Scope
Volume 13 Number 02 June 2023

In Agrarian Realms Selecting Crops Using Rough Set Based Decision Techniques

Vairal Kailas L

Padmashri Vikhe Patil College of Arts, Science and Commerce, Pravaranagar,
Loni-413713, Ahmednagar, Maharashtra, India. Affiliated to Savitribai Phule, Pune University, Pune

Abstract:
Rough Set Theory (RST) is innovative mathematical techniques for dealing with inexact, imprecise knowledge. RST was delivers powerful several interesting applications that can be used for a wide range of purposes. This paper studies strength, certainty factor, and coverage factor of the decision rule signifies a though flow the relevant branch. In this article used rough set based techniques in agrarian fields. A proposed task we sort out better sugarcane varieties for planting using rough set theory. Certainty factor of the decision rules and coverage factors of the decision rules to find out varieties. Using RST the outcome we got convinces that the agrarians used sugarcane varieties V_3, V_{11}, and V_6.

Keywords: Rough Set, Decision Rule, Strength, Certainty Factor, Coverage Factor.

Introduction

Z. Pawlak, proposed (Rough Set Theory) RST in 1982 as one of them. In the analysis of classification datasets, this way of addressing rough sets is crucial. By assessing whether all objects with the same features described by attributes are categorised into the same class or not, we may discern between credible, consistently categorised data and dubious, contradicting data. We can get the bare minimum of attributes by decreasing attributes that describe the feature of groups of objects while keeping all creditable data. By lowering the amount of trustworthy data described in a class while maintaining classification accuracy. Rough sets representation of ability constraint, human thinking in the face of uncertainty, is incomprehensible to classical logic. Non-classical logics including modal logic, many-valued logic, intuitionistic logic, and para consistent logic have been explored and developed since Aristotle's time. Rough set theory is investigated from the perspectives of algebras and non-classical logic. Furthermore, the linkages between non-monotonic reasoning, association rules in conditional logic, and background information were investigated using a granularity-based reasoning framework, which is a wide approach to reasoning with rough sets. It's a variant of (ordinary) set theory in which a slice of a universe is formalised by two sets, the LA and UA. These approximations can be described by two operators on subgroups of the field. Recognize that an equivalency relation plays key role in RST. This approximation is frequently used to represent data that is incomplete. Certainly, relations other than equivalence relations can be used to create RST using an equivalence relation, on the other hand, allows for a more elegant formalisation and simple applications. Following Pawlak's work, multiple variations of RST have been established in the literature, each based on a different set of relationships. Rough set theory is particularly beneficial for extracting knowledge from data tables, and it has been successfully used to a variety of disciplines, including data analysis, decision making, and machine learning. In actuality, rough set and non-classical logic, particularly modal logic, have a variety of links. Great deal of effort has gone into laying a logical framework for rough set theory. Orlowska [5] devised a Logic for thinking approximately concepts constructed on RST in 1980s, effectively the modal logic S5. Yao and Lin [18] developed RST using modal logic and Kripke semantics. RST is among the most widely used models for reasoning from ambiguous and imprecise data. It also has something to do with granular computing. In reality, rough set theory raises a slew of challenges involving numerous sorts of reasoning. Pawlak presented an information system in Pawlak [13] in 1981, observed as a predecessor of RST, because it shares many principles with it. Pawlak created the rough set notion in 1982 to react with arguments based on erroneous facts [14]. Pawlak [15] gathered his contributions in a monograph released in 1991. Pawlak's main focus is a rigorous categorization of knowledge. Rough set theory and knowledge logics are inextricably intertwined. Orlowska [5] in 1988 investigated logical features of learning concepts. The research broadens the scope of rough set theory to include "probabilistic data" or "inconsistent data". In Yao and Lin [18], in 1996, Kripke models were used to study the association amongst the generic RS model, and modal logics. Their research revealed that opportunity and higher extrapolation in rough sets are closely related. Because both theories can manage vagueness, it is appropriate to investigate combining RST and FST. In 1989, Dubois and Prade [3] defined the distinctions between FST and RST. The former utilises UA and LA on fuzzy sets, whereas the latter fuzzifies an equivalence relation. We could pick one of them depending on the application. In 1996, Pagliani proposed using Nelson algebras as a basis for RST; Pagliani [10]. Duntsch [4] created a rough set logic for the first time in 1997. Based on his discoveries, Pomykala proposed a probabilistic logic for rough sets with such an arithmetic interpretation following established twofold Stone algebras. Pomykala and Pomykala [17] demonstrated that a regular double Stonealge brain 1998 is formed by the gathering RS of an assessment universe. The statement that a Boolean algebra is formed by the collection of all subsets of a set and that its logic is equivalent to classical propositional logic is well-known. RST can be utilised as a semantic foundation in non-classical logics. For instance, Akama and Murai [2] developed a crude set semantics for numerous three-valued logics in 2005. A table-based information system, such as a relational database, is the other, work created multi-rough sets, which are rough sets that have been generalised. As a continuation of Duntsch's rudimentary set logic,Akama et al. presented a Heyting-Brouwer rough set logic in 2013 [1]. Because it includes an implication, the logic is beneficial for reasoning about ambiguous data. According to Akama et al. [2], its subsystem can also be used as ambiguous logic. Rough set theory has been used to solve a wide range of problems, and there is a substantial amount of literature on the subject, RST was the handling imprecision (Pawlak, 2004). The Latin word for this concept, the tertiumnon datur principle of old logic, defines vagueness; the current version of this concept is referred to as the law of excluded middle. Rough sets, in this sense, express ambiguous notions in intermediate sense. Orlowska introducing several information relations, Nondeterministic information systems [6,7,8]. As a result, indiscernibility relations determined by these types of information systems are identity relations. Comprehensive information structures are the foundations for generating perfect understanding, which is depicted by identity interactions, because wisdom is built on the differentiating among things, according to Pawlak [14,15,16].

Knowledge Systems and Decision Rule (DR)

Apiece DT(Decision Table)depiction judgements (activities, outcomes etc.) that are taken when certain circumstances are met. We describe decision rule as follows

A decision table is stated as, $S = (U, G, N)$, $\forall$ e $\in U$ finding a string of $g_1(e)$, $g_2(e)$,…,$g_r(e)$.$n_1(e)$, $n_2(e)$,…,$n_d(e)$. The above sequence known as, DR stimulated by e $\in S$, it signify $G \rightarrow_e N$.
The numeral

$supp_e\,(G,N) = |\,G(e) \cap D(e)\,|$

is known as DR on support, $G \rightarrow_e N$ and the

$\sigma_e(G, N) = \frac{supp_e\,(G,N)}{|U|}$

represent as power on DR $P \rightarrow_x N$ we indicate the CF of the DR as

$cert_e\,(G,D) = \frac{supp_e\,(G,N)}{|G(e)|}$,

and we use CFs of the DR detailed as
$cov_e\,(G,D) = \frac{supp_e\,(G,N)}{|N(e)|}$

Decision Tables and Flow Graphs

We allocate a focussed division "e" involving the involvement node G(e) and outcomes node N(e) to every DT, which is a engaged, linked, acyclic-graph constructed as below: with each DR $G \rightarrow_e N$. Each decision rule's strength, certainty, and coverage indicate the thought flow of the relevant branch.

Decision Table:

The data are collected from few farmers, asked "Which quality of sugarcane crop varieties better for plantings?" The data is given as a table, with columns labelled by Attributes and rows marked by varieties like as Co09002, Co09003, Co09004, Co09005, Co09006, Co09007, CoN09071, CoN09072, Co850004, Co94008, CoC671 say as V_1,V_2,V_3, V_4, V_5, V_6,V_7, V_8,V_9, V_{10}, V_{11}. In the table containing information about sugarcane varieties. The following attributes sugarcane yields (t/ha), CCS yields (t/ha), CCS%, sucrose%, Brix%, Purity%, and 5 cane weights. Information system presented data about sugarcane as shown in below table.

Varieties	Better Varieties	Age of Farmers	Sugarcane Yields	CCS %	Brix %	5 Cane Weights	Quality of Varieties	Support
V_1	*No*	*Old*	*M*	*H*	*H*	*H*	*NVG*	*75*
V_2	*No*	*Middle*	*M*	*H*	*VH*	*H*	*NVG*	*75*
V_3	*Yes*	*Old*	*VH*	*VH*	*H*	*VH*	*E*	*320*
V_4	*Yes*	*Middle*	*M*	*H*	*H*	*H*	*VG*	*90*
V_5	*No*	*Young*	*H*	*H*	*H*	*L*	*NE*	*20*
V_6	*Yes*	*Middle*	*VH*	*H*	*H*	*VH*	*E*	*240*
V_7	*No*	*Young*	*L*	*M*	*H*	*H*	*NVG*	*70*
V_8	*No*	*Old*	*H*	*M*	*H*	*M*	*NG*	*30*
V_9	*Yes*	*Young*	*M*	*M*	*H*	*H*	*VG*	*90*
V_{10}	*Yes*	*Old*	*H*	*H*	*VH*	*H*	*VG*	*90*
V_{11}	*Yes*	*Middle*	*H*	*VH*	*H*	*VH*	*E*	*300*

Decision Algorithm:

The following decision algorithm associated with above table

a) *if(Sugarcane parameter of better varieties(Sugarcane yields-Moderate, CCS% - Highest, Brix % - Highest, 5 Cane weights - Highest), No) and (age, old man) then (Quality of Varieties , NVG).*

b) *if(Sugarcane parameter of better varieties(Sugarcane yields-Moderate, CCS% - Highest, Brix % - Very Highest, 5 Cane weights - Highest), No) and (age, middle man) then (Quality of Varieties , NVG).*

c) *if(Sugarcane parameter of better varieties(Sugarcane yields-Very Highest, CCS% - Very Highest, Brix % - Highest, 5 Cane weights - Very Highest), Yes) and (age, old man) then (Quality of Varieties , E).*

d) *if(Sugarcane parameter of better varieties(Sugarcane yields-Moderate, CCS% - Highest, Brix % - Highest, 5 Cane weights - Highest), Yes) and (age, middle man) then (Quality of Varieties , VG).*

e) *if(Sugarcane parameter of better varieties(Sugarcane yields-Highest, CCS% - Highest, Brix % - Highest, 5 Cane weights - Low), No) and (age, young man) then (Quality of Varieties , NE).*

f) *if(Sugarcane parameter of better varieties(Sugarcane yields-Very Highest, CCS% - Highest, Brix % - Highest, 5 Cane weights - Very Highest), Yes) and (age, middle man) then (Quality of Varieties , E).*

g) *if(Sugarcane parameter of better varieties(Sugarcane yields-Low, CCS% - Moderate, Brix % - Highest, 5 Cane weights - Highest), No) and (age, young man) then (Quality of Varieties , NVG).*

h) *if(Sugarcane parameter of better varieties(Sugarcane yields-Highest, CCS% - Moderate, Brix % - Highest, 5 Cane weights - Moderate), No) and (age, old man) then (Quality of Varieties , NG).*

i) *if(Sugarcane parameter of better varieties(Sugarcane yields-Moderate, CCS% - Moderate, Brix % - Highest, 5 Cane weights - Highest), Yes) and (age, young man) then (Quality of Varieties , VG).*

j) *if(Sugarcane parameter of better varieties(Sugarcane yields-Highest, CCS% - Highest, Brix % - Very Highest, 5 Cane weights - Highest), Yes) and (age, old man) then (Quality of Varieties , VG).*

k) *if(Sugarcane parameter of better varieties(Sugarcane yields-Highest, CCS% - Very Highest, Brix % - Highest, 5 Cane weights – Very Highest), Yes) and (age, middle man) then (Quality of Varieties , E).*

Inverse Decision Algorithm:

The following decision algorithm associated with above table

a) *if(Quality of Varieties , NVG) then (Sugarcane parameter of better varieties(Sugarcane yields-Moderate, CCS% - Highest, Brix % - Highest, 5 Cane weights - Highest), No) and (age, old man).*

b) *if(Quality of Varieties , NVG) then (Sugarcane parameter of better varieties(Sugarcane yields-Moderate, CCS% - Highest, Brix % - Very Highest, 5 Cane weights - Highest), No) and (age, middle man).*

c) *if (Quality of Varieties, E) then (Sugarcane parameter of better varieties(Sugarcane yields-Very Highest, CCS% - Highest, Brix % - Very Highest, 5 Cane weights - Highest), Yes) and (age, old man).*

d) *if (Quality of Varieties, VG) then (Sugarcane parameter of better varieties(Sugarcane yields-Moderate, CCS% - Highest, Brix % - Highest, 5 Cane weights - Highest), Yes) and (age, middle man).*

e) *if (Quality of Varieties, NE) then (Sugarcane parameter of better varieties(Sugarcane yields-Highest, CCS% - Highest, Brix % - Highest, 5 Cane weights - Low), No) and (age, young man).*

f) *if(Quality of Varieties , E) then (Sugarcane parameter of better varieties(Sugarcane yields-Very Highest, CCS% - Highest, Brix % - Highest, 5 Cane weights-Very Highest), Yes) and (age, middle man).*

g) *if(Quality of Varieties , NVG) then (Sugarcane parameter of better varieties(Sugarcane yields-Low, CCS% - Moderate, Brix % - Highest, 5 Cane weights - Highest), No) and (age, young man).*

h) *if(Quality of Varieties, NG) then (Sugarcane parameter of better varieties(Sugarcane yields-Highest, CCS% - Moderate, Brix % - Highest, 5 Cane weights - Moderate), No) and (age, old man).*

i) *if(Quality of Varieties , VG) then (Sugarcane parameter of better varieties(Sugarcane yields-Moderate, CCS% - Moderate, Brix % - Highest, 5 Cane weights - Highest), Yes) and (age, young man).*

j) *if(Quality of Varieties , VG) then (Sugarcane parameter of better varieties(Sugarcane yields-Highest, CCS% - Highest, Brix % - Very Highest, 5 Cane weights - Highest), Yes) and (age, old man).*

k) *if* (*Quality of Varieties , E) then* (*Sugarcane parameter of better varieties*(*Sugarcane yields-Highest, CCS% - Very Highest, Brix % - Highest, 5 Cane weights-Very Highest*), *Yes*) *and* (*Age, Middle Man*).

Flow Graph:

The flow diagram for the decision algorithm is shown below.

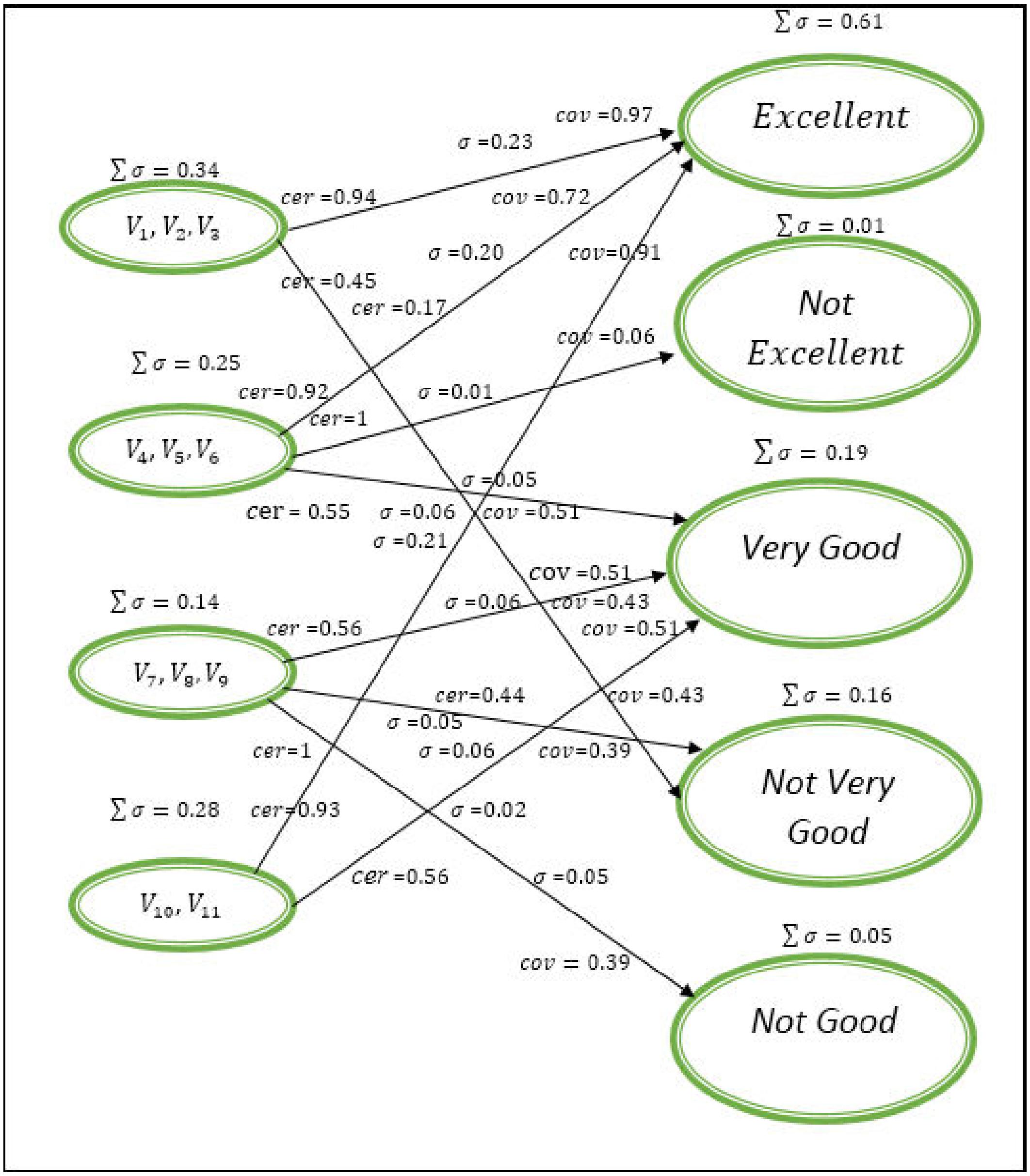

Figure No.1: Flow Graph

Conclusion

The RST is new mathematical techniques for dealing with inexact, unclear knowledge. A proposed task we sort out major sugarcane varieties produces more yields using rough set theory. This task used strength, certainty factor, and coverage factor of the DR signifies a though flow the relevant branch. Certainty factor of the decision rules to find out excellent varieties; 94% is V_3 , 93% is V_{11} and 92% is V_6 , very good varieties; 55% is V_4 , and V_9, and 56% is V_{10} , varietiesV_1 , V_2 and V_7 is not very good, varietiesV_5 and V_8 is neither excellent and not good. The coverage factors of the decision rules to find out excellent varieties; 97% is V_3 , 91% is V_{11} and 72% is V_6 , very good varieties; 51% is V_4 , V_9, and

V_{10} , varietiesV_1 , V_2 and V_7 is not very good, varietiesV_5 and V_8 is neither excellent and not good. Using RST the result we got assures that the farmers used sugarcane varieties V_3, V_{11}, and V_6.

References

1. Akama, S., Murai, T. and Kudo, Y.(2013). Heyting-Brouwer Rough Set Logic. In Proceedings of the KSE2013, Hanoi, Springer, Heidelberg.135-145.
2. Akama, S., Murai, T.(2005). Rough set semantics for three-valued logics. In Nakamatsu, K., Abe, J.M. (eds.) Advances in Logic Based Intelligent Systems. IOS Press, Amsterdam. 242-247.
3. Dubois, D., and Prade, H. (1990). Rough fuzzy sets and fuzzy rough sets .International Journal of General Systems.17: 191-209.
4. Orlowska, E. (1995). Information algebras. In V. S. Alagar, and M. Nivat. (Eds.), Algebraic methodology and software technology, Lectures Notes in Computer Science. 936: 50-65.
5. Orlowska, E. (1997). Studying incompleteness of information: A class of information logics. In K. Kijania- Placek, and J. Wolenski (Eds.). The Lvov-Warsaw School and contemporary philosophy. Dordrecht: Kluwer Academic Publishers. 303-320.
6. Orlowska, E. (1998). Introduction: What you always wanted to know about rough sets. In E. Orlowska (Ed.), Incomplete information: Rough set analysis. New York Physica-Verlag Heidelberg.
7. Pagliani, P. (1993). From concept lattices to approximation spaces Algebraic structures of some spaces of partial objects. Fundamental Informaticae.18: 1-25.
8. Pagliani, P.(1998). Intrinsic co-Heyting boundaries and information incompleteness in rough set analysis. In, Polkowski, L., Skowron, A. (eds.). Rough Sets and Current Trends in Computing. Springer, Berlin. 123–130.
9. Pawlak, Z. (1982). Rough sets, International Journal of Computer and Information Sciences.11: 341-356.
10. Pawlak, Z. (1991). Rough sets-Theoretical aspects of reasoning about data. Boston: Dordrecht: Kluwer Publishers.
11. Pawlak, Z. (2005). Flow graphs and data mining. Lecture Notes in Computer Science. 3400:1-36.
12. Prerna Mahajan, Rekha Kandwal, Ritu Vijay. (2012). Rough Set Approach in Machine Learning: A Review, International Journal of Computer Applications. 56(10): 1-13.
13. Yao, Y.Y. (1998). On generalizing Pawlak approximation operators. In Proceedings of 1st International Conference on Rough Sets and Current Trends in Computing (RSCTC'98). 298-307.