Crop Prediction Framework Using Rough Set Theory

Hetal Patel#1, Dharmendra Patel#2

#Faculty of Computer Science and Applications,
Charotar University of Science and Technology (CHARUSAT), Changa - 388421, Gujarat, India
1hetalpatel.mca@charusat.ac.in
2dharmendrapatel.mca@charusat.ac.in

Abstract—**The agriculture sector contains the vast amount of data which require the development of specialized framework to store, clean, and analysis of the stored data to convert it into the knowledge such that hidden pattern can be identified from the data. Here, the basic concept of Rough Set Theory which is applied to the agriculture data set to make the decision. The Rough Set Theory (RS) offers a feasible approach for extraction of decision rules from data sets. These rules can be used for doing forecasting of crop-yield in the agriculture sector. In this paper, the RS framework ispresentedto generate the classification rules from 640 sets of agriculture data for crop forecasting.In proposed framework, the collected data are preprocess and then information table is generated. After this, decision table is generated. The reduction method is employed for finding out the reduct of the data set which holdsthe minimal subset of attributes accompanying with a class label. By applying the LEM2 algorithm, the rules are generated from the reduct. The study shows that the theory of rough sets is the one of the best technique for rule generation and decision making.**

Keyword- decision making, knowledge discovery, Rough Set Theory, crop-yield forecasting, rule generation and reduction, rule classification

I. INTRODUCTION

Now a day, the RST is applied in various domains, such asmachine learning [1], knowledge acquisition and knowledge discovery from database[2], decision analysis [3], expert systems[4], inductive reasoning and pattern recognition[5], data mining[6], and many more.The rough set methodology is applied to many applications like legal reasoning for drawing conclusion from the fact data, churn modeling in telecommunications and analysis of medical, finance and military dataset[7].

The central objective of the analysis using RST is to induce of (learning) approximations of concepts [8]. It gives mathematical tools to discover the hidden patterns in data. It can be used for data reduction[9], feature selection[10], feature and pattern extraction [11], decision rule generation [12]. Moreover, it can be employed to recognize partial or total dependencies in data,dynamic data,removes redundant data, missing data, give approach to null values , [13] and others.

The best of our knowledge, a very little work is done by employing rough set in agriculture sector. Therefore, we motivated to develop a framework by applying rough set in this domain. Advantages of employing this technique are explained as follows [14][15]:

- There is no prior or additional information about the data set is required
- It provides a valuable analysis
- It provides the interpretation in form of quantitative and qualitative data.

The main objectives of this study is to build an appropriate framework to access the performance of rough set classifiers, to do the forecasting of crop in the agriculture domain and to produce understandable decision rules to be applied on crop.

Use of Rough Set in Various Domains:

The RST has many properties which makes the one and only option for solving the various real problems like pattern recognition in which it is used for improvement in the classification ability of a hybrid pattern recognition system [16]. The designand development ofa mobile support system to triage abdominal pain in the emergency room of a hospital was done by the use of rough sets [17]. The rough sets concept is also applied to generalize the rules that explain the association between acoustical parameters ofconcert halls and sound processing algorithms [18]. The RST is employed to do the extractionof facts and rules for the power system operation [19].The hierarchical learning method based on RST is applied to the problem of sunspot classification from satellite images [20]. The author Shen and Jensen have identified the other area where rough set is successfully applied like prediction of business failure, financial investment, bioinformatics and medicine and fault diagnosis [21]. The rough set rules applied in forming the meta-structures of interest to semiconductor applications [22].

The paper is organized as follows: Sect. 2 describes the basics of RST and data mining with its applications; Sect. 3 presentsthe material and methodology; Sect. 4 describes the experimental results and discussion; Finally, Sect. 5 includes the conclusions.

II. OVERVIEW OF ROUGH SET THEORY

In 1992, the RST was initially proposed by ZdzislawPawlak [23]. The methodology of RST is deal with unclear or imperfect information and knowledge, analysis and classification vague, which is consider as non-statistical methods in data analysis.RST is a new method that deals with vagueness and uncertainty emphasized in decision making. This is the new technique to do analysis of the data. The advantages of RST to data analysis are as under[24]:

- It offers efficient algorithms which are able to find out the hidden patterns from the data
- It finds reduced sets of data so data reduction is done easily
- The significance of at the data is evaluated
- The minimal sets of decision rules are generated from data set
- It offers straightforward understanding of results
- The quantitative data analysis and the qualitative data analysis can be prepared
- It recognizesassociationswhich is not possible by applying statistical methods

The basic perception behind RST is the lower approximation and upper approximation of a set. The subset which is produced by lower approximations is the objects of interested subset. The subset produced by upper approximation is the objects which can possibly make a chunk of an interested subset. These subsets, defined by the lower approximation and upper approximation is known as Rough Set. The hidden knowledge in the systems can be discovered and expressed in the form of decision rules [25].

A. *Concept of RST*

The concept of Rough Set and Basic Terms used in this theory are discussed as follows:

A set is a collection of various objects of interest for instance collection of magazine, paintings, people etc. Suppose the given set of object O is a finite set of objects, called the universe. The relation R, R ⊆ O × O, is an indiscernibility relation which represent the lack of knowledge about the element of O. S is a subset of U.Now we are going to describe the set S with respect to R.

Definition 1: (Lower Approximation). The lower approximation of a set S with respect to relationR is the set of entirely facts that can be classified as S in view of the Relation R. Mathematically, it can be expressed as:

$$RL_*(s) = \bigcup_{x \in U} \{R(s) : R(s) \subseteq S\} \text{.............}(1)$$

Definition 2: (Upper Approximation).The upper approximation of a set S with respect to Relation R is the set of all facts which is certainly classified asS in view of the Relation R. Mathematically, it can be expressed as:

$$RU^*(s) = \bigcup_{x \in U} \{R(s) : R(s) \cap S \neq \emptyset\} \text{.............}(2)$$

Definition 3: (Boundary Region). A set of all the objects that is classified neither S nor not-X with respect to R of the boundary region of a set S with respect to R. It can be stated as:

$$R_B(s) = \bigcup_{x \in U} \{RL_*(s) - RU^*(s) \text{.............}(3)$$

In a simple word, we can say that granules of knowledge can be represented by the lower and upper approximation. The lower approximation of a set is union of all granules which are entirely included in the setwhereas the upper approximation is union of all granules which have non-empty intersection with the set.The difference between the upper and the lower approximation is the boundary region. This definition is representing in the Figure 1.

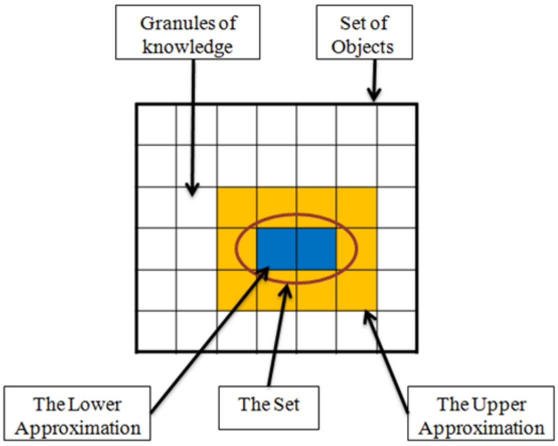

Fig. 1. Upper Approximation, Lower Approximation and Boundary Region

The Crisp Set and Rough Set can be expressed as under.

Definition 4: (Crisp Set).Set S is crisp (same with respect to R), if the boundary region of S is empty, i.e. $R_B(s) = \varnothing$ with respect to R.

Definition 5: (Rough Set).Set S is rough (inexact), if the boundary region of S is nonempty,
i.e. $R_B(s) \neq \varnothing$ with respect to R.

Definition 6: (Rough Membership Function). The rough membership function states the conditional probability of s which belongs to S certain in R and can be clarify like a degree at which s fits to S such that sexpressed by R.Below formula defines the Rough Membership Function.

$$\mu_S^R(s) = \frac{|S \cap R(s)|}{|R(s)|} \qquad \text{..................(4)}$$

where $\mu_S^R(s)$ denotes the cardinality or degree of certainty of S or a measure of significance.The meaning of rough membership function can be depicted as shown in Fig.2.

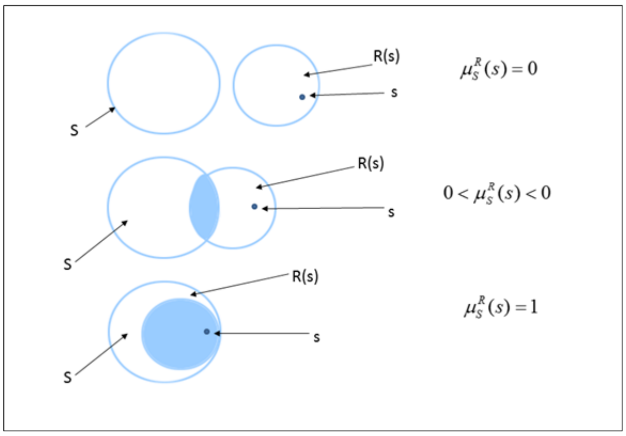

Fig.2. The Rough Membership function

Definition 7: (Degree of Consistency). The degree of consistency must be preserved while data reduction is being done. The degree of consistency of a table data is given below:

$$\mathrm{k} = \frac{\text{number of all consistent cases}}{\text{number of all cases}} \ldots\ldots\ldots\ldots (5)$$

Where, $0 \leq k \leq 1$.

B. Rough Set Attribute Reduction

In an information system, there may be a chance of some condition attributes that actually do not provideany additional information about the objects in O. So, it is required to remove thoseattributes. By doing this, we can reduce the complexity and cost of decision process.

Definition 8: (Reduct). It is a set of essential lowest amount of data, seeing as proprieties of the original data of information table are uphold. In other words, the reduct should be contain the ability of classifying objects, deprived of varying the form of representing the knowledge. Mathematically, it can be represented as $\gamma(J,K)=\gamma(R,D)$ where, J is variables of a set labelling D and $R \subseteq J$.

Definition 9: (Significance). The Significance of an attribute is calculated by determining the effect of removing the attribute from an information table. Suppose J and K are the sets of condition attributes and decision attributes respectively. The j is the condition attribute, i.e., $j \in J$.

The significant σs(J ,K)of set Sis calculated as shown in the below equation:

$$\sigma s(J,K)=\gamma(J,K)-\gamma(J-\{s\},K).$$

Definition 10: (Minimal Reduct).The minimal reduct isthe reduction of significance attributes induced by the reduction of the unnecessary and unused attributes without mislaying the information.

Definition 11: (Accuracy).The accuracy of the set S in RST calculated by the ratio of lower approximation and the upper approximation as shown in the below equation. In other word,

$$\alpha_s(X)=\frac{RL_*(s)}{RU^*(s)} \text{.............}(6)$$

III. Framework for Crop Prediction: The Rough Set Approach

A framework is presented below for data analysis using the rough set approach, in Figure 3. Each phase of the framework is explained next.

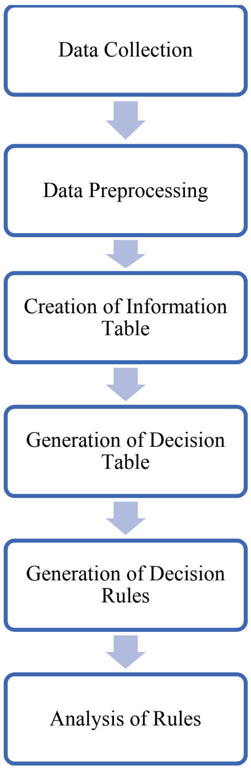

Fig. 3. Framework for data analysis using Rough Set

A. Phase 1. Collect the Data.

The dataset used in our experiment consists of 640 samples, collected from various sources. We have collected data from various government websites as shown in Table 1. Each sample data consists of three condition attributes or feature that represents its class which is wheat. The two class of each instance are either yes or no. If yes, then wheat is cultivated and no if wheat is not cultivated in the given conditions. In table 2, the class label c1 to c8 are attributes and c9 is the condition attributes.

Table I
Data Source

Sr. No	Name of the Organization	URL
1	National Portal of India	https://india.gov.in/topics/agriculture/crops
2	Farmer's Portal	http://farmer.gov.in/FarmerHome.aspx
3	Agriculture & Cooperation Department, Government of Gujarat	https://ikhedut.gujarat.gov.in/
4	Accuweather Website	http://www.accuweather.com/en/in/anand/188164/weather-forecast/188164
5	Anand Agriculture University	http://shc.aau.in/home/soil
6	Agriculture Portal, Tamil Nadu	http://agritech.tnau.ac.in/agriculture/crop_production_varieties.html
7	ENVIS Centre : Status of Environment & Related Issues - Gujarat	http://www.gujenvis.nic.in/PDF/soil.pdf
8	Government of India, Department of Fertilizer, Ministry of Chemical & Fertilizer	http://fert.nic.in/node/1452
9	National Food Security Mission, Ministry of Agriculture & Farmer Welfare, Government of India	http://nfsm.gov.in/

Table II
Condition and Decision Attributes of Crop Dataset

Label	Attribute	Domain
C1	Season	Rabi, Kharif
C2	Soil	Light clay, Sand, Heavy loam, Sandy Loam, Silt loam to clay, Sandy clay loam to clay
C3	Soil fertility	Low, Medium, High, Average
C4	Weather	Hot, Rainy, Windy, Cloudy
C5	Wind Speed	Low, Medium, High
C6	Water Source	Well, Canal, irrigation
C7	Fertilizer	Liquid, Granular
C8	Pesticides	Solid, Liquid
d=C9	Class	Wheat or Not

B. Phase 2. Data Preprocessing

Data preprocessing is very important phase as it removes those attributes which are not important or give significant role for the prediction. The incomplete data are removed from the data set.

C. Phase 3. Creation of Information Table

Knowledge representation using rough sets is done by information table. The sample data set of Information table is as shown in Table 3.

Table III
Information System for Crop Dataset

Crop	Atr1	Atr2	Atr3	Atr4	Atr5	Atr6	Atr7	Atr8	Decision Class
s1	Silt loam to clay	medium	cloudy	low	well	g	s	rabi	Yes
s2	Silt loam to clay	medium	cloudy	low	well	g	s	rabi	No
s3	Sandy clay loam to clay	High	cloudy	Medium	well	g	s	rabi	Yes
s4	Sandy loam to sandy clay loam	medium	windy	low	canal	g	s	rabi	Yes
s5	Sandy loam to sandy clay loam	High	windy	Medium	canal	g	s	rabi	Yes
s6	Silt loam to clay	High	windy	low	well	g	s	rabi	Yes

D. Phase 4. Creation of Decision Table

The two types of attributes, which are known as condition and decision, are exists in the information table. If they are extricate then the new table is known as decision table. Each individual row of such table is known as decision rules based on that the decision can be taken. The table 4 is the decision table [26].

Table IV
Decision Table Derived from Table III

Crop	Atr1	Atr 2	Atr 3	Atr 4	Atr5	Atr6	Atr7	Atr8	Decision Class
s1	Silt loam to clay	medium	cloudy	low	well	g	s	rabi	Yes
s2	Sandy clay loam to clay	High	cloudy	Medium	well	g	s	rabi	Yes
s3	Sandy loam to sandy clay loam	medium	windy	low	canal	g	s	rabi	Yes
s4	Sandy loam to sandy clay loam	High	windy	Medium	canal	g	s	rabi	Yes
s5	Silt loam to clay	High	windy	low	well	g	s	rabi	Yes

E. Phase 5. Generation of Decision Rule.

The approximations are very useful to draw the conclusion from the data. The relationship we have found between the condition attributes are {s1, s3, s4, s6} and {s2,s5}. In our example we have, with respect to the condition attributes, following facts:

1. Lower Approximation

The set of fact {s1,s2, s3,s4, s5} is certainly classified as wheat can be cultivated with the given condition. The set {s1, s3, s4} is the lower approximation of the set {s1, s2, s3, s4, s5}.

2. Upper Approximation

The set of fact {s1,s3, s4, s5, s6} is possibly cab be classified as wheat can be cultivate with the given condition. The set {s1, s2, s3, s5} is the upper approximation of the set {s1, s2, s3, s4, s5}.

3. Boundary Region

The set of fact {s2} is classified as neither as wheat nor no wheat (boundary region). The set {s4, s5} is the boundary region of the set {s1, s2, s3, s4, s5}.

4. Decision Rule

It is required to reduct the data for making the decision rule. Below step describes how to create reduct from information table.

Step 1: Verification inconclusive data

The crop data s5 and s6 are excluded as they hold equal values of conditions attributes with a value of decision attribute that is different.

Step 2: Verification of equivalent information

Thereis no data exist in the table 3 that possess equivalent information. The reduct of information table is as under.

Table 5
Reduct of Information Table

Crop	Atr1	Atr 2	Atr 3	Atr 4	Atr5	Atr6	Decision Class
s1	Silt loam to clay	Medium	cloudy	low	well	rabi	Yes
s2	Sandy clay loam to clay	High	cloudy	Medium	well	rabi	Yes
s3	Sandy loam to sandy clay loam	Medium	windy	low	canal	rabi	Yes
s4	Sandy loam to sandy clay loam	High	windy	Medium	canal	rabi	Yes
s5	Silt loam to clay	High	windy	low	well	rabi	Yes

IV. EXPERIMENT AND RESULT

Here in this research, the LEm2 algorithm is apply to the reduct data set which generated the 26 significant decision rules as shown in the following figure 4.

By using the information reduct shown above, the necessary decision rules are generated by applying the LEM2 algorithm for crop prediction. The obtained significant rules are presented as under.

```
A set consisting of  26  rules:
1. IF SoilFertility is High and SoilType is clay and WindSpeed is low THEN  is Yes;
                (supportSize=87; laplace=0.98876404494382)
2. IF WaterSource is well and SoilFertility is medium and SoilType is Sandy loam to sandy clay loam THEN  is Yes;
                (supportSize=16; laplace=0.944444444444444)
3. IF SoilFertility is medium and SoilType is Sandy clay loam to clay THEN  is Yes;
                (supportSize=19; laplace=0.952380952380952)
4. IF Weather is windy and WindSpeed is Medium and SoilType is clay THEN  is Yes;
                (supportSize=10; laplace=0.916666666666667)
5. IF SoilFertility is High and WaterSource is well and SoilType is Clay loam to clay and WindSpeed is low THEN  is Yes;
                (supportSize=6; laplace=0.875)
6. IF WindSpeed is High and SoilType is clay THEN  is Yes;
                (supportSize=14; laplace=0.9375)
7. IF Weather is cloudy and WaterSource is canal and SoilFertility is medium THEN  is Yes;
                (supportSize=10; laplace=0.916666666666667)
8. IF WindSpeed is Medium and WaterSource is well and Weather is cloudy THEN  is Yes;
                (supportSize=12; laplace=0.928571428571429)
9. IF Weather is windy and SoilFertility is High and WindSpeed is Medium and SoilType is Clay loam to clay THEN  is Yes;
                (supportSize=9; laplace=0.909090909090909)
10. IF Weather is windy and WaterSource is well and WindSpeed is High and SoilType is Silt loam to clay THEN  is Yes;
                (supportSize=6; laplace=0.875)
11. IF SoilType is Sandy loam to sandy clay loam and WaterSource is well and WindSpeed is Medium THEN  is Yes;
                (supportSize=8; laplace=0.9)
12. IF Weather is cloudy and WindSpeed is low and SoilType is clay THEN  is Yes;
                (supportSize=35; laplace=0.972972972972973)
13. IF SoilFertility is High and WindSpeed is High and SoilType is Sandy loam to sandy clay loam THEN  is Yes;
                (supportSize=4; laplace=0.833333333333333)
14. IF Weather is cloudy and SoilType is Sandy loam to sandy clay loam THEN  is Yes;
                (supportSize=8; laplace=0.9)
15. IF WaterSource is canal and Weather is cloudy and SoilType is Silt loam to clay THEN  is Yes;
                (supportSize=10; laplace=0.916666666666667)
16. IF Weather is windy and SoilType is Clay loam to clay and WindSpeed is High and WaterSource is canal THEN  is Yes;
                (supportSize=2; laplace=0.75)
17. IF WaterSource is well and SoilType is Clay loam to clay and WindSpeed is Medium THEN  is Yes;
                (supportSize=5; laplace=0.857142857142857)
18. IF WaterSource is well and SoilType is Clay loam to clay and Weather is cloudy THEN  is Yes;
                (supportSize=4; laplace=0.833333333333333)
19. IF WaterSource is well and SoilType is clay THEN  is Yes;
                (supportSize=45; laplace=0.978723404255319)
20. IF Weather is windy and SoilType is Sandy clay loam to clay and WindSpeed is low THEN  is Yes;
                (supportSize=12; laplace=0.928571428571429)
21. IF SoilFertility is Medium and SoilType is clay and WindSpeed is Medium THEN  is Yes;
                (supportSize=2; laplace=0.75)
22. IF WindSpeed is High and SoilFertility is Medium THEN  is Yes;
                (supportSize=2; laplace=0.75)
23. IF SeasonType is kharif THEN  is No;
                (supportSize=199; laplace=0.995024875621891)
24. IF Weather is windy and SoilType is Clay loam to clay and SoilFertility is Medium and WaterSource is canal THEN  is No;
                (supportSize=3; laplace=0.8)
25. IF SoilFertility is High and Weather is windy and WindSpeed is Medium and SoilType is Sandy clay loam to clay THEN  is No;
                (supportSize=1; laplace=0.666666666666667)
26. IF WindSpeed is High and Weather is cloudy and SoilType is Clay loam to clay THEN  is No;
                (supportSize=1; laplace=0.666666666666667)
```

Fig. 5. Rules Generated using LEM2 Algorithm

The laplace of the rules are obtained which are shown as below:

```
##    Rule_1    Rule_2    Rule_3    Rule_4    Rule_5    Rule_6    Rule_7
## 0.9887640 0.9444444 0.9523810 0.9166667 0.8750000 0.9375000 0.9166667
##    Rule_8    Rule_9   Rule_10   Rule_11   Rule_12   Rule_13   Rule_14
## 0.9285714 0.9090909 0.8750000 0.9000000 0.9729730 0.8333333 0.9000000
##   Rule_15   Rule_16   Rule_17   Rule_18   Rule_19   Rule_20   Rule_21
## 0.9166667 0.7500000 0.8571429 0.8333333 0.9787234 0.9285714 0.7500000
##   Rule_22   Rule_23   Rule_24   Rule_25   Rule_26
## 0.7500000 0.9950249 0.8000000 0.6666667 0.6666667
```

The support of a rule is indicates that how often theantecedent and the consequent of a rule appear together in the transaction. Theconfidence of a rule indicates that how often the antecedent and the consequent exist together. The support of the rules is obtained which are shown as below:

```
##    Rule_1    Rule_2    Rule_3    Rule_4    Rule_5    Rule_6    Rule_7
## 0.1359375 0.0250000 0.0296875 0.0156250 0.0093750 0.0218750 0.0156250
##    Rule_8    Rule_9   Rule_10   Rule_11   Rule_12   Rule_13   Rule_14
## 0.0187500 0.0140625 0.0093750 0.0125000 0.0546875 0.0062500 0.0125000
##   Rule_15   Rule_16   Rule_17   Rule_18   Rule_19   Rule_20   Rule_21
## 0.0156250 0.0031250 0.0078125 0.0062500 0.0703125 0.0187500 0.0031250
##   Rule_22   Rule_23   Rule_24   Rule_25   Rule_26
## 0.0031250 0.3109375 0.0046875 0.0015625 0.0015625
```

The confidence of the rules is obtained which are shown as below:

```
## Rule_1 Rule_2 Rule_3 Rule_4 Rule_5 Rule_6 Rule_7 Rule_8 Rule_9
##      1      1      1      1      1      1      1      1      1
## Rule_10 Rule_11 Rule_12 Rule_13 Rule_14 Rule_15 Rule_16 Rule_17 Rule_18
##      1      1      1      1      1      1      1      1      1
## Rule_19 Rule_20 Rule_21 Rule_22 Rule_23 Rule_24 Rule_25 Rule_26
##      1      1      1      1      1      1      1      1
```

The lift of the rules is obtained which is shown as below:

```
##   Rule_1   Rule_2   Rule_3   Rule_4   Rule_5   Rule_6   Rule_7   Rule_8
## 1.615233 1.542833 1.555798 1.497455 1.429389 1.531489 1.497455 1.516903
##   Rule_9  Rule_10  Rule_11  Rule_12  Rule_13  Rule_14  Rule_15  Rule_16
## 1.485080 1.429389 1.470229 1.589437 1.361323 1.470229 1.497455 1.225191
##  Rule_17  Rule_18  Rule_19  Rule_20  Rule_21  Rule_22  Rule_23  Rule_24
## 1.400218 1.361323 1.598831 1.516903 1.225191 1.225191 2.565486 2.062651
##  Rule_25  Rule_26
## 1.718876 1.718876
```

V. CONCLUSION

In this research, RST is used to create the decision rules which are very useful for the prediction of the crop. The LEM2 algorithm of RST is applied as it provides the efficient result in comparison of association rule mining algorithm. The experiment is conducted on 640 dataset collected from various sources. The LEM2 algorithm is applied to on reduct dataset which is generated from the decision table. As a result, 26 significant rules are generated by the algorithm. These rules are used to forecast that whether wheat should be cultivated or not on the given condition. For instance, the rules tell that when the season is kharif then wheat cannot cultivate. When the wind speed is high, then the wheat crop will not give the best return of the investment. The framework can be extended for the various types of crop like flowers, vegetables, cereals, pulses and many more.

REFERENCES

[1] Wenshan, W., &Haihua, L. (2010*). Machine Learning Applications in Rough Set Theory.* In 2010 International Conference on Internet Technology and Applications.
[2] Leung, Y., Wu, W. Z., & Zhang, W. X. (2006). Knowledge acquisition in incomplete information systems: a rough set approach. European Journal of Operational Research, 168(1), 164-180.
[3] Greco, S., Matarazzo, B., &Slowinski, R. (2001). Rough sets theory for multicriteria decision analysis. European journal of operational research, 129(1), 1-47.
[4] Tsumoto, S. (1998). Automated extraction of medical expert system rules from clinical databases based on rough set theory. Information sciences, 112(1), 67-84.
[5] Pawlak, Z. (2012). Rough sets: Theoretical aspects of reasoning about data (Vol. 9). Springer Science & Business Media.
[6] Chan, C. C. (1998). A rough set approach to attribute generalization in data mining. Information Sciences, 107(1), 169-176.
[7] Pawlak, Z. (2002). Rough set theory and its applications. Journal of Telecommunications and information technology, 7-10.
[8] Yao, Y. (2008). Probabilistic rough set approximations. International journal of approximate reasoning, 49(2), 255-271.
[9] Zhu, W., & Wang, F. Y. (2003). Reduction and axiomization of covering generalized rough sets. Information sciences, 152, 217-230.
[10] Swiniarski, R. W., &Skowron, A. (2003). Rough set methods in feature selection and recognition. Pattern recognition letters, 24(6), 833-849.
[11] Zhai, L. Y., Khoo, L. P., &Fok, S. C. (2002). Feature extraction using rough set theory and genetic algorithms—an application for the simplification of product quality evaluation. Computers & Industrial Engineering, 43(4), 661-676.
[12] Stefanowski, J. (1998). On rough set based approaches to induction of decision rules. Rough sets in knowledge discovery, 1(1), 500-529.
[13] Qian, Y., Liang, J., Pedrycz, W., & Dang, C. (2011). An efficient accelerator for attribute reduction from incomplete data in rough set framework. Pattern Recognition, 44(8), 1658-1670.
[14] Chen, Y. S., & Cheng, C. H. (2013). Application of rough set classifiers for determining hemodialysis adequacy in ESRD patients. Knowledge and information systems, 34(2), 453-482.
[15] Tay, F. E., & Shen, L. (2002). Economic and financial prediction using rough sets model. European Journal of Operational Research, 141(3), 641-659.
[16] Cyran, K. A., &Mrozek, A. (2001). Rough sets in hybrid methods for pattern recognition. International Journal of Intelligent Systems, 16(2), 149-168.
[17] Michalowski, W., Rubin, S., Slowinski, R., & Wilk, S. (2003). Mobile clinical support system for pediatric emergencies. Decision Support Systems, 36(2), 161-1
[18] Kostek, B. (1999). Assessment of Concert Hall Acoustics using Rough Set and Fuzzy Set Approach, in:Soft computing in acoustics: applications of neural networks, fuzzy logic and rough sets to musical acoustics, Pal, S. &Skowron, A. (Ed.), pp. 381-396, Springer-Verlag Co., ISBN 981-4021-00-8, Secaucus- USA.
[19] Lambert-Torres, G. (2002). Application of rough sets in power system control center data mining. In Power Engineering Society Winter Meeting, 2002. IEEE (Vol. 1, pp. 627-631). IEEE.
[20] Nguyen, S. H., Nguyen, T. T., & Nguyen, H. S. (2005, August). Rough set approach to sunspot classification problem. In International Workshop on Rough Sets, Fuzzy Sets, Data Mining, and Granular-Soft Computing (pp. 263-272). Springer Berlin Heidelberg.
[21] Shen, Q., & Jensen, R. (2007). Rough sets, their extensions and applications. International Journal of Automation and Computing, 4(3), 217-228.

[22] Kusiak, A. (2001). Rough set theory: a data mining tool for semiconductor manufacturing. IEEE transactions on electronics packaging manufacturing, 24(1), 44-50.
[23] Pawlak, Z. (1982). Rough sets. International Journal of Computer & Information Sciences, 11(5), 341-356.
[24] Pawlak, Z. (1997). Rough sets and data mining. In Proceedings of the Australiasia-Pacific Forum on.
[25] Rissino, S., & Lambert-Torres, G. (2009). Rough set theory–fundamental concepts, principals, data extraction, and applications. Data mining and knowledge discovery in real life applications, 438.
[26] Ziarko, W. (2002). Rough set approaches for discovery of rules and attribute dependencies. Handbook of data mining and knowledge discovery, 328-338.