

## LITHOLOGY PREDICTION USING WELL LOGS: A GRANULAR COMPUTING APPROACH

TOUHID MOHAMMAD HOSSAIN<sup>1,\*</sup>, JUNZO WATADA<sup>1</sup>, IZZATDIN ABDUL AZIZ<sup>1</sup>  
MAMAN HERMANA<sup>2</sup>, SHEIKH TANZIM MERAJ<sup>3</sup> AND HIROSHI SAKAI<sup>4</sup>

<sup>1</sup>Department of Computer and Information Sciences

<sup>2</sup>Department of Geosciences

Faculty of Science and Information Technology

<sup>3</sup>Electrical and Electronics Engineering Department

Faculty of Engineering

Universiti Teknologi PETRONAS

Seri Iskandar, Perak Darul Ridzuan 32610, Malaysia

\*Corresponding author: [touhidmhossain@gmail.com](mailto:touhidmhossain@gmail.com)

[junzo.watada@gmail.com](mailto:junzo.watada@gmail.com); { [izzatdin](mailto:izzatdin); [maman.hermana](mailto:maman.hermana); [@utp.edu.my](mailto:sheikh.19001724) }

<sup>4</sup>Department of Basic Sciences

School of Engineering

Kyushu Institute of Technology

1-1 Sensui-cho, Tobata-ku, Kitakyushu-shi, Fukuoka 804-8550, Japan

[sakai@mns.kyutech.ac.jp](mailto:sakai@mns.kyutech.ac.jp)

Received June 2020; revised October 2020

**ABSTRACT.** *With the advancement of machine learning and artificial intelligence, the automated estimation of a bed's complex lithology has become one of the most crucial requirements in petroleum engineering because of its important role in reservoir characterization. In the past geophysical modelling, petro-physical analysis, artificial intelligence and several statistical approaches have been implemented to estimate lithology since prediction of lithology from recorded continuous cores are very expensive and unprofitable. Geoscience researchers often encounter uncertain, inexact, and vague data in the process of lithology identification that results in inefficient classification. Additionally, the complexities that are coupled to the lithology trends and their equivalent fluid responses, produce ambiguity and confuse the models. The goal of this work is to develop a lithology prediction technique by applying rough set theory (RST) as a granular computing approach to construct logical rules from an inconsistent information system that includes data from several well log attributes including the lithology indicator,  $SQ_p$  and the fluid indicator,  $SQ_s$  that have noticeable contribution in lithology classification. In addition, the rules will be established as a baseline for application in practice and future developments for multivariate well-log analysis. The results were validated with cutting data, and it was proved that the proposed approach has classified the lithology effectively with misclassification rate less than 18% which is less than other methods in comparison. Moreover, the result has confirmed that the method has a promising prospect as a lithology prediction tool, especially in real-time operation, because of the white-box nature of the module that represents the ability of describing the model's calculation steps and results in easily understandable form.*

**Keywords:** Lithology prediction, Rough sets, Granular computing, Knowledge acquisition, Decision making, Explainable AI

1. **Introduction.** Lithology refers to the composition or type of rock in the Earth's subsurface. The term lithology is used as a gross description of a rock layer in the subsurface and uses familiar names, including sandstone, siltstone, mudstone, etc. The lithology of a layer can be identified by drilling holes, although this method often does not provide exact information. We can also obtain the classification results of lithology from recorded continuous cores that are very expensive and might be unprofitable. The lithology can also be estimated by geophysical inversion and geophysical modelling methods. Lithology prediction can be performed using petro-physical well logs. The estimation of lithology from well logs multi-attribute data has become one of the most prominent techniques used by several sectors of petroleum engineering, including geological studies for reservoir characterization, reservoir modelling and formation evaluation, well planning including drilling and enhanced oil recovery processes, well completion management, etc. This study shows granular computation works effectively on a basis of rough sets (RS) to recognize the pattern of a few well-log attributes to predict lithology. In the era of smart data mining and analysis, granular computation involves the partitioning of an object into granules, with a granule being a clump of elements defined by similarity, indistinguishability, functionality or proximity [1, 2, 3]. Rough set theory (RST) [4, 5, 6] is a rising granular computing technique with a wide range of applications in many sectors, particularly in decision studies, inductive inference, conflict resolution, machine learning, knowledge discovery and acquisition, pattern recognition and inductive reasoning [7, 8]. It is very difficult for geoscientists to reduce the size of the dataset and to obtain the associated data simultaneously. RST addresses this challenge by performing as an efficient and effective module that reduces the data size in its computational process and discovers the hidden data patterns in database(s), which is an approach called knowledge discovery (KD) in the field of granular computing. KD has been used for the development of information systems that assist in extracting concealed data patterns and other important information in the datasets. RST performs granular computation from a vague idea (set) depending on two vivid concepts, which are lower approximations and upper approximations (discussed in Methodology section). To perform granular computation, RST requires only the provided data [9]. RST performs by employing a granular understanding of the provided dataset. The most significant benefit of white-box models like RST over black-box approaches is that the detailed knowledge of the classification process is available for better understanding the problem under study. RST includes numerous other advantages [10, 11]. Some are given here.

- Provide effective algorithms for finding invisible patterns inside the dataset.
- Generate a nominal dataset called the data reduction and, thus, show the significance of data.
- Determine relations, which are not found in statistical methods.
- Find partial or total dependencies among the objects and discover the significance of attributes.
- Generate set of significant decision rules from the datasets that are easily understandable and explainable.

Certainly, RST has been successfully applied in numerous fields of smart real-life applications, as an independent module or a combined module with other soft computing method(s) [12], including fuzzy logic (FL), neural network (NN), etc., to deal with uncertainty and vagueness. RST has also been used in several areas, such as cluster analysis [13], fault diagnosis [14], image processing [15], classification and prediction [11, 16, 17, 18].

Recently rule based white-box classification modules such as RST and other modules based on RST frameworks such as NIS (non-deterministic information system) Apriori

algorithm have been used in lithology interpretation and classification problems and achieved satisfactory outcomes [11, 17, 19]. Hossain et al. [19] have used NIS-Apriori rules to identify lithology classes directly from well logs and has achieved satisfactory results. In another research Hossain et al. have employed RST to classify electrofacies and interpret lithology classes from the electrofacies using RST rules [18]. This research proposes a white-box novel architecture to deal with the uncertainty and vagueness in the inconsistent well log datasets by using RST as a granular computing approach and to construct logical rules to classify ten different lithology classes. These interpretable rules will be useful for establishing a baseline for lithology identification and related applications in practice and future developments for multivariate well-log analysis.

The rest of the study is organized as follows. Section 2 contains the problem and the literature review. In Section 3 RST is explained as a form of granular computing and then the methodology RST based rule induction is described in detail. Section 4 contains the experimental steps and in Section 5 the experimental results are shown. In Section 6 the comparison study is given where three other soft computing methods are used. Section 7 includes the overall discussion and concludes the paper.

### Nomenclature

Variable	Description	Acronyms	Description
$CN_c$	Boundary region	ANN	Artificial Neural Network
$C^*$	Upper approximation	DS	Depth Sequence
$C_*$	Lower approximation	DT	Decision Tree
$G$	Shear modulus	FL	Fuzzy Logic
$I$	Information system	KD	Knowledge Discovery
$K$	Attributes' set	LDA	Linear Discriminant Analysis
$M$	Compressional modulus	ML	Machine Learning
$N$	Knowledge base	MLP	Multi-Layer Perception
$S$	Concept	NN	Neural Network
$SQ_p$	Scale of quality factor of P-wave	PNN	Probabilistic Neural Network
$SQ_s$	Scale of quality factor of S-wave	RBF	Radial Basis Function
$U$	Universe	RST	Rough Set Theory
$r$	Rough set rule	RS	Rough Sets
$\alpha$	Confidence of a particular rule	SOM	Self Organizing Maps
$\mu$	Rough membership function	SVM	Support Vector Machine
$\rho$	Density	TOC	Total Organic Carbon

**2. Problem and Background.** In earth science, the prediction of subsurface properties, such as lithology composition, has always been among the basic problems. An important concern in the oil and gas industry is reservoir characterization, and this is conducted to ensure the improvement of hydrocarbon prediction; this task requires that lithology be classified efficiently. Traditional methods that employ seismic data for the estimation of reservoir lithology consist of finding a physical relationship between the lithology properties to be identified and the seismic attributes, and then employ those attributes over the entire seismic dataset in order to predict the target. However, usually the seismic datasets include noise values due to sensor's noisy responses or equipment mis-measurements and hence the conventional methods provide inaccurate means to make lithology prediction. In some cases where the functional relationships between the attributes and the target properties can be found, the physical foundation is not often clear or understandable. Thus, relying on well logs is more efficient and recording such properties by well logging is considered to be much more reliable but costly, time consuming and challenging. There are close relationships between well log data and formation and conventionally, for lithology prediction by employing the recorded well log data, a wide

range of soft computing methods has been proposed through combinations of various measurements both in the qualitative and quantitative evaluation and hence automated lithology prediction using well logs has achieved outstanding contribution in the field of oil and gas. For solving lithology prediction and classification problems, several soft computing methods such as support vector machines using conventional wire-line well logs [20], cross plot interpretation and statistical analysis based on histogram plotting [21], FL for association analysis, NN and multivariable statistical methodologies [22], artificial intelligence approaches and multivariate statistical analysis [23], NIS apriori algorithm [19], hybrid NN methods [24], self organizing maps (SOM) [25], FL methods [26], artificial neural network (ANN) methods [27], fuzzy curves and ensemble neural networks [28], determination of the total organic carbon (TOC) using ANN [29], multi-agent collaborative learning architecture approaches [30], random forest [31, 32], generative adversarial network [33], multivariate statistical methods [34], aggregation of principal component, clustering and discriminant analysis [35], statistical characterization, and discrimination and stratigraphic correction methodologies [36] have been suggested by the researchers. However, majority of these methods are black box which makes it difficult to understand the models for further analysis [37]. The performance of ANN and FL methods are superior compared with statistical methods [20, 27, 38, 39]. SOM methods provide better results in lithology classification compared to other machine learning techniques [40]. Other kinds of NN are faster than probabilistic neural networks (PNN) because PNN involves more computational steps [39].

Moreover, due to high nonlinearity, uncertainty and the complexities that are coupled to the lithology trends and their equivalent fluid responses, the system eventually produces ambiguity that confuses the models making it difficult to get well-defined expressions using the conventional soft computing methods and improvement in prediction accuracy is still a major concern. Hermana et al. [41, 42] have addressed this problem, and proposed a unique transformation where decoupling can be performed by using the output of a constrained elastic inversion attributes,  $SQ_p$  and  $SQ_s$  which are derived from attenuation attributes through rock physics approximation by applying basic elastic properties: S-wave, P-wave and density. According to their method,  $SQ_p$  and  $SQ_s$  attributes are equivalent to the gamma ray and resistivity log response respectively and they can contribute for facies prediction or classification and  $SQ_p$  and  $SQ_s$  would have a noticeable effect on lithology classification and hydrocarbon prediction [41, 42]. The variations in lithology can be predicted accurately by the integration of well logs including  $SQ_p$  and  $SQ_s$  and geological core description data using RST rules applied to the wireline data.

Granular computing has a position of centrality in rough set theory. An interesting motivation of using granular computing is that it involves perception-based arithmetic. In this arithmetic, the objects of arithmetic operations are perceptions of numbers rather than numbers themselves. Informally, granulation involves partitioning of an object into granules, with a granule being a clump of elements drawn together by indistinguishability, equivalence, similarity, proximity or functionality. In this research a rough set theory (RST) based granular computing approach is proposed for special data pattern recognition system by using well log attributes along with the special attributes  $SQ_p$  and  $SQ_s$  in order to efficiently predict lithology. In this approach granulation is done in two steps, partitioning the raw attribute values into bins or intervals to reduce the noise and non-linearity; and partitioning of the objects with binned or discretized values into different nonoverlapping equivalence classes to represent the indiscernibility of the objects that are constructed based on distinct subsets of attributes. Under this viewpoint, rough set theory deals with approximation and reasoning with partitions of different levels of granularity to find patterns and generates significant descriptive rules showing the attributes'

dependencies, and these rules are used for predicting lithology in diverse field conditions with similar log responses.

### 3. Methodology.

**3.1. Granular computing and rough set theory (RST).** As given by Zadeh, the following definitions may assist us in knowing the scopes and the purpose of granular computing:

“Granulation of an object  $A$  leads to a collections of granules of  $A$ , with a granule being a clump of points (objects) drawn together by indistinguishability, similarity, proximity or functionality” [43].

“Granular Computing is a superset of the theory of fuzzy information granulation, rough set theory and interval computations, and is a subset of granular mathematics” [44].

From the quotations a fundamental idea of granular computing is the use of clusters, groups or classes of elements named as granules. In 1982 RST was proposed by Pawlak and since then it is in a state of continuous growth. The methodology of RST deals with the categorization and incorporation of uncertain, incomplete or imprecise knowledge and information, and has been considered as one of the first non-statistical approaches in data-science [45]. RST is a mathematical approach having granular structures to deal with imprecise or uncertain knowledge, knowledge discovery, quality evaluation, finding data patterns and consistencies, evaluation and recognition of data reasoning and dependencies on the basis of reduct of the information dataset. We consider the main concept of granular computing is studied from two interrelated features, the formation of the granules and computation with those granules for problem solving. Granular computing with RST is originated on the equivalence classes which are basic granules, and the lower and upper approximation spaces of a set can be computed by using those basic granules. In other words, these two approximation spaces (upper and lower) of any given set are the proper categorization of knowledge concerning the domain of interest [14]. The internal arrangements or the patterns of a given condition-decision dataset can be minimized to a set of nominal rules by employing the concept of *core* and *reduct* [45].

**3.2. Information system and decision table.** In RST the dataset is prepared in the form of a table where the columns of the table consist of a dependent attribute and some independent attributes. Each column consists of the information of a measurable property for the objects or observations. In the rows there are the data samples or observations. The entries in the table contain the attribute values. This table is called an information system. Formally, an information system is a pair  $I = (U, K)$ , where  $U$  is called universe and  $K$  is the attribute set and both of these sets are basically nonempty finite, i.e.,  $k : U \rightarrow V_k$  for  $k \subset K$ , where  $V_k$  is recognized as the domain of  $k$ . A decision table is a special kind of information system which expresses all the knowledge about the model. Formally, decision table,  $I = (U, K \cup D)$ , where attributes in  $K$  are called condition attributes and  $D$  is the decision attribute. The conditional attributes, also known as independent attributes contain the conditional values of an observation to form a decision. Decision attributes may have binary values like 0 and 1 or more than two values.

Table 1 below is an RS decision table, where  $D$  is the dependant or decision attribute and  $e, f, g, h$  are the independent or conditional attributes.

From Table 1 it is visible that in examples VII and VIII the independent attributes have the same values but the dependent or the decision attribute values are not the same. Table 1 is inconsistent because the objects VII and VIII are conflicting each other. In

TABLE 1. Decision table

<i>Objects</i>	<i>e</i>	<i>f</i>	<i>g</i>	<i>h</i>	<i>D</i>
I	N	O	N	7	5
II	N	O	P	5	5
III	O	O	N	5	5
IV	O	K	N	7	6
V	N	K	N	6	6
VI	O	K	O	7	7
VII	P	L	P	6	6
VIII	P	L	P	6	5
IX	P	K	O	7	7
X	P	L	O	7	7

Section 3.3 and Section 3.4 Table 1 is used to describe the basic steps to generate rules using RST methodology.

**3.3. Rough sets approximations.** The primary part of RS is the relation of indiscernibility that is produced by information that concerns the objects of interest. The intention of the indiscernibility relation is to define that by reason of the deficiency of knowledge it is not possible to separate some objects that participate in the available information system. Approximations are also among the vital concepts in RST. There are two different approximations in RST. *Lower Approximation* is defined as the domain of the sample objects which certainly belong to the interest subset. On the contrary, *Upper Approximation* consists of the sample objects that might belong to the interest subset. *Boundary Region* is formed with the sample objects that cannot be classified under any of the approximation sets.

The types of approximations in RST are described below. Let a set  $Y \subseteq U, C$  be an equivalence relation and a knowledge base  $N = (U, C)$ . So, mathematically the approximations are denoted as below:

$$C_*(Y) = \{y \in U : C(y) \subseteq Y\} \quad (1)$$

$$C^*(Y) = \{y \in U : C(y) \cap Y \neq \emptyset\} \quad (2)$$

Here, two data sets  $C_*(Y)$  and  $C^*(Y)$  are identified as the  $C$ -lower-approximation and the  $C$ -upper-approximation of  $Y$  respectively.

The *boundary region* is defined as

$$CN_C(Y) = C^*(Y) - C_*(Y) \quad (3)$$

If the borderline area of  $Y$  is empty which means if  $CN_C(Y) = \emptyset$ , then data set  $Y$  is said to be exact in  $C$ . Else if  $CN_C(Y) \neq \emptyset$ , then the data set  $Y$  is a rough set which relates to  $C$ . This definition is clearly demonstrated in Figure 1.

The definition of RS can also be given using a rough membership function [4], defined as

$$\mu_y^C(y) = |Y \cap C(y)| / |C(y)| \quad (4)$$

Apparently,

$$\mu_y^C(y) \in [0, 1] \quad (5)$$

The membership function value  $\mu_y^C(y)$  is a type of conditional probability, and it can be represented as a degree of certainty to which  $y$  belongs to  $Y$  (or  $1 - \mu_y^C(y)$  as a degree of uncertainty).

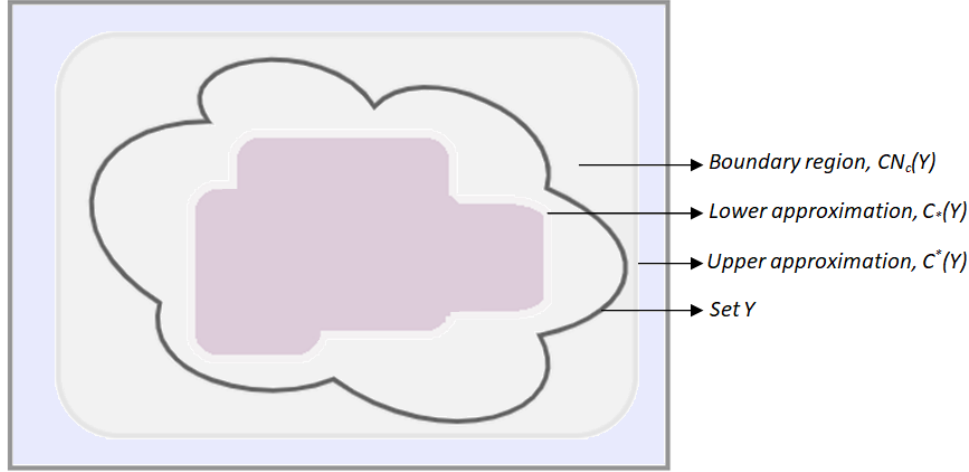


FIGURE 1. Rough sets

For defining *approximations* and the *boundary region* of a set the rough membership function can also be used which is shown below:

$$C_*(Y) = \{y \in U : \mu_y^C(y) = 1\} \quad (6)$$

$$C^*(Y) = \{y \in U : \mu_y^C(y) > 0\} \quad (7)$$

$$CN_C(Y) = \{y \in U : 0 < \mu_y^C(y) < 1\} \quad (8)$$

Now we can give two definitions for rough sets shown as follows.

**Definition 3.1.** Set  $Y$  is rough with respect to  $C$  if,

$$C_*(Y) \neq C^*(Y) \quad (9)$$

**Definition 3.2.** Set  $Y$  is rough with respect to  $C$  if for some  $y$ ,

$$0 < \mu_y^C(y) < 1 \quad (10)$$

In Table 1, {I}, {II}, {III}, {IV}, {V}, {VI}, {IX}, {X} and {VII, VIII}; these sets are the  $C$ -elementary sets and each one of these 9 sets is a granule of equivalence class. Now, in order to discover the concept from Table 1, first we need to classify *decision attributes* and  $C$ -elementary set associated with the *decision* as the subset of the example set having the same decision values. We call these subsets as *concept*. Three different *concepts* are found in Table 1:  $S_i = \{I, II, III, VIII\}$  where *decision* value is 5,  $S_{ii} = \{IV, V, VII\}$  where *decision* value is 6 and  $S_{iii} = \{VI, IX, X\}$  where *decision* value is 7.

Now let us find the *lower-* and *upper-approximation* and *boundary region* for three of these concepts.

$C$ -lower-approximation,

$$C_*(Y) = \{I, II, III, IV, V, VI, IX, X\}$$

$C$ -upper-approximation,

$$C^*(Y) = \{I, II, III, IV, V, VI, VII, VIII, IX, X\}$$

The *boundary-region*,

$$\begin{aligned} CN_C(Y) &= C\text{-upper-approximation} - C\text{-lower-approximation} \\ &= C^*(Y) - C_*(Y) \\ &= \{VII, VIII\} \end{aligned}$$

Hence, the consistent part of Table 1 turns into Table 2. It is found by abolishing the samples VII & VIII because they conflict each other.

TABLE 2. Consistent portion of Table 1

<i>Objects</i>	<i>e</i>	<i>f</i>	<i>g</i>	<i>h</i>	<i>D</i>
I	N	O	N	7	5
II	N	O	P	5	5
III	O	O	N	5	5
IV	O	K	N	7	6
V	N	K	N	6	6
VI	O	K	O	7	7
VII	P	K	O	7	7
VIII	P	L	O	7	7

**3.4. Reduction of attributes and rule generation.** *Attribute reduction* is a vital application information system and a core application of RST. The key concept is reducing the redundant information from the system but at the same time to keep the relation of indiscernibility.

Let us assume,  $L$  is a subset of  $K$  and  $k$  belongs to  $L$ . By definition:

- 1) If  $I(L) = I(L - \{k\})$ , then  $k$  is *discernible* in  $L$ ; or else  $k$  is *indiscernible* in  $L$ .
- 2) If attributes are necessary, then  $L$  is *independent*.
- 3) If  $I(L') = I(L)$  and  $L'$  is independent, then  $L'$ , the subset of  $L$  is an  $L$ 's reduction.

*Reduction* includes several functionalities. Let us explain them. The first one is finding the core of the attributes. *Core* is the attribute set where the attributes are common to all the *reducts*. Hence it consists of the attributes that are indispensable from the information system devoid of breaking the structure of the similarity class. We set  $L$  to be a subset of  $K$ . Then the set of the necessary attributes in  $L$  is the *core* of  $L$ . The main relationship between *core* and *reduction* is as follows:

$$Core(L) = \cap Reduct(L) \quad (11)$$

where  $Reduct(L)$  is all of the reductions of  $L$ .

So, *Reducts* of Table 2 will be  $\{e, g, h\}$ ,  $\{e, f, g\}$  and  $\{f, g, h\}$  and the core is attribute  $g$ . Since attribute  $g$  is the most important attribute in the table it is irremovable. From the table another indispensable attribute is obtainable by calculating the strength or confidence ( $\alpha$ ). The confidence for a particular rule  $r \rightarrow D$  is the ratio of number of example(s) which has  $r \cup D$  to the number of example(s) which has  $r$ .

The confidence  $e$ ,  $f$  and  $h$  attributes can be calculated for Table 2.

The confidence ( $\alpha$ ) of the rules for attribute  $e$  is obtained as:

- $(e = N) \rightarrow (D = 5)$ , value of  $\alpha = .66$ .
- $(e = O) \rightarrow (D = 5)$ , value of  $\alpha = .33$ .
- $(e = O) \rightarrow (D = 6)$ , value of  $\alpha = .33$ .
- $(e = N) \rightarrow (D = 6)$ , value of  $\alpha = .33$ .
- $(e = O) \rightarrow (D = 7)$ , value of  $\alpha = .33$ .
- $(e = P) \rightarrow (D = 7)$ , value of  $\alpha = 1.00$ .

Similarly the confidence ( $\alpha$ ) of rules for attribute  $f$  and  $h$  are obtained as:

- $(f = O) \rightarrow (D = 5)$ , value of  $\alpha = 1.00$ .
- $(f = K) \rightarrow (D = 6)$ , value of  $\alpha = .50$ .
- $(f = K) \rightarrow (D = 7)$ , value of  $\alpha = .50$ .



- $(f = L) \rightarrow (D = 7)$ , value of  $\alpha = 1.00$ .

and

- $(h = 5) \rightarrow (D = 5)$ , value of  $\alpha = 1.00$ .
- $(h = 6) \rightarrow (D = 6)$ , value of  $\alpha = 1.00$ .
- $(h = 7) \rightarrow (D = 5)$ , value of  $\alpha = .20$ .
- $(h = 7) \rightarrow (D = 6)$ , value of  $\alpha = .20$ .
- $(h = 7) \rightarrow (D = 7)$ , value of  $\alpha = .60$ .

Based on the confidence ( $\alpha$ ) it can easily be concluded that  $f$  is indispensable among the other attributes because of having the highest confidence. So, the *reduct* set of the set  $\{e, f, g, h\}$  becomes  $\{f, g\}$ . Hence, the *reduced* version of Table 2 becomes Table 3.

TABLE 3. Attribute reduction

<b>Objects</b>	<b>f</b>	<b>g</b>	<b>D</b>
I	O	N	5
II	O	P	5
III	O	N	5
IV	K	N	6
V	K	N	6
VI	K	O	7
VII	K	O	7
VIII	L	O	7

From Table 3 we can combine the rows having exactly the same values in conditional and decision attributes which we denote as *Row Reduction*. *Row Reduction* is shown in Table 4.

TABLE 4. Row reduction

<b>Objects</b>	<b>f</b>	<b>g</b>	<b>D</b>
I	O	P	5
II	O	N	5
III	K	N	6
IV	K	O	7
V	L	O	7

Next step is to find the *core* for the examples. From Table 4 *core* can be derived but the condition is that the table still needs to be consistent. Now, if we try to remove  $g = N$ , two  $D$  values are found (5 & 6). That implies, we cannot take an exclusive decision depending on  $g$ ; hence the  $g$  cannot be eliminated. Similarly, there are two  $D$  values (6 & 7) if we want to eliminate  $f = K$ . This means, we cannot make an exclusive decision based on attribute  $f$ . Hence, the value of  $f$  cannot be eliminated. Therefore, Table 4 can be reconstructed as Table 5 which shows the *core* of the examples.

By combining the identical rows we can again rewrite Table 5 as Table 6.

Now, reduction is not possible anymore. From Table 6 we can obtain the decision rules. On the basis of *reduct* and *core* we found the decision rules as the followings:

- 1) iff  $f \rightarrow O$  then  $D \rightarrow 5$
- 2) iff  $f \rightarrow K$  and  $g \rightarrow N$  then  $D \rightarrow 6$
- 3) iff  $g \rightarrow O$  then  $D \rightarrow 7$

TABLE 5. *Core finding*

<i>Objects</i>	<i>f</i>	<i>g</i>	<i>D</i>
I	O	*	5
II	O	*	5
III	K	N	6
IV	*	O	7
V	*	O	7

TABLE 6. Merging identical rows

<i>Objects</i>	<i>f</i>	<i>g</i>	<i>D</i>
I	O	*	5
II	K	N	6
III	*	O	7

## 4. Experiments.

### 4.1. Lithology dataset preparation.

#### 1) Data Collection:

In data collection, a total of 5,560 samples and 11 well log attributes from the well named “DMP Harvey 3” are considered in the digital well log dataset retrieved from WAPIMS, a Petroleum Exploration Database [47]. The core descriptions of “DMP Harvey 3” are also retrieved and they showed that the corresponding well is composed of ten major lithology classes, namely, Claystone, Mudstone, Siltston, Sandstone, Sandy Mudstone, Sandy Siltstone, Silty Sandstone, Silty Mudstone, Muddy Sandstone and Granulestone which are the items in the prediction class set [47]. From the core description dataset, the lithology classes are assigned for each depth interval. The corresponding values for the important well log attributes in each depth interval are also achieved by using the well log dataset. By combining these two datasets the main dataset for the experiment is created where each object has well log attribute values and corresponding lithology class information. For the main well log dataset, we have considered the following attributes as explanatory variables that contribute directly or indirectly to lithology classification.

- a) Gamma Ray Log (GR): This variable is used for measuring radioactivity of rocks and is helpful for geological correlations, separation of depth correlations between clayey zones and clean zones, etc. Shale-free sandstones and carbonates give low gamma ray readings. The gamma ray log response increases with growing shale content.
- b) Porosity or Neutron Log (NPHI): This variable is used for measuring the reaction of the rock to fast neutron bombardment, and its unit is dimensionless. The recorded factor is an index of hydrogen for a particular formation of lithology.
- c) Density Log (RHOB): The density log usually labelled ‘RHOB’ measures the density of the borehole and the rocks penetrated by the drill bit. The unit for density is gram per cubic centimetre.
- d) Photo Electric Effect Log (PE): The photo electric effect log is useful for determining lithology directly, because it has definitive matrix values, and linear interpolation between two end points works well. The unit is b/e (barns per electron).

- e) Density Porosity Log (DPHI): In the standard sequential interpretation process, the analyst determines porosity directly from the density log. It is conceptually the easiest of the porosity logs to interpret. The unit of DPHI is %, v/v decimal.
- f) Resistivity Log (RT10) and Conductivity Log (CT10): There are new tools that can cope with extremely highly resistive muds (oil-based muds or gas as the borehole fluid), which rely upon electromagnetic coupling and an induced alternating current (induction logs). The induction log actually measures resistivity and conductivity. The coil separation value is 10 inches for the attribute RT10 and CT10 and the unit is ohm and mm.ohm/m.
- g) Compressional Sonic Log:  
The sonic or acoustic log measures the travel time of an elastic wave through the formation. This information can also be used to derive the velocity of elastic waves through the formation. The tool measures the time it takes for a pulse of “sound” (i.e., and elastic wave) to travel from a transmitter to a receiver, which are both mounted on the tool. When the sound energy arrives at the receiver, having passed through the rock, it does so at different times in the form of different types of wave. The transmitter fires at  $t = 0$  and after some time the first type of wave arrives which is called compressional or longitudinal or pressure wave (P-wave). It is usually the fastest wave, and has a small amplitude. In the experiment, the attribute DTC represents compressional sonic log.
- h) Shear Sonic Log:  
The next wave, usually, to arrive after the P-wave is the transverse or shear wave (S-wave). This is slower than the P-wave, but usually has a higher amplitude. The shear wave cannot propagate in fluids, as fluids do not behave elastically under shear deformation. In the experiment, the attribute DTRS represents shear sonic log.
- i)  $SQ_p$ :  $SQ_p$  is a special attribute that is developed from seismic-attenuation rock physics using the concept of elastic inversion.  $SQ_p$  is used as lithology indicator [41].

$$SQ_p = \frac{5 (M/G - 2)^2}{6\rho (M/G - 1)} \quad (12)$$

where  $M$  and  $G$  are the compressional and shear modulus, respectively, which are measured under different conditions (high and low frequency conditions) and  $\rho$  is the density.

- j)  $SQ_s$ :  $SQ_s$  is a fluid indicator that is also derived from seismic-attenuation rock physics using the concept of elastic inversion [41].

$$SQ_s = \frac{10}{3\rho} \frac{M/G}{(3M/G - 2)} \quad (13)$$

## 2) Data Preparation:

The number of samples in the main dataset is 5560. We divided the main dataset into two subsets, training as DTr (70% or 3892) and testing as DTst (30% or 1668 samples). The training dataset, DTr was used to extract rules by using the RST methodology, and for validation or prediction, the testing dataset, DTst was used. The procedure for training the system with RST rules and predicting lithologies will be described in the next section and the overall workflow is shown in Figure 2.

**4.2. RST implementation.** To implement RST these following steps are required (as shown in Figure 3).

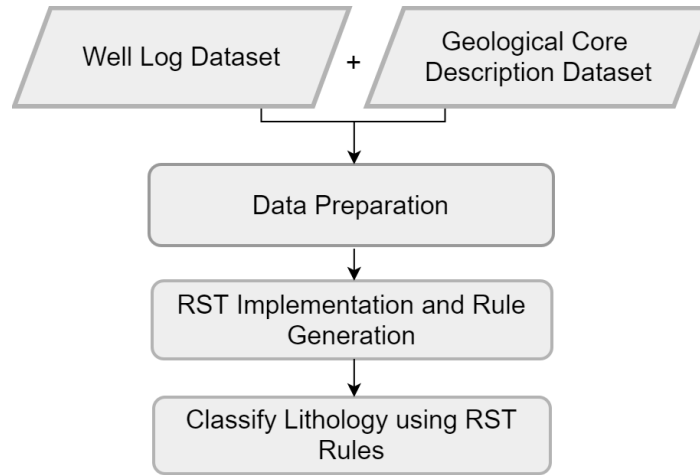


FIGURE 2. Workflow of RST-based lithology prediction module

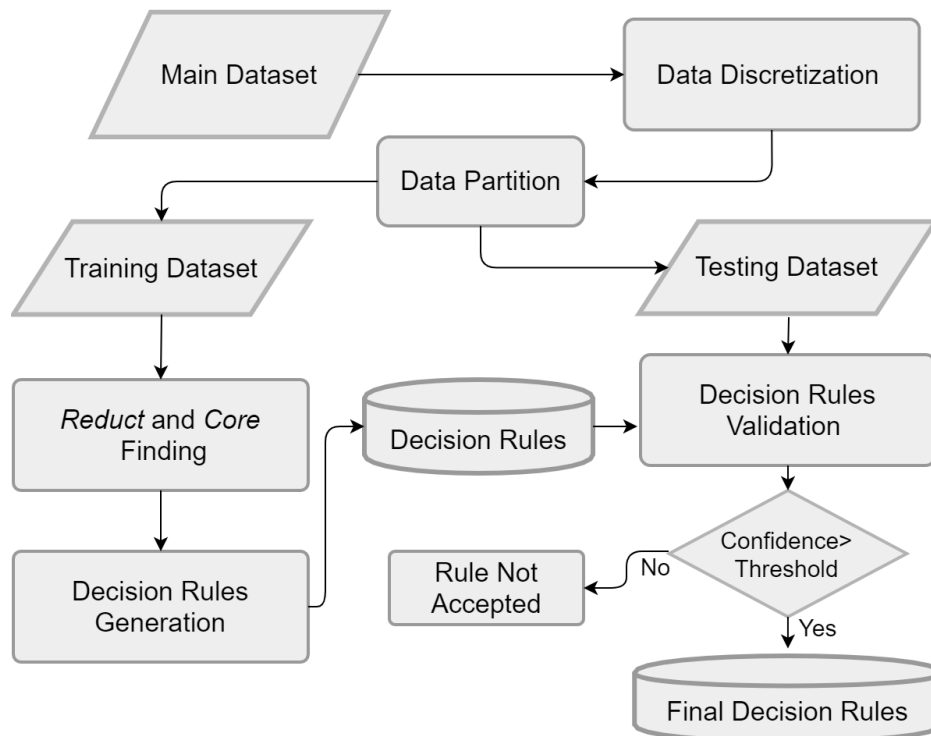


FIGURE 3. Flow chart of RST decision rules generation

1) Binning or Discretization: Binning or discretization is the process of transforming continuous variables into categorical variables. In this step the continuous explanatory attributes of the training dataset are discretized or binned into 24 bins or 25 equal length intervals. However, the decision attribute lithology does not need to be discretized since it consists of the lithology class information only. In Table 7 the cut points of all the attributes are shown. To discretize the values of an attribute  $A$  into  $N$  equal length intervals the following steps are taken:

- Minimum( $A$ ) and Maximum( $A$ ) are calculated.
- For each interval, width,  $w$  is calculated from the following equation:

$$w = [\text{Maximum}(A) - \text{Minimum}(A)]/N \quad (14)$$

- The cut points or bins are found from the following equation:

$$nth \text{ Cut Point} = \text{Minimum}(A) + w \times n \quad (15)$$

where  $n = \text{bin number} = 1, 2, 3, \dots, N$ .

- For  $N$  number of equal length intervals, there are  $N - 1$  cut points. For an example, if there are 2 bins or cut points, there are 3 intervals. The first interval contains the numbers between the minimum value and bin 1; the second interval contains the numbers between bin 1 and bin 2 (excluding bin 1) and the third interval contains the numbers between bin 2 and the maximum value (excluding bin 2). The minimum and maximum values of each attribute are denoted as  $-Inf$  and  $+Inf$  respectively.

For example, for the attribute GR in our experiment, the minimum and maximum values are 37.84 and 528.82 respectively. Since we have considered 25 equal length intervals, the width of each interval for GR is  $(528.82 - 37.84)/25$  or 19.64. So, the 1st cut point is  $37.84 + 19.64 \times 1$  or 57.48 (as shown in Table 7). Similarly, the last or 24th cut point for GR is  $37.84 + 19.64 \times 24$  or 509.2.

TABLE 7. Cut points for each well log attribute

Cut	GR	NPHI	RHOB	PE	DPHI	RT10	CT10	DTC	DTRS	$SQ_p$	$SQ_s$
1	57.48	0.150	1.401	0.596	0.064	80.10	381.1	64.77	122.6	0.138	0.430
2	77.12	0.195	1.453	1.019	0.095	160.1	761.7	68.72	125.5	0.154	0.446
3	96.76	0.240	1.505	1.442	0.125	240.1	1142	72.67	128.3	0.169	0.462
4	116.4	0.286	1.557	1.865	0.156	320.1	1523	76.62	131.1	0.185	0.478
5	136.0	0.331	1.609	2.288	0.186	400.1	1903	80.57	134.0	0.201	0.494
6	155.7	0.376	1.661	2.711	0.217	480.1	2284	84.53	136.8	0.216	0.510
7	175.3	0.421	1.714	3.134	0.247	560.1	2665	88.48	139.6	0.232	0.526
8	195.0	0.466	1.766	3.557	0.278	640.1	3045	92.43	142.5	0.248	0.542
9	214.6	0.511	1.818	3.979	0.308	720.1	3426	96.38	145.3	0.263	0.558
10	234.2	0.556	1.870	4.402	0.339	800.1	3806	100.3	148.1	0.279	0.574
11	253.9	0.601	1.922	4.825	0.369	880.1	4187	104.3	151.0	0.295	0.590
12	273.5	0.646	1.974	5.248	0.400	960.1	4568	108.2	153.8	0.311	0.606
13	293.2	0.691	2.026	5.671	0.430	1040	4948	112.2	156.6	0.326	0.623
14	312.8	0.736	2.079	6.094	0.461	1120	5329	116.1	159.5	0.342	0.639
15	332.4	0.781	2.131	6.517	0.491	1200	5709	120.1	162.3	0.358	0.655
16	352.1	0.826	2.183	6.939	0.522	1280	6090	124.0	165.1	0.373	0.671
17	371.7	0.871	2.235	7.362	0.552	1360	6471	128.0	168.0	0.389	0.687
18	391.3	0.916	2.287	7.785	0.583	1440	6851	131.9	170.8	0.405	0.703
19	411.0	0.961	2.339	8.208	0.613	1520	7232	135.9	173.6	0.420	0.719
20	430.6	1.006	2.391	8.631	0.644	1600	7612	139.9	176.5	0.436	0.735
21	450.3	1.051	2.443	9.054	0.674	1680	7993	143.8	179.3	0.452	0.751
22	469.9	1.096	2.496	9.477	0.705	1760	8374	147.8	182.1	0.468	0.767
23	489.5	1.141	2.548	9.899	0.735	1840	8754	151.7	185.0	0.483	0.783
24	509.2	1.187	2.600	10.32	0.766	1920	9135	155.7	187.8	0.499	0.799

- 2) Generation of *Reduct*: In this step *cores* and *reducts* are generated by using DTr. At the end of this step we found a minimal subset consists of the features and as shown in the methodology part, and this subset still provides the same quality of information that was present in the main dataset.

- 3) **Generation of Rules:** In this step significant rules are generated from the dataset DTr. According to the described methodology in Section 4. the support and confidence of the rules are also counted.
- 4) **Validation of the Generated Rules:** In this step validation of the reduced rules is performed, which turns the rules into final decision rules by incorporating the threshold accuracy or confidence which is also denoted as laplace. Confidence or laplace is an indication of how often the rule has been found to be true. (Laplace > .15 or 15% for our experiment). If this threshold is raised, the number of rules to solve the problem decreases resulting in the overall prediction accuracy to be less.

## 5. Experimental Results.

**5.1. Explanations of RST rules.** According to the described methodology, we have found significant rules from dataset DTr. These rules are applied to DTst to perform the task of prediction. We have found a total of 866 rules that differentiates 10 different lithology classes. In Table 8 the number of rules for describing each lithology class is shown. Some sample rules found from the experiment are given in Table 9.

TABLE 8. Lithology classes and their corresponding information

Lithology	Class No.	No. Rules	No. Samples in DTr
Claystone	1	11	27
Mudstone	2	119	286
Siltston	3	10	16
Sandstone	4	421	2682
Sandy Mudstone	5	96	239
Sandy Siltstone	6	51	145
Silty Sandstone	7	97	289
Silty Mudstone	8	3	5
Muddy Sandstone	9	9	32
Granulestone	10	49	172
Total		866	3893

In Table 9 we can see that, the model can show the contributions of the antecedent attributes in the rules in leading towards the decisions to classify the objects into different lithology classes.

**5.2. Calculating lithology prediction accuracy.** In this step, the 866 decision rules that we have found from the experiment by using the training dataset DTr are applied to the sample testing dataset DTst and the prediction accuracy or hit rate is calculated for each rule. Prediction accuracy (PA) of the module is calculated from DTst by employing the decision rules and following the equation below:

$$PA = \frac{\text{No. of correctly classified instances}}{\text{No. of instances in the test dataset}} \times 100\% \quad (16)$$

By applying the decision rules to DTr according to Equation (16), training score is calculated. Similarly, by applying the decision rules to the training dataset DTst, training score is calculated. In our experiment RST scores 0.8273 and 0.8261 for training and cross-validation which is shown in Table 10 as well.

**6. Comparison Study.** For comparison we have used three other techniques to obtain prediction accuracy of the lithology prediction problem: support vector machine (SVM), artificial neural network (ANN) and linear discriminant analysis (LDA).

TABLE 9. Sample RST rules

GR	NPHI	RHOB	PE	DPHI	IF	CT10	DTC	DTRS	$SQ_p$	$SQ_s$	THEN Lithology	Support Size	Laplace Value
(77.1, 96.8]					(880, 960]			(142, 145]			Mudstone	2	0.25
(77.1, 96.8]	(0.286, 0.331]				(400, 480]		(92.4, 96.4]				Mudstone	3	0.31
(77.1, 96.8]	(0.15, 0.195]							(148, 151]			Sandstone	4	0.36
	(0.195, 0.24]	(2.29, 2.34]	(1.44, 1.87]							(0.478, 0.494]	Sandstone	26	0.75
(96.8, 116]	(0.24, 0.286]		(6.09, 6.52]		(240, 320]				(0.201, 0.216]		Granulestone	3	0.31
				(0.186, 0.217]				(145, 148]			Granulestone	5	0.40
		(2.39, 2.44]	(5.25, 5.67]		$[-Inf, 80.1]$	$[-Inf, 381]$	(88.5, 92.4]				Silty Sandstone	20	0.70
(136, 156]		(2.23, 2.29]			(240, 320]		(84.5, 88.5]			(0.185, 0.201]	Sandy Siltstone	1	0.18
(136, 156]		(2.23, 2.29]	(4.4, 4.83]		(400, 480]			(159, 162]			Claystone	2	0.25
					(320, 400]						Claystone	1	0.18

## 6.1. Comparison methods.

6.1.1. *Support vector machine (SVM)*. SVM is a machine learning tool [48, 49, 50] proposed by Vladimir Vapnik in 1996 that has been used for 20 years to solve several problems, including lithology prediction [40, 51, 52]. To examine how SVM performs, the same datasets that we used for RST have been used. For training and testing the datasets, the same training and testing ratio (70 : 30) has been selected. For SVM the following settings have been selected.

- 1) Kernel = Radial basis function kernel.
- 2) C value = 1.
- 3) Cache Size = 200.

In Table 10 the training and cross validation scores for lithology prediction are shown and Figure 4 illustrates the results.

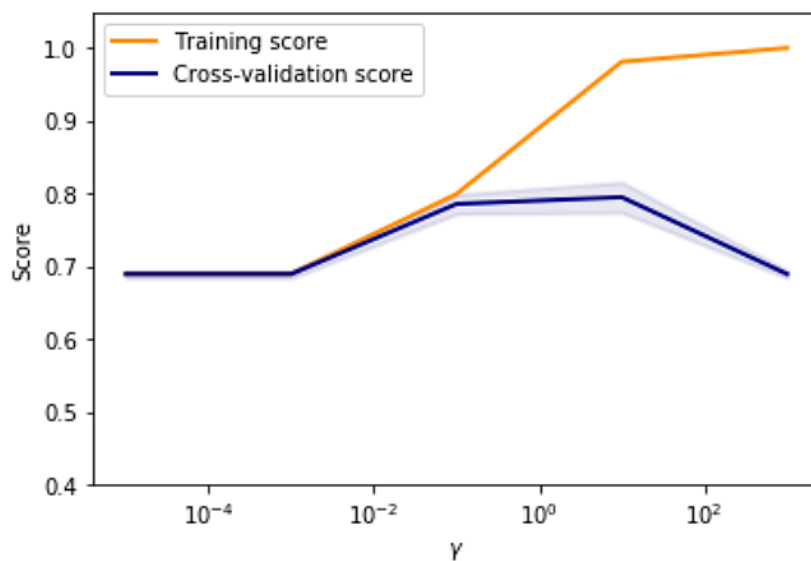


FIGURE 4. SVM scores for lithology prediction

6.1.2. *Artificial neural network (ANN)*. Several popular geophysical researchers suggest using ANN [53]. ANN can solve problems efficiently and discover very complex relationships between several variables. Over the years ANN has been used for solving data pattern recognition and classification problems [54, 55]. ANN methods have a remarkable ability to establish a complex mapping between nonlinearly coupled input and output data; hence they can perform well in prediction [56]. For analyzing how an ANN performs, the same training and testing datasets that we used for RST have been used and we have applied two different methods: multilayer perception (MLP) and radial basis function (RBF).

For the ANN with MLP, the following settings are selected:

- 1) *Input layer*  
Number of layers = 2.  
Activation Function = Hyperbolic tangent.
- 2) *Output layer*  
Activation Function = Sigmoid.

The prediction accuracy of the ANN is described in Section 6.2.



6.1.3. *Linear discriminant analysis (LDA)*. In the machine learning domain linear discriminant analysis (LDA) [57] is a very widely used technique for dimensionality reduction and pattern classification applications. It is used for modeling differences in groups, i.e., separating two or more classes. It is used to project the features in higher dimension space into a lower dimension space. For lithology classification several researchers used this technique [35, 58]. It is usually used as a black box and hence the model is not well understandable. In Table 10 the training and cross validation scores using LDA are shown.

6.2. **Comparison results.** Table 10 compares lithology prediction accuracy among RST, SVM, ANN and LDA and Figure 5 illustrates the results.

TABLE 10. Prediction accuracies for different methods

<i>Accuracy</i>	<i>RST</i>	<i>SVM</i>	<i>ANN (MLP)</i>	<i>ANN (RBF)</i>	<i>LDA</i>
Training accuracy	0.8273	0.7975	0.7690	0.7038	0.6945
Validation accuracy	0.8261	0.7793	0.7680	0.7073	0.6738

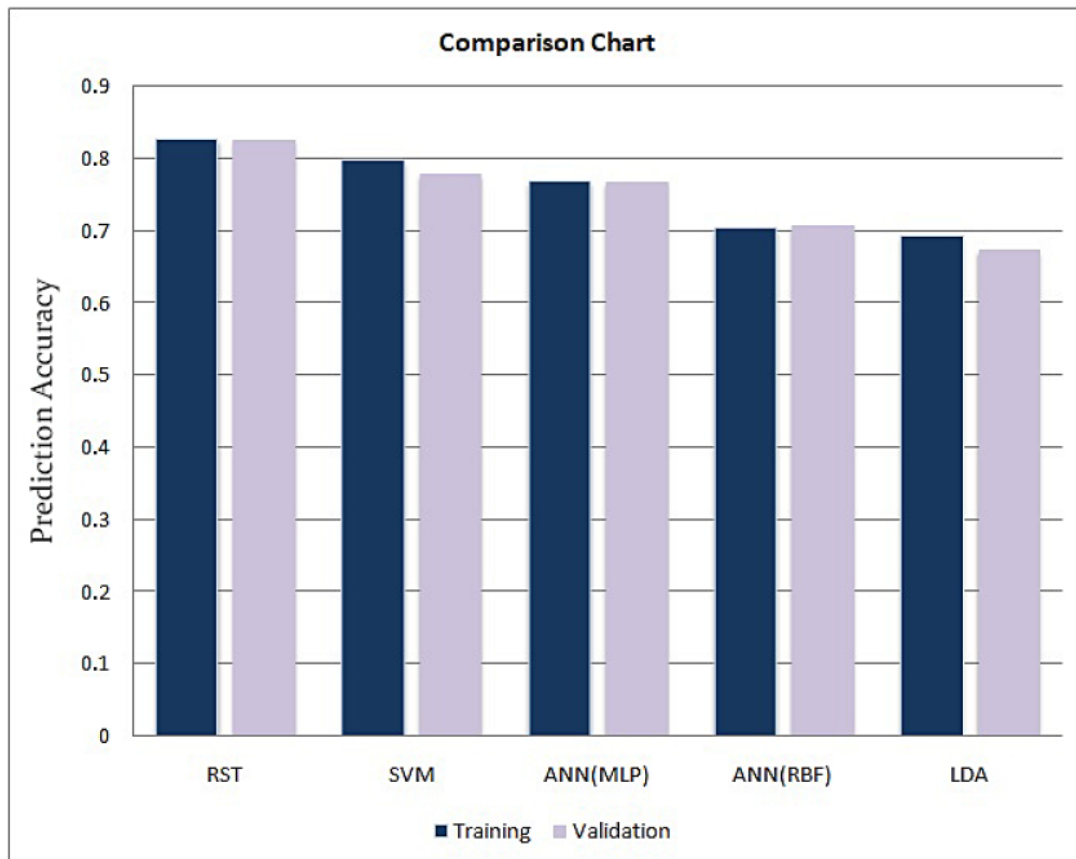


FIGURE 5. Pictorial representation of Table 10

From Figure 5 it is clear that ANN performs better than LDA, SVM performs better than ANN and LDA, and RST performs better than the others.

7. **Discussion and Conclusion.** In this paper we proposed a unique and efficient RST based granular computing method using well log attributes including the lithology indicator,  $SQ_p$  and the fluid indicator,  $SQ_s$ , to classify ten lithology classes. We also provided

the comparison study on the same dataset by employing some other renowned methods that have been widely used in this field it is vivid that granular computing based on the rule induction algorithm of RST offers a unique white-box ML approach for making prediction on lithology classes from the well log dataset including  $SQ_p$  and  $SQ_s$  by generating explainable decision rules. Lithology prediction has been a challenging problem in recent decades and our approach to lithology prediction has many noticeable positive outcomes. RST provides efficient algorithms for finding hidden patterns in the well log dataset and identifies their relationships in lithology. The prediction accuracy by RST is more than the other methods in comparison. Moreover, using RST-based rule induction, the decision-making process is performed in terms of easily interpretable and understandable rules and the rules themselves can explain the contributions of the antecedent well log attributes in defining the lithology classes for each object and this explainability and transparency of the algorithm makes it a white-box prediction module whereas almost all of the available lithology prediction modules are blackbox or hardly interpretable.

The reliability of RST-based prediction module strongly depends on the input training dataset. To guarantee a prediction module with better accuracy, an adequate training dataset is required. Choosing the appropriate explanatory attributes among a large number of well-log attributes is also challenging for the prediction modules. However, a larger dataset with more samples and datasets from different wells could also ensure a better prediction outcome. However, despite RST being independent in its numerous achievements to its tribute, in the future, for better accuracy, we will be working on combined modules of RST with other methods.

**Acknowledgment.** The authors would like to express their deepest gratitude to Petroleum Research Fund (PRF), Cost Center 0153AB-A33 and the project leader Dr. Eswaran Padmanabhan for supporting the research. The first author would like to thank and appreciate PRF and Universiti Teknologi PETRONAS (UTP) for providing Graduate Research Assistance (GRA) scheme.

## REFERENCES

- [1] J. T. Yao, A. V. Vasilakos and W. Pedrycz, Granular computing: Perspectives and challenges, *IEEE Trans. Cybernetics*, vol.43, no.6, 2013.
- [2] W. Pedrycz, Granular computing for data analytics: A manifesto of human-centric computing, *IEEE/CAA Journal of Automatica Sinica*, vol.5, no.6, 2018.
- [3] A. Bargiela and W. Pedrycz, *Granular Computing*, Handbook on Computational Intelligence, 2016.
- [4] Z. I. Pawlak, Rough sets and intelligent data analysis, *Information Sciences*, vol.147, pp.1-12, 2002.
- [5] S. K. Pal and A. Skowron, *Rough Fuzzy Hybridization – A New Trend in Decision Making*, Springer Verlag, Singapore, 1999.
- [6] T. Y. Lin, Y. Y. Yao and L. A. Zadeh, *Data Mining, Rough Sets and Granular Computing*, Physica Verlag, Heidelberg, 2002.
- [7] M. Magnani, “*Technical Report on Rough Set Theory for Knowledge Discovery*” in *Data Bases*, [www.researchgate.net/publication/technical\\_report\\_on\\_Rough\\_Set\\_Setr\\_Knowledge\\_Discovery\\_in\\_Data\\_Bases](http://www.researchgate.net/publication/technical_report_on_Rough_Set_Setr_Knowledge_Discovery_in_Data_Bases), Accessed on 2 April, 2019.
- [8] I. Dunsch, G. Gediga and H. S. Nguyen, *Rough Set Data Analysis in the KDD Process*, [citeseer.nj.nec.com/387773.html](http://citeseer.nj.nec.com/387773.html), Accessed on 21 May, 2019.
- [9] S. Tan, Y. Wang and X. Cheng, Text feature ranking based on rough-set theory, *IEEE/WIC/ACM International Conference on Web Intelligence*, 2007.
- [10] Z. Pawlak, Rough set theory for intelligent industrial applications, *Proc. of the 2nd International Conference on Intelligent Processing and Manufacturing of Materials*, Honolulu, HI, USA, 1999.
- [11] T. M. Hossain, J. Watada, M. Hermana and H. Sakai, A rough set based rule induction approach to geoscience data, *International Conference of Unconventional Modelling, Simulation & Optimization on Soft Computing and Meta Heuristics (UMSO 2018)*, Fukuoka, Japan, 2018.

- [12] A. Dasgupta, S. R. Broderick, C. Mack, U. Kota, R. Subramanian, S. Setlur, V. Govindaraju and K. Rajan, Probabilistic assessment of glass forming ability rules for metallic glasses aided by automated analysis of phase diagrams, *Scientific Reports*, vol.9, no.1, article no.357, 2019.
- [13] R. Li and Z. O. Wang, Mining classification rules using rough sets and neural networks, *European Journal of Operational Research*, vol.157, no.2, pp.439-448, 2004.
- [14] S. Asharaf, M. M. Narasimha and S. K. Shevade, Rough set based incremental clustering of interval data, *Pattern Recognition Letters*, vol.27, pp.515-519, 2006.
- [15] A. Petrosino and G. Salvi, Rough fuzzy set based scale space transforms and their use in image analysis, *International Journal of Approximate Reasoning*, vol.41, pp.212-228, 2006.
- [16] S. C. K. Shiu, Y. Li, S. K. Pal and J. N. K. Liu, A rough set-based case-based reasoner for text categorization, *International Journal of Approximate Reasoning*, vol.41, pp.229-255, 2006.
- [17] T. M. Hossain, J. Watada, M. Hermana and I. A. Aziz, Supervised machine learning in electrofacies classification: A rough set theory approach, *Journal of Physics: Conference Series*, 2020.
- [18] T. M. Hossain, J. Watada, I. A. Aziz and M. Hermana, Machine learning in electrofacies classification and subsurface lithology interpretation: A rough set theory approach, *Appl. Sci.*, DOI: 10.3390/app10175940, 2020.
- [19] T. M. Hossain, J. Watada, Z. Jian, H. Sakai, S. Rahman and I. A. Aziz, Missing well log data handling in complex lithology prediction: An NIS apriori algorithm approach, *International Journal of Innovative Computing, Information and Control*, vol.16, no.3, pp.1077-1091, 2020.
- [20] M. A. Sebtosheikh, R. Motafakkerfard, M. A. Riahi, S. Moradi and N. Sabety, Support vector machine method, a new technique for lithology prediction in an Iranian heterogeneous carbonate reservoir using petrophysical well logs, *Carbonates Evaporites*, vol.30, no.1, pp.59-68, 2015.
- [21] J. M. Busch, W. G. Fortney and L. N. Berry, Determination of lithology from well logs by statistical analysis, *SPE Form. Eval.*, vol.2, no.4, pp.412-418, 1987.
- [22] A. Carrasquilla, J. da Silvab and R. Flexa, Associating fuzzy logic, neural networks and multivariable statistic methodologies in the automatic identification of oil reservoir lithologies through well logs, *Rev. Geol.*, vol.21, no.1, pp.27-34, 2008.
- [23] J.-S. Lim, J. M. Kang and J. Kim, Interwell log correlation using artificial intelligence approach and multivariate statistical analysis, *SPE Asia Pacific Oil and Gas Conference and Exhibition*, Society of Petroleum Engineers, Jakarta, Indonesia, <https://doi.org/10.2118/54362-MS>, Accessed on 11 April, 2019.
- [24] S. Chikhi and M. Batouche, Hybrid neural network methods for lithology identification in the Algerian Sahara, *World Acad. Sci. Eng. Tech.*, vol.4, pp.774-782, 2007.
- [25] S. Chikhi and M. Batouche, Using probabilistic unsupervised neural method for lithofacies identification, *Int. Arab J. Inf. Technol.*, vol.2, no.1, pp.58-66, 2005.
- [26] S. J. Cuddy, Litho-facies and permeability prediction from electrical logs using fuzzy logic, *SPE Reserv. Eval. Eng.*, vol.3, no.4, pp.319-324, 2000.
- [27] M. Raeesi, A. Moradzadeh, F. D. Ardejani and M. Rahimi, Classification and identification of hydrocarbon reservoir lithofacies and their heterogeneity using seismic attributes, logs data and artificial neural networks, *J. Petrol. Sci. Eng.*, vol.82, pp.151-165, 2012.
- [28] I. Siregar, Y. Niu, P. Mostaghimi and R. T. Armstrong, Coal ash content estimation using fuzzy curves and ensemble neural networks for well log analysis, *International Journal of Coal Geology*, vol.181, no.1, pp.11-22, 2017.
- [29] A. A. Mahmoud, S. Elkatatny, M. Mahmoud, M. Omar, A. Abdurraheem and A. Ali, Determination of the total organic carbon (TOC) based on conventional well logs using artificial neural network, *International Journal of Coal Geology*, vol.179, pp.72-80, 2017.
- [30] C. M. Gifford and A. Agah, Collaborative multi-agent rock facies classification from wireline well log data, *Eng. Appl. Artif. Intell.*, vol.23, no.7, pp.1158-1172, 2010.
- [31] I. Priezzhev and E. Stanislav, Application of machine learning algorithms using seismic data and well logs to predict reservoir properties, *The 80th EAGE Conf. Exhib.*, 2018.
- [32] Y. Kim, R. Hardisty, E. Torres and K. J. Marfurt, Seismic-facies classification using random forest algorithm, *Proc. of SEG Tech. Program Expanded Abstr.*, Society of Exploration Geophysicists, pp.2161-2165, 2018.
- [33] G. Li, Y. Qiao, Y. Zheng, Y. Y. Li and W. Wu, Semi-supervised learning based on generative adversarial network and its applied to lithology recognition, *IEEE Access*, vol.7, 2019.
- [34] E. Puskarczyk, Application of multivariate statistical methods and artificial neural network for facies analysis from well logs data: An example of miocene deposits, *Energies*, vol.13, no.7, p.1548, 2020.

- [35] W. Teh, G. P. Willhite and J. H. Doveton, Improved reservoir characterization in the ogallah field using petrophysical classifiers within electrofacies, *Proceedings SPE Improved Oil Recovery Symposium*, Tulsa, Oklahoma, USA, <https://doi.org/10.2118/154341-MS>, Accessed on 11 April, 2019.
- [36] M. Borsaru, B. Zhou, T. Aizawa, H. Karashima and T. Hashimoto, Automated lithology prediction from PGNA and other geophysical logs, *Appl. Radiat. Isot.*, vol.64, no.2, pp.272-282, 2006.
- [37] G. Chen, M. Chen, G. Hong, Y. Lu, B. Zhou and Y. Gao, A new method of lithology classification based on convolutional neural network algorithm by utilizing drilling string vibration data, *Energies*, vol.13, no.4, DOI: 10.3390/en13040888, 2020.
- [38] H. Tang and C. D. White, Multivariate statistical log-log-facies classification on a shallow marine reservoir, *J. Petrol. Sci. Eng.*, vol.61, no.2, pp.88-93, 2008.
- [39] H. Tang, Successful carbonate well log facies prediction using an artificial neural network method: Wafra Maastrichtian reservoir, Partitioned Neutral Zone (PNZ), Saudi Arabia and Kuwait, *SPE Annual Technical Conference and Exhibition*, New Orleans, LA, 2009.
- [40] C. Deng, H. Pan, S. Fang, A. A. Konaté and R. Qin, Support vector machine as an alternative method for lithology classification of crystalline rocks, *J. Geophys. Eng.*, vol.14, pp.341-349, 2017.
- [41] M. Hermana, J. Ngui, C. W. Sum and D. P. Ghosh, Feasibility study of *SQp* and *SQs* attributes application for facies classification, *Geosciences*, vol.8, no.1, p.10, 2018.
- [42] M. Hermana, D. P. Ghosh and C. Weng, Discriminating lithology and pore fill in hydrocarbon prediction from seismic elastic inversion using absorption attributes, *The Leading Edge*, vol.36, no.11, pp.902-909, 2017.
- [43] L. A. Zadeh, Towards a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic, *Fuzzy Sets and Systems*, vol.19, pp.111-127, 1997.
- [44] L. A. Zadeh, *Announcement of GrC*, 1997, <http://www.cs.uregina.ca/~yyao/GrC/>, Accessed on 11 May, 2019.
- [45] Z. Pawlak, Rough sets, *International Journal of Computer and Information Science*, vol.11, no.5, pp.341-356, 1982.
- [46] R. Deja, Conflict analysis, rough set methods and applications, in *Studies in Fuzzyness and Soft Computing*, L. Polkowski, S. Tsumoto and T. Y. Lin (eds.), Physica-Verlag, 2000.
- [47] *Western Australian Petroleum and Geothermal Information Management System (WAPIMS)*, DMP Harvey 3, <https://wapims.dmp.wa.gov.au/WAPIMS/Search/Wells>, Accessed on 12 September, 2019.
- [48] D. Sun and J. Watada, Detecting pedestrians and vehicles in traffic scene based on boosted HOG features and SVM, *IEEE the 9th International Symposium on Intelligent Signal Processing (WISP) Proceedings*, Siena, Italy, <https://ieeexplore.ieee.org/document/7139161>, Accessed on 14 September, 2019.
- [49] M. Wu, Z. Xu and J. Watada, A memetic algorithm based support vector machine classification, *International Journal of Innovative Management Information & Production*, vol.5, no.1, pp.99-117, 2014.
- [50] S. Maldonado and J. López, Dealing with high-dimensional class-imbalanced datasets: Embedded feature selection for SVM classification, *Applied Soft Computing*, vol.6, pp.94-105, 2018.
- [51] Wikipedia, *Support Vector Machine*, [https://en.wikipedia.org/wiki/Support\\_vector\\_machine](https://en.wikipedia.org/wiki/Support_vector_machine), Accessed on 11 April, 2019.
- [52] A. Gasmi, C. Gomez, H. Zouari, A. Masse and D. Ducrot, PCA and SVM as geo-computational methods for geological mapping in the southern of Tunisia, using ASTER remote sensing data set, *Arab. J. Geosci.*, vol.9, p.753, 2016.
- [53] Y. Imamverdiyev and L. Sukhostat, Lithological facies classification using deep convolutional neural network, *Journal of Petroleum Science and Engineering*, vol.174, pp.216-228, 2019.
- [54] S. Chikhi and M. Batouche, Using probabilistic unsupervised neural method for lithofacies identification, *Int. Arab J. Inf. Technol.*, vol.2, no.1, pp.58-66, 2005.
- [55] S. Sahoo and M. K. Jha, Pattern recognition in lithology classification: Modeling using neural networks, self-organizing maps and genetic algorithms, *Hydrogeol. J.*, vol.25, pp.311-330, 2016.
- [56] J. R. Quinlan, Introduction of decision tree, *Machine Learning*, vol.1, pp.86-106, 1986.
- [57] A. Tharwat and A. Ibrahim, Linear discriminant analysis: A detailed tutorial, *AI Communications*, vol.30, no.2, pp.169-190, 2017.
- [58] S. Dong, Z. Wang and L. Zeng, Lithology identification using kernel Fisher discriminant analysis with well logs, *J. Pet. Sci. Eng.*, vol.143, pp.95-102, 2016.