield methods based on available data, employing machine learning techniques to predict yields and enhance crop profitability.

a) Support VectorMachines:[8] Support vector machines (SVMs) constitute a set of supervised learning techniques utilized for classification, regression, and outlier detection, primarily serving as a classification tool. During the algorithm execution, information regarding the n items in the n-dimensional space is mapped to specific coordinates, with each feature represented by the corresponding coordinate value. SVM, as a discriminative classifier, effectively segregates data points using a hyperplane determined by labeled training data, aiming to best classify new examples. In the context of rainfall forecasting, an external dataset containing historical rainfall data is initially loaded and preprocessed[1]. Following data preprocessing, the SVM classifier with a radial basis function kernel is trained using the dataset. Subsequently, the classifier is applied to the training set, and after fitting and testing, the model is capable of predicting the forthcoming annual rainfall. The predicted rainfall serves as a crucial input parameter for the crop forecasting system.

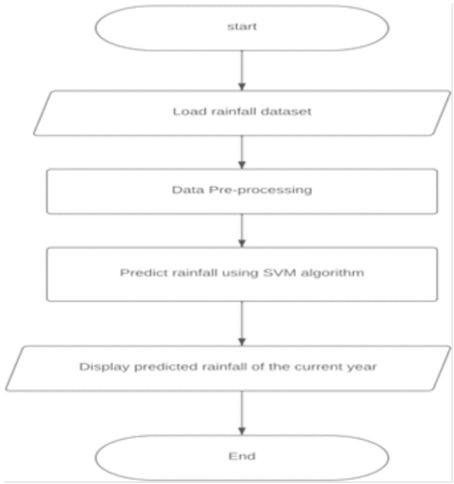


Fig.6. Rainfall prediction flow chart

b) Decision tree: [2] The decision tree stands as a versatile classification algorithm within the realm of machine learning, crafting a model from observations based on defined parameters. This tree structure unfolds recursively from top to bottom, with each node representing either a leaf node or a decision node. Decision tree techniques are renowned for their simplicity, ease of comprehension, and relevance to decision-making processes. Each decision tree comprises interconnected internal and external nodes, where the former serve as decision-making entities, and the latter act as subsequent nodes to traverse. Leaf nodes, conversely, lack child nodes and are associated with specific labels. The overarching objective is to distill overarching rules from contextual experiences. The decision tree algorithm constructs a model that predicts the value of a target variable by discerning straightforward decision rules from data features. [7]Following dataset ingestion, preprocessing occurs across several stages, culminating in model training using the Decision Tree Classifier on the training set. In the context of crop forecasting, variables such as temperature, humidity, soil pH, and projected rainfall are considered, either manually inputted or sourced from sensors, and subsequently integrated into a list alongside predicted rainfall. Leveraging this data, the Decision Tree algorithm forecasts the crop type.

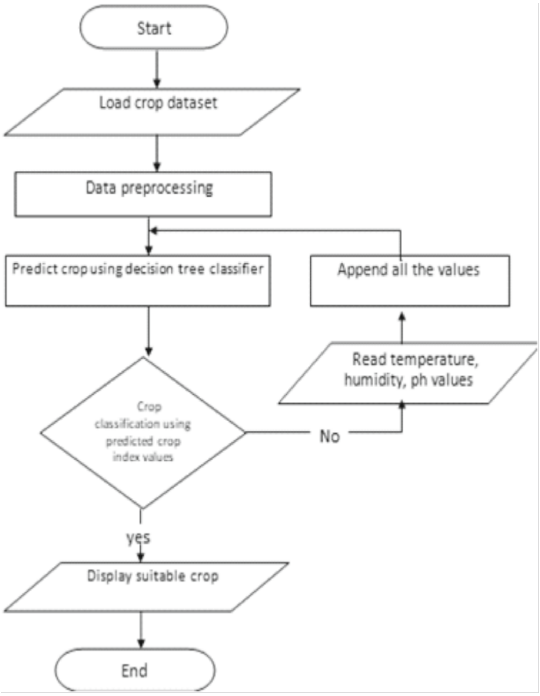


Fig.7. Flow chart for crop prediction.

D. Crop Recommendation

[4]The system recommends the optimal crop for cultivation by considering predicted rainfall, soil composition, and climatic conditions. Additionally, it offers information on necessary fertilizers, including Nitrogen (N), Phosphorus (P), and Potassium (K) per kilogram per hectare, along with seed requirements for each crop. Furthermore, it suggests the recommended crop yield in kilograms per acre for optimal harvest. These comprehensive details empower farmers to select the most lucrative crop options.

III. RESULT ANALYSIS

The proposed system suggests the optimal crop for a given land, considering factors such as annual rainfall, temperature, humidity, and soil pH. To predict annual rainfall, the system utilizes the SVM algorithm based on previous year's data, while users input other parameters. In the Output section, the system showcases the recommended crop, necessary seeds per acre, market price, approximate yield, and uses NPK values from the System Input section to indicate the required NPK for the suggested crop.

Required parameter for crop prediction			Predicted Crop	Entered Soil nutrients(Kg/ha)			Required nutrients for Crop (Kg/ha)			Required seed for cultivation (Kg/acre)	Approximated yield (quintal/acre)	Market price (Rs /quintal)
pH (0-14)	Temperature (°c)	Humidity (%)		N	P	K	N	P	K			
6.6	28	88	GROUNDNUT	00	16	173	40	24	-113	45	3-4	4000-5000
7.96	27	79	WATERMELON	00	16.95	613.0	200	83.5	-513	0.3	180-200	800-1200
7.6	23	80	SUGARCANE	00	4.5	245.0	200	145.5	-95	1000-1500	400-600	2000-2500
7.04	25	89	ONION	00	56.5	442.0	60	3.5	-412	350	80-100	800-1200
9	29	82	GREEN GRAM	316.68	22.2	163	-291.32	27.5	-138	6-8	2-3	700-1000

Fig.8.Tested Output result

IV. CONCLUSION

A specific technology aids farmers in determining the most suitable crop for cultivation in their fields, aiming to ensure efficient and productive harvesting. Given the current underutilization of technology and analysis among farmers, there is a risk of selecting inappropriate crops, which can adversely impact their income. To mitigate such losses, we've developed a user-friendly system utilizing a graphical user interface (GUI). This system predicts the ideal crop for a particular piece of land, recommends necessary nutrient additions, specifies the required seeds for cultivation, and estimates the expected yield market price. By empowering farmers to make informed decisions, this system fosters agricultural development through innovative concepts. Additionally, in the future, rainfall forecasting will be utilized to assess the need for additional water availability. Leveraging an enhanced dataset with a multitude of attributes, the system facilitates improved yield forecasts, thereby contributing to food security by averting potential food crises.

REFERENCES

[1] Nischitha, K., Mahendra N. Dhanush Vishwakarma, and Manjuraju MRAshwini. ”Crop prediction using machine learning approaches.” Int. J.Eng. Res. Technol.(IJERT) (2020): 23-26.

[2] Rajak, Rohit Kumar, et al. ”Crop recommendation system to maximizecrop yield using machine learning technique.” International ResearchJournal of Engineering and Technology 4.12 (2017): 950-953

[3] http://scikit-learn.org/stable/modules/tree.html

[4] https://www.data.gov.in

[5]Ashok,Tatapudi,andV.P.Suresh.”Prediction of Crops based on Environmental Factors using IoT Machine Learning Algorithms.”Int. J.Innov.Technol.Explor.Eng9(2019):5395-5401.

[6] https://www.javatpoint.com/machine-learning-support-vector-machinealgorithm

[7] https://en.wikipedia.org/wiki/Agriculture

[8] https://openweathermap.org

[9]Kumar, Aayush, OmenRajendraPooniwala, and Swapneel Chakraborty. Intelligent Crop Recommendation System Using ML. Diss. CMR Insti-tute of Technology. Bangalore, 2020

[10] AnakhaVenugopal, Aparna S, Jinsu Mani, Rima Mathew, VinuWilliams, 2021, Crop Yield Prediction using Machine Learning Al-gorithms, INTERNATIONAL JOURNAL OF ENGINEERING RE-SEARCHTECHNOLOGY(IJERT)NCREIS–2021 (Volume09–Issue13)