Prediction Based Sugar Cane Farming in Western Maharashtra Using **Data Mining**

Prof.M.M.Hajare¹, Sucheta S. Patil², Snehal S. Patil³, Rutuja A. Jadhav⁴, Aishwarya S. Sawant⁵, Mohini A. Chougule⁶

¹Assistant Professor, Department of Computer Science & Engineering, Sanjeevan Engineering & Technology *Institute Panhala(MH, India)*

^{2,3,4,5,6} Student, Department of Computer Science & Engineering, Sanjeevan Engineering & Technology Institute Panhala(MH, India).

***_____

Abstract: In India, Most of Indians have agriculture as their occupation. The present paper deals with the prediction based farming by using data mining method. This research focuses on evolution of a prediction model which may be used to predict sugar production. The proposed method use data mining technique to predict the sugar production based on the association rules and kmeans clustering.

The main objective of this paper is to study on Sugar prediction and farmers issues related to their sugarcane crop. Farmers can register their crop through paper or they will particular go to in sugar industry to registration. In our model, farmers can register to the system at any time and provide accurate information to industry. Through this information our model predict the sugar production in that industry and also help to harvesting team to make cluster data.

Key Words: Sugarcane crop information, Prediction analysis, Profit margins, Market price, Association rules, Kmeans clustering algorithm.

1. INTRODUCTION

Agriculture is very important because it produces food and feed which is necessity to animals and human beings. It fulfill the basic need of billions of people. It is one of the major contributor to the country's GDP and economic growth. Hence, it is widely practiced in India.

India is the second largest producer of sugarcane crop and produces about 20% of the world's sugarcane. With new regulatory policies and amendments in the older ones, there is a greater chance for farmers and millers to increase their efforts toward cane sugar production and processing. The sugar industry supports an estimated 12% of the rural population in the nine states of the country namely, Punjab, Uttar Pradesh, Maharashtra, Andhra Pradesh, Bihar, Gujarat, Haryana, Karnataka, and Tamil Nadu.

Normally, farmers can guess the final yield by their experience of growing particular crop again and

again. Farmers yield prediction accuracy is low and not cost effective. To meet the sugar requirements of the entire population of the country and for the export to other countries, it is important to practice modern methods of farming by using technology instead of practicing traditional farming methods. Modern methods allows the farmers to cultivate the crops in small area with minimum amount of water ,fertilizers and pesticides ,which finally produces good yield and profit to the farmers.

Our system is developed based on the data mining concepts to predict sugarcane yield. With the data obtained from industries, it is divided into sample and verification data sets. System is tested on verification data and the predicted values are compared to actual values. System takes current location of farm land, number of hectares of land and crop decided by farmer to grow as an input from the farmer in the web application which is an interface between farmer and the system. By the provided data and earlier trained data, the system is well understood to produce solution. Thus, algorithm analyzes the final values and predicts the yield per hectare and total values of cultivated crop based on the current market price.

2. PROBLEM STATEMENT

To Design, Develop and Implement the training model by using different inputs data. So system will able to learn the features and extract the sugarcane crop yield from the data by using data mining techniques.

3. OBJECTIVE

The proposed system aims at predicting or forecasting the crop yield by learning the past data of the farming land. Here we make use of different data mining techniques such as Association Rules, K-means clustering. Performance is evaluated based on predicted accuracy.



4. CHALLENGES

Some of the challenges faced during crop yield prediction are:

• Choosing appropriate dataset, after choosing dataset tuning of the parameters which makes project more efficient to get the desired results.

• Model must be trained by taking consideration of less computational efficiency and power.

• Increase of error rate due to dynamically changing the environment.

5. ARCHITECTURE

5.1 System Architecture

Training Phase

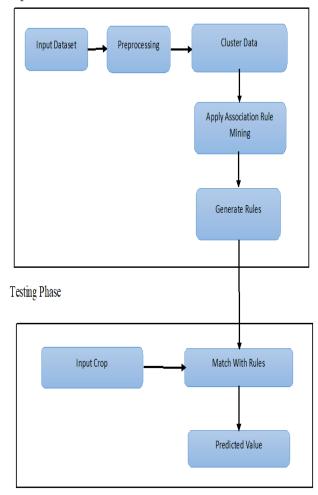


Diagram-1: Block Diagram of Proposed Work

5.2 Methodology

In our project includes following modules-

1. Farmer Module

In farmer module, farmer can register to the system. Through registration they can provide information to industry and also farmers enter to the system through username and password. After login they can see various options on the basis of information.

2. Admin Module

In this module, admin has the access to approval and disapproval which is controlled by the sugar industry. Staff admin can perform various operations like add, update, delete, etc.

3. Sugarcane Harvesting Module

In this module, the information of various harvesting tools are present like harvesting machine, tractors, teams etc. System add the various information related to harvesting tools and also view that information. Farmers crop information also viewed by this module and system make the cluster of same land location.

- Data Clustering: In clustering step, the preprocessed data was clustered using K-Means clustering algorithm.
- Association Rule Mining: Once the data has been collected and converted, association rule mining may begin. In this step, rules are created using frequent pattern mining. Association rules identify the relationships among a set of items or object in a database. The proposed method use data mining technique to predict the crop yield prediction based on the association rules.
- Crop Yield Prediction: Data Mining is widely applied to agricultural issues. Data mining is used to analyze large data sets and establish useful classifications and patterns in the data sets. The overall goal of the data mining is to extract the information from a data set and transform it into understandable structure for further use. This paper analyzes the crop yield production based on available data. The data mining technique was used to predict the crop yield for maximizing the crop productivity. Figure shows the flow of proposed crop yield prediction.



5.3 System working

- 1. The System takes input from user i.e. user location, sugarcane crop type, number of Hectares of land and also other information.
- 2. Based on the location and chosen crop, system takes previous years data from the repositories and analyzes the data and predicts the results.
- 3. The algorithm predicts the yield of crop and calculates the total weight of sugarcane crop in tons and generate sugar production in kilo.
- 4. Finally System produces the predicted output in the application.

6. IMPLEMENTATION

6.1 Experimental Setup

- Import the libraries.
- Import the data set.
- Get the basic information about the data set.
- And also need to purify the data set.
- Check for nulls in the data set.
- If any drop those rows.
- Check for any special characters in the columns.
- Just convert them in to nulls and drop those rows.
- Now need to select the features for the project.

• Taking all the column as input and sugarcane weight and sugar production as output

- Now need to split the data set into sample and verification data set.
- And finally using this trained model we can create an application to work.

• To take the inputs from the user and display the result.

6.2 Tools Used

Project Overview:

Front End: HTML, CSS, JavaScript

Back End: Java 1.8

Hardware Specification:

Processor: Intel CORE i3

RAM: 4GB

Hard Disk: 500 GB

Software Specification:

Software: Apache Tomcat 8.0.27.0

Operating System: Windows10

RESULTS



Fig -1: Home Page

REGISTRATION FORM

Name* :	Enter first name	Enter middle name	Enter last name
Gender* :	○ Male	• Female	
Address* :	Enter your current	address	
Mobile Number* :	Enter your mobile nur	nber	
Aadhar Number* :	Enter your aadhar nu	mber	
Email ID :	Enter correct Email-Id		
Survey Number* :	Enter your survey nur	nber	
Area* :	Enter area in hectares		
Sugarcane Type* :	Enter the current type	of sugarcane	
Cultivation Date* :	m/dd/yyyy		
Expected Harvesting Date*	: m/dd/yyyy		
BANK DETAILS :			
Bank Name* :	Please enter your bar	k name	
Account Number* :	Enter your account nu	mber	
Account Type* :	which account type is	that?	
IFSC Number* :	Enter IFSC number		
Username* :			
Password* :			
Re-enter Password* :			

Fig -2: Registration



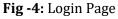
International Research Journal of Engineering and Technology (IRJET)e-ISSVolume: 08 Issue: 06 | June 2021www.irjet.netp-ISS

e-ISSN: 2395-0056 p-ISSN: 2395-0072

						ALL INFO	DEMATION OF E	ARMERS .					
~	Inter	Mails laws	La fan	Genter	later.	101	Intelle	Destr.	Grag Serier	4.00	igene (p	Caterio Das	lperior
	Dark .	Cambra dia	N.		AT GRADIERA DE LA SER	897534	875000	and pairs a	0		802	279.5	29.10
0	Ser.	functionals	14	-	44-bechilana.dekinez	454740	0000100	ortiptigesion			402	21945	
0	Sec.e.	Hotes	24		show wheeler	875+39	+5366-0	Sched Videolog	28	4	802	2152	2551
Di	400	No.	14	-	APOINT/PRODUCTION	81010	690000	and pright in	28		80	279-0	29-10
8	ike .	Dapar	14	-	AT GROWTH AND DR GROWT	01010	CREATER	derydl[prd.ce	128		812	2192	29-02
	-	Serie-	14		AT STRUCTURE DRIVING	PRICH	0760-0	modignics	28		802	219-5	290.8
Di	Ser.	Station	Name of Street S	-	44-beck-interdentier	0563347	363.629	addited to geni on	100		85	219-5	290.8
10	Sena	84	14	-	APOINT/Patient/Patient	5798124	BRO-DH	unitalgerian	12		802	2100	29-98
121	Berly	theoty	746	-	Alter Schele Service	10000	CREADER	netwoliged on		3	85	20102	2914
	144	theory .	14	-	AT 4 & Tathen Delicher	107702	C 100 C 100	utrystigention	e		-	20.00	2914
	inter .	Variation	24		sheelahma.lastaya	80.00	19-87610	winter informations		4	802	210.0	2445
2	Artist	No.	14	-	AF-HAShinadar-Atua	578-038	12-07612	universal grant can	-		802		
2	100	Owner	14	1.0	AF-tage Science Socienti	\$70×71	352-0100	another is a	18	20	80	21945	2942
	100	Owner	14		A Princips Tectional Descharged	100.041	87.04 DH	andalgedue	94	20	802	279.0	
Di	les .	later .	14		sifetings facilities for large	MARCH 1	+800 101	etterigenion	-D1	20	802	2165	2942
	Partner .	later .	14		ad tanga ha insa beringai	808.749	+849.70	preferencial preferences	124	20	402	2190	2912
2	Internet	Notice -	Debet	1.0	AF-targe Selfere De Grapi	(96.04	0100215	enter Odenium	100	-12	802	2140	2942
	5.04	24	14		AP-Date Translation Statistics	172423		sales all grad an	10		852	27.90	
	Derina .	Dava	N	10	AT COMP. Transfer Sectionary	100704	10(2)1	and the second s	(2)	20	-	2790	2912
	in the second se	4000	14		ad sizes have been best they	474707	10000	artiget (proton)	10		802	20.95	29.52
	Date:	Owter	24	1.00	Advisory Teconomic Servicement	679-057	48134412	bdeprigraine	10		80	21-95	2942
	140	tree .	14	-	Adversely Toward and Service Insur	407.04	4512411	entprigning	-			21940	2912
	Diff of	Test	24	10	AP-Orte Taxantel Services	877.0	05/2011	beloop digestion	101		852	2190	2910
	1010	No.	14		al size following and a size of the size o	8037104	7960-0	otherstignalism			802	219.3	29.52
9	ior.	Parks	Set a	-	show when better	40-01	+596-0X	reported project		4	802	2144	2942
	-	Vana	1418		whereas behavior (behavior	10,000	+59603	- desurt ad posicion			40	219.8	2942
0	104	205	5454	10	show branchester	10 13 10	00786005	abatathagealan	116	4	852	2791	29.00
	244	Takin .	late .	-	AT LONG THE PARTY OF LONG	10100	000000	enducted pairs			10	2104	
0		Date:			adding below lot they	10000	CHARLE	stratignal on	-01		-	2142	29.04
	0.01	(unit)	210	-	Address Tailance Devices	10010	62667	orthological or	19	-	40	2142	29-2-5
0	2144	Tourse .	210	1.0	APPENDER DEFINISION OF A Decision	10010	12428	protocol and			852	2742	2994
	Law .	Gent	14	10	AT the Terraria Terraria	1007.0	100900	percentigentian	102		-	2740	29.07
c .	Long Long	Ad	14		al tao la la tana Jan Kingar	180746	+00903	umterilprice	-		80	20.90	29.02
0	-	40.9	(hot	-	45-Dispalle-Data et Seriettepr	452471	10740	nimplegration		20	802		
6	4	42.9	Chait		44-bits brits at brits or	452-494	81185	structure	101	-	85	214.0	2942
	7645	Deleter	14	1.0	Afferia Tribund Defities	73e-C8	882.027	paragraphics	10		-	2194	291-5
	And	Delarry	14	-	Ministration Continue	The Cal	880.025	ant signal on	10	4	85	2194	2914
	ingen .	low			Mark School Deliter	1737-01	DOGGO	transferring and the second se				219.2	2912
	Debier	Line	244		views/binned/initial	10100	Deviden	menintelegences	-		85	216.0	29112
	And an	Seator Sector	100		(Hora billion Delites	61010	101100	and a construction	10		-	2101	29-94
	Test.	(and)	Cone .	-	Africa Tribust Deliter	107121	CREDON	patrice	-		-	2004	29-12
	in i		Calment .	-	stan Schen Setting	1000	OFDOS	statuted on				2102	29.04

Fig -3: Input Dataset Information





	welco	ome ganesh	patil Log O
	ADD CROP		
Farmer ID	Enter ID		
Area	Area		
Type Of Sugarcane	Type of sugarcane		
Total Weight	Total weight		
Date	mm/dd/yyyy		
Email ID	abc@gmail.com		
Mobile Number	Mobile number		
	Submit		

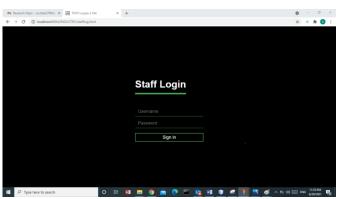


Fig -6: Admin/Staff Login

		2017-2018 INFORMATION E	ST DATE	
Sugarcase Type	Harvesing Date	Tand Area(Ja.419)	Tanal Weight(Je Jose)	Tend Ingor(3x Elic)
8005	2018-02-26	1	72	8640
	2015-16-01	5	216	
8005	2015.10.20	2	138	16260
BIRDS	2016-11-01	2	56	11520
	2015.11.15	3	168	
8005	2015-11-30	5	235	33120
BOOD	2018-12-01	8	384	4433
8605	2018-12-13		8	5750
8003	and the	38	1458	134360
80012	2018-02-26	3	144	17280
8/032	2018-06-25	2	%	11920
	2015-10-01	13	660	
80032	2015-16-10		48	5760
80032	2018-10-30		36	4120
	2018-10-31	5	240	
80032	2015-11-01	6	288	34560
80032	2018-11-30	3	148	20160
	2018-12-01	114	5484	658380
86032	2018-12-15	29	963	115200
80032	14	169	8124	914330
0		199	9982	1141640

Fig -7: Date wise Clustered Information

	2017-2018 1	FORMATION BY VILLAGE	TELEVISION DE	The second
Sugarcase Type	Allens	Tand Arms(Jacker)	Tanal Weight(In Stee)	Total Sugar(/w.Kiit)
hing.	1/P-Arie 3al-Pantala, Dist-Kalhapat	A. C.	12	56-40
94932	A.PA.No. Xal Psycholo, 25:54 & officient	1	~	11,520
		1 3	168	293.00
RARE	A P. Main, Tal-Pariheia, Jitos Kothapar	1	148	(728)
94832	A.W.Malu, Tai Panhala, Dini Kuthapat	10	34	42100
	A.P. Multi, Tail Parthela, Elini K. elbapar	19		129440
NH(D)	A P.K. Aidus, Eul-Pondulu Dess Ardiopur	1	139	18739
598.12	In F.K. Alabe, Sali Postkala 25 in Kolbepter	19	104	112229
	A P-K Alder, Sci-Pankola 2014 Andiopart	22	1102	.111140
KNOT	A P.K. Alder, Tal-particular, Dissi-Korthagene		104	48129
848.12	A P-R Aldu, Iai-pantula Don Koltupar	a.	214	20480
	A P-K-Mite: Tal-pathala.Disi-K-chapar	12	600	72900
89815	A/P Mangia Aid (Metala Det Sangal)	3	240	28800
54832	A.P.Mangla Tal.Shiradi.Hvit.Sangali	116	3585	ROMAND.
8	A W. Mungle, Sal. (Metala, Deat. Surgal)	(3)	1430	2/9K800
Relation (Control of Control of C	A (P. Kandas, Tal-Nudwerad, Dist-Keilioper		.79	13400
94832	A 'P Kandus Ta's Waltervall, Dist Kellisper		144	17200
1	A (P.Kanda), Ial Nudwead, Dist-Kelluper	4	214	29583
8001	A 95 Shinga Jal Shihuwaii Dist Kollogur	3	240	28600
98832	A #-Sitespackal-Baltgerad Dive Kolkapar	1	100	48139
	A W. Shinyu, Ed. Shihuwath J.Net & Ologue	12	176	##1.39

Fig -8: Location Wise Clustered Information

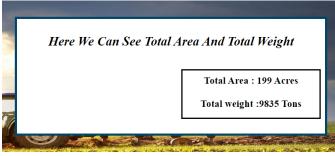
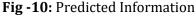


Fig -9: Actual Information

Total Area(In Acres)	Total Weight(In Tons)	Total Sugar(In Kilo
199	9582	1149840



International Research Journal of Engineering and Technology (IRJET)Volume: 08 Issue: 06 | June 2021www.irjet.net

7. CONCLUSIONS

Our system predicts the sugar production and total weight of sugarcane crop. Through this method we can predict all country sugar production and also it helpful for farmer and industry. Sugar industry can managed easily and farmers does not go in loss.

In future we add some extra features to our system like we list all the new and modern equipment used for cultivating the predicted best sugarcane crop.

ACKNOWLEDGEMENT

I mainly thanks to my guide **"Prof. M.M.Hajare"** who has provided the extensive guidance and expertise throughout the study. Thanks to all who helped we in complete this work.

REFERENCES

[1] D Ramesh, B Vishnu Vardhan, "Analysis of Crop Yield Production using Data Mining Techniques", International Journal of Research in Engineering and Technology(IJRET),Vol.4,2015.

[2] Dakshayini Patil, Dr.M.S.Shirdhonkar, "Rice Crop Yield Prediction Techniques: An Overview", International Journal of Advanced Research in Computer Science and Software Engineering, Volume 7, Issue 5, ISSN: 2277 128X, 2017.

[3] Dr. Rakesh Poonial, Sonia Bhargava "Prediction of Crops Methodology using Data Mining Techniques", International Research of Innovative Research in Science, Engineering and Technology, Vol.6, Issue 10,October 2017.

[4] Sujata R. and Issaki P (2016). A study on crop yield forecasting using classification Techniques, International Conference on Computing Technologies and Intelligent Data Engineering (ICCTIDE), pp.1-4.

[5] Sellam V. and Poovammal, E (2016). Prediction of Crop Yield using regression analysis. Indian J. Sci. Tech.9 (38):110-113.

[6] Aakunuri Manjula and Dr. G. Narsimha."Crop Yield Prediction with Aid of Optimal Neural Network in Spatial Data Mining: New Approaches", International Journal of Information & Computation Technology, ISSN 0974-2239 volume 6, Number1, pp. 25-33, 2016.

[7] B. Milovi and V. Radojevi,"APPLICATION OF DATA MINING IN AGRICULTURE", Bulgarian journal of agricultural science, 21(No 1)26-34 agricultural academy, 2015.

[8] R. S. Kodeeshwari, K. Tamil,"Different Types of Data Mining Techniques Used in Agriculture-A Survey

Т

"International Journal of Advanced Engineering Research and Science (IJAERS), vol-4, Issue-6s, Jun-2017.

[9] Suvidha Jambekar, Shikha Nema, Zia Saquib,"Data Mining Techniques for Prediction of Crop Production in India", International Journal of Innovations & Advancement in Computer.

[10] Dr. A.Senthil Kumar, P.Arun,"A survey on agriculture for crop yield prediction using data mining techniques", "IOSR journal of computer engineering (ISOR-JCE) p.no23- 25, vol 6, issue 3, March 2017.

[11] S.Veenadhari, D.Bharat Mishra, D C.Singh,: Data mining techniques are used in agriculture to study the data of weather and soil parameters. In this more emphasis is given to study the influence of climatic parameters on crop productivity using decision tree technique. The Decision trees are used for better understanding. This type of studies will help the researchers, policy makers and farmers in predicting crop yield in advance.

[12] N. Kumar, G, P.O. REDDY, S. CHATTERJI, D. SARKAR: Land Classification is described in detail in which soil units are created for effective utilization of farm land. In which planning makes easier and sustainable for agricultural practice and management. Machine learning systems are used in planning the use of farm land.

[13] M.S.Dahikar, D.V.Rode,: In this paper ,research is performed on climatic conditions of various parts of the world which directly affect the crop yield. This provides work on how to tackle the environmental conditions. Artificial neural networks have been used to study on the varying climate conditions day by day This work increases the effectiveness of the prediction solution.

[14] M. Kaur, H. Gulati, H. Kundra: For agriculture crop price analysis, Data mining techniques are very useful. This paper presents the applications and techniques of Data mining in agriculture. There are various data mining techniques such as K-Means, KNN, ANN and Support Vector Machines (SVM) which are used in yield and price prediction of the crops.



BIOGRAPHIES



Sucheta Patil. **BTech(Computer)** Kolhapur-416114.



Aishwarya Sawant. **BTech(Computer)** Kolhapur-416012.



Snehal Patil. BTech(Computer) Sangali-415408.



Rutuja Jadhav. **BTech(Computer)** Sangali-415408.



Mohini Chougule BTech(Computer) Kolhapur-416214.