



Cairo University  
**Egyptian Informatics Journal**

[www.elsevier.com/locate/eij](http://www.elsevier.com/locate/eij)  
[www.sciencedirect.com](http://www.sciencedirect.com)



**FULL-LENGTH ARTICLE**

# Rough – Granular Computing knowledge discovery models for medical classification



Mohammed M. Eissa<sup>a</sup>, Mohammed Elmogy<sup>b,\*</sup>, Mohammed Hashem<sup>c</sup>

<sup>a</sup> Information Systems Dept., Faculty of Business Administration and Information Systems, French University in Egypt, Cairo, Egypt

<sup>b</sup> Information Technology Dept., Faculty of Computers and Information, Mansoura University, Mansoura, Egypt

<sup>c</sup> Information Systems Dept., Faculty of Computers and Information, Ain Shams University, Cairo, Egypt

Received 20 October 2015; revised 15 December 2015; accepted 22 January 2016

Available online 26 February 2016

## KEYWORDS

Granular Computing;  
Genetic Algorithm;  
Knowledge discovery;  
Rough Mereology;  
Rough Sets

**Abstract** Medical domain has become one of the most important areas of research in order to richness huge amounts of medical information about the symptoms of diseases and how to distinguish between them to diagnose it correctly. Knowledge discovery models play vital role in refinement and mining of medical indicators to help medical experts to settle treatment decisions. This paper introduces four hybrid Rough – Granular Computing knowledge discovery models based on Rough Sets Theory, Artificial Neural Networks, Genetic Algorithm and Rough Mereology Theory. A comparative analysis of various knowledge discovery models that use different knowledge discovery techniques for data pre-processing, reduction, and data mining supports medical experts to extract the main medical indicators, to reduce the misdiagnosis rates and to improve decision-making for medical diagnosis and treatment. The proposed models utilized two medical datasets: Coronary Heart Disease dataset and Hepatitis C Virus dataset. The main purpose of this paper was to explore and evaluate the proposed models based on Granular Computing methodology for knowledge extraction according to different evaluation criteria for classification of medical datasets. Another purpose is to make enhancement in the frame of KDD processes for supervised learning using Granular Computing methodology.

© 2016 Production and hosting by Elsevier B.V. on behalf of Faculty of Computers and Information, Cairo University. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## 1. Introduction

The usage of Knowledge Discovery in the Database (KDD) for the growing databases is essential for facing the challenge of the great growing size of data and its developed complexity. KDD seeks to gather knowledge by identifying relations between data attributes for predictions. In addition, knowledge discovery techniques based on rule induction are important for the growing size of data [1].

\* Corresponding author.

E-mail address: [melmogy@mans.edu.eg](mailto:melmogy@mans.edu.eg) (M. Elmogy).

Peer review under responsibility of Faculty of Computers and Information, Cairo University.



Production and hosting by Elsevier

Many of the research-based knowledge discovery helped medical experts to identify relations among different medical indicators, to prevent false results, find out the minimum sets of medical indicators that affect in detecting of diseases, diagnostic, and disease prediction and improve the process of treatment decision making.

Granular Computing methodology for classification task is based on the terminology of “granule” as the building block of the model structure. Datasets for classification are represented in information table in which the data are separated into conditional attributes and decision class. Objects in dataset can construct the elementary granules of the classification system. Using similarity measures defined in Rough Sets (indiscernability), Fuzzy Logic (similarity) and other techniques Granular Computing methodology builds relationships between objects inside the same granule, between granules and between family of granules for classification.

Rough Sets and Fuzzy Logic and its extension of theories developed the similarity measure as the base of Granular Computing methodology. Using hybrid techniques based on the idea of Granular Computing in classification makes a difference by taking the advantages of each technique to improve the classification accuracy in modeling of KDD.

In this section, we highlight two common diseases in the modern world, which are Coronary Heart Disease (CHD) and Hepatitis C Virus (HCV).

Coronary Heart Disease (CHD) is the most widely recognized type of heart disease, it's a leading cause of death in adults worldwide, and specialists expect that CHD will be the first cause of death in many countries by the year 2020 [2]. Coronary Heart Disease is also called artery disease or atherosclerosis. It is the solidification of the vessels by fatty deposits called blockage in the artery. Oxygen and nutrients are essential for heart functionality. Blood carries the oxygen and nutrients to the heart through the blood vessels called arteries. Chest pain is defined as shortness of breath up to a heart attack, which is gained as a result of the accumulation of plaque on coronary and blood flow to the heart decreased [3].

The second common disease worldwide is the HCV, which is a liver disease caused by the Hepatitis C Virus. It sometimes results in intense illness, however regularly turns into a silent, chronic infection that can prompt liver failure and liver cancer that lead to death [4].

The paper is structured as follows: Section 2 reviews related work of knowledge discovery models. Section 3 demonstrates the structure of knowledge discovery models for classification of medical datasets. The evaluation performance measures are demonstrated in Section 4. Section 5 describes the used medical datasets, the experimental results, and comparison with other models. Finally, the conclusion is illustrated in Section 6.

## 2. Related work

Many researchers developed models based on Artificial Neural Networks, Fuzzy Logic, Genetic Algorithm techniques, and hybrid models for different knowledge discovery tasks. Because of Rough Sets strengths as a data analysis tool the scientists focused on building hybrid Rough based models for knowledge discovery (Rough Neural models–Fuzzy Rough models–Rough Genetic Algorithm models).

Dongbo et al. [5] proposed a model that uses Rough fuzzy model to construct loosely coupled Rough Neural network using adaptive Gustafson–Kessel clustering algorithm for dimension reduction and feature selection that combined the sample disturbance and attribute disturbance, and then using neural network LVQ learning algorithm for classification, the model obtained very good results in the test.

Rough-Mereology as an extension of Rough Set Theory is used in different application areas in knowledge discovery. For example, Zheng and Zhan [6] explored a model that Granular Computing can improve the performance of rule-based classification using Rough Mereology. However, the model was not extended more classifiers and working to develop it based on minimum length principle.

Polkowski and Artiemjew [7] developed a classifier for Coronary Heart Disease, first using data pre-processing techniques for dealing with the missing values and then applying the granular classifier to discover the absence or presence of coronary disease.

Polkowski and Artiemjew developed a model using granular reflections in the frame of Rough – Mereology for rules induction. A comparative analysis is made with exhaustive Rough Sets classifier whose accuracy is less than the proposed model [8].

Zaki et al. [9] introduced a Rough Set based model for classifying Hepatitis C Virus. The model utilizes Rough Set in data pre-processing by discretizing continuous medical indicators and a generating set of decision rules that determine the absence or presence of Hepatitis C Virus. However, the approach did not use any reduction algorithm for attribute reduction, and the accuracy of the classifier was acceptable.

Badria et al. [10] proposed a Rough based Granular model using the fundamental concepts of Rough Sets to discover dependencies between the attribute, discrete-continuous attributes, and the dynamic reduct for reduction. Then, they generated set rules for classifying HCV.

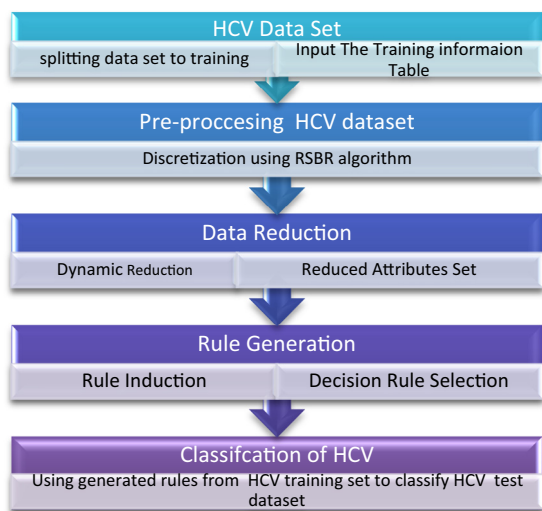
Eissa et al. [11] proposed a Rough-Genetic model for HCV classification using Rough Sets to induce a set of rules that reflect the whole universe. Then, the support measure is computed for each rule to filtering the rules and selecting the most frequent rules. The most promising rule is coded in chromosomes and using GAs operations to get better classification accuracy.

A comparative study was introduced by Ding et al. [12]. Between two models based on the concept of Granular Computing the concept of Rough Neural networks RNN and fuzzy neural networks FNN is used, and then the models are compared for making future development to improve the performance of them.

## 3. Knowledge discovery models

### 3.1. Rough based Granular Computing Model

The architecture of the Rough based Granular model is illustrated in Fig. 1. This model is divided into five phases: first, medical datasets are formulated into decision table. Second, discretization phase converts continuous data to intervals using Rough Sets algorithm. In the third phase, dynamic reduct algorithm is used to remove superfluous attributes and decrease the complexity of data. Finally, rules are



**Figure 1** The Rough based Granular Approach (RGA) used in medical classification.

generated from produced reducts with associated measures such as rule strength, rule certainty and rule coverage and the classification accuracy calculated, and rules evaluated by medical experts [10].

### 3.2. Rough Genetic Algorithm based Granular Computing Model

This model incorporates two novel data-mining methods: Rough Set Theory having a great power in data pre-processing and rule generation and gas to improve the generated rules for the classification processes, and provides three major phases. The architecture of the model is introduced and demonstrated as shown in Fig. 2, as follows [11]:

Phase 1 data pre-processing: It consists of removing redundant data using Rough indescribability relation and discretization of a continuous attribute by using Rough Sets and

Boolean Reasoning algorithm for improving the performance of the model.

Phase 2 data reduction: It is the Rough Set analysis phase that computes the minimal set of attributes that preserve the indescribability relation.

Phase 3 rule generation: this phase contains two sub-phases: first sub-phase is using the Rough Set for rule generation. Second sub-phase uses Genetic Algorithms (GAs) to improve the rules that are produced by the Rough Set algorithm. This phase contains two sub-phases, sub-phase 1: it generates initial chromosomes in which rules are encoded into chromosomes in an initial population as shown in Eq. (1).

Sub-phase 2: Genetic Algorithm operations (selection, crossover, and mutation) are applied for rules filtering and refinement.

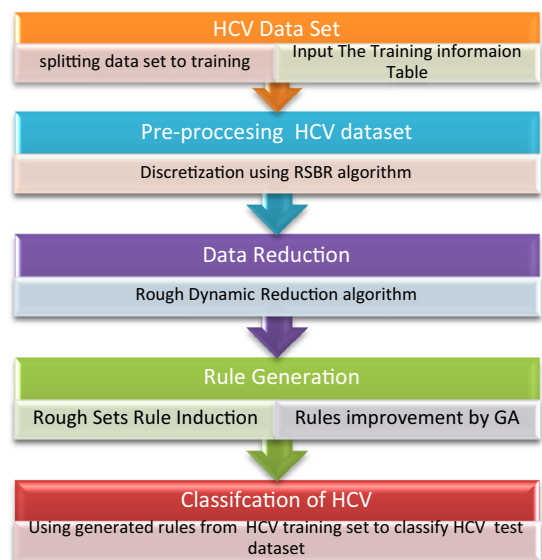
Phase 4 classifications of medical data: after calculating improved rules by Genetic Algorithm these rules are used for classification. These rules are used to predict unseen cases to make the treatments' decision of the new medication.

$$F = \frac{\text{(The number of classified correctly by the chromosome)}}{\text{(All of the observations in the training data)}} \quad (1)$$

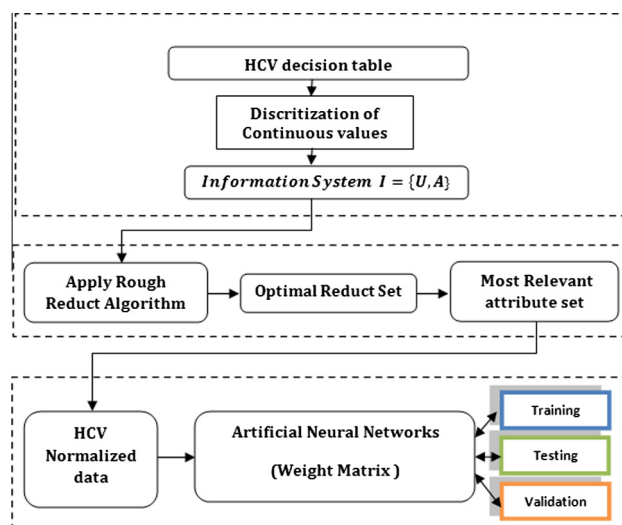
### 3.3. Rough – Granular neural network model

Rough Neural networks are based on the Rough Granular information and traditional neural network. The Granular information can be viewed as a collection of the same or similar properties or characteristics. Coarse granularity measure can solve complex problems in the real world. Based on the proposed concepts aforesaid, this paper suggests hybrid loosely Coupled Rough – neural network model. The model integrates two novel data-mining methods: Rough Sets in pre-processing (data discretization and reduction) and transformation and feed-forward back-propagation neural network algorithm for classification and prediction [13].

The architecture of the model to classify medical dataset is introduced and demonstrated in Fig. 3, as follows:



**Figure 2** The hybrid Rough Genetic model for classification.



**Figure 3** The proposed Rough-Granular neural model used in Medical classification.

Phase 1 data pre-processing: Discretization algorithm is based on the class label in its operations called Rough Sets and Boolean Reasoning algorithm is used. It converts continuous value attribute to discrete one to enhance the classification and decrease the complexity of data.

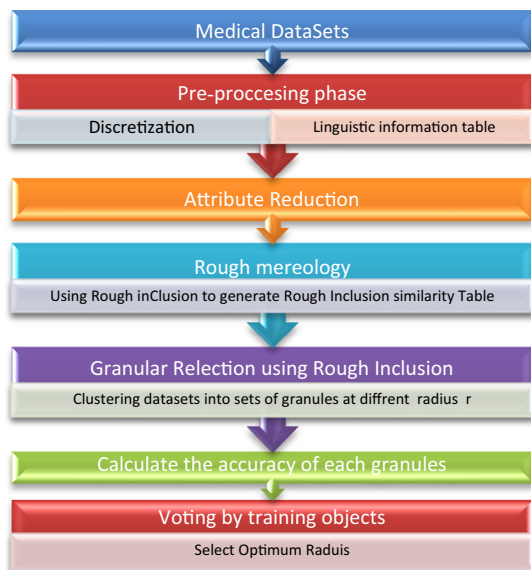
Phase 2 data reduction: reduction of attributes is used to find out the minimal set of medical indicators that the Rough Set analysis phase computes the minimal set of attributes that describe each object in datasets without losing the indescribability relation.

Phase 3 classification and prediction of medical dataset using ANN Back-Propagation Algorithm. In the classification phase, medical datasets were split into training and testing subsets, and then, data transformation to be suitable for neural network input layer and finally, network construction, training and model testing.

### 3.4. Rough Mereology based Granular Computing Model

The development processes of the proposed framework are demonstrated in Fig. 4. It requires pre-processing of medical dataset to remove redundancy and inconsistency, and convert continuous data to discretized one to be more suitable for processing. In addition, attribute reduction is needed to find the optimum attributes that represent the datasets without losing the value of the data. On the other hand, Granular Computing in the frame of Rough Mereology formalized the idea of granular reflection of medical datasets. First, Rough Mereology concept is applied in the frame of rough inclusion to medical datasets to induce rough inclusion similarity table.

Second, a set of rough inclusion tables are constructed by re-applying the first step with different radius for clustering the datasets into sets of granules with different radius. In Coronary Heart Disease dataset, 7 Rough inclusion tables are produced and 9 Rough inclusion tables are produced in Hepatitis C Virus datasets as shown in Tables 1 and 2.



**Figure 4** The proposed framework of the Rough-Mereology model.

**Table 1** Coronary Heart Disease granules accuracy.

Radius $r_{\text{gran}}$	Accuracy measure $\text{acc}_g$
$r0 = 0$	90.0
$r1 = 0.166667$	85.7
$r2 = 0.333333$	92.1
$r3 = 0.5$	96.2
$r4 = 0.666667$	93
$r5 = 0.833333$	0
$r6 = 1$	0

**Table 2** Hepatitis C Virus granules accuracy.

Radius $r_{\text{gran}}$	Accuracy measure $\text{acc}_g$
$r0 = 0$	88.9
$r1 = 0.111111$	74.5
$r2 = 0.222222$	94.7
$r3 = 0.333333$	93.6
$r4 = 0.444444$	95.4
$r5 = 0.555556$	96.6
$r6 = 0.666667$	90.2
$r7 = 0.777778$	0
$r8 = 0.888889$	0
$r9 = 1$	0

After the granular reflection of medical datasets reflecting inclusion tables into a set of granules, voting by training object is applied in two steps. First, compute the accuracy measure for each Rough inclusion table with different radius  $r$ . Second, select the optimum radius with the highest accuracy that represents the optimized granules.

The accuracy rate of each granule produced by Rough inclusion achieved its best score 96.2, 96.6 at granule radius  $r = r3$ ,  $r = r5$  for Coronary Heart Disease and Hepatitis C Virus datasets, respectively as shown in Tables 1 and 2.

## 4. Evaluation performance measures

The performance evaluation of knowledge discovery models to determine which of the models the superior classifier is should be straightforward. In this paper, we analyze different measures of determining a better model in the classification of datasets [14–16].

The natural performance measure for any classification problems is accuracy computed as shown in Eq. (2). However, higher accuracy rates do not necessarily imply better performance.

$$\text{Accuracy} = (\text{TP} + \text{TN}) / (\text{TP} + \text{FP} + \text{TN} + \text{FN}) \quad (2)$$

According to classification outcomes of binary decision four potential possibilities are expected. The True Positive (TP) and True Negative (TN) are correct classification outcomes. A False Positive (FP) happens when the result is mistakenly anticipated as positive when it is actually negative. A False Negative (FN) happens when the result is inaccurately anticipated as negative when it is really positive. True Positive Rate and False Positive Rate are figured as demonstrated in Eqs. (3) and (4):

$$\text{True Positive Rate} = \text{TP} / (\text{TP} + \text{FN}) \quad (3)$$

$$\text{False Positive Rate} = \text{FP} / (\text{FP} + \text{TN}) \quad (4)$$

Precision measure: it is the probability of retrieved instances that are relevant in Eq. (5).

$$\text{Precision measure} = \text{TP}/(\text{TP} + \text{FP}) \quad (5)$$

Recall measure: it is the probability of relevant instances that are retrieved in Eq. (6).

$$\text{Recall measure} = \text{TP}/(\text{TP} + \text{FN}) \quad (6)$$

*F*-measure: it is a measure that combines precision and recall is the harmonic mean of precision and recall, and *F*-measure is defined as in Eq. (7):

$$F\text{-measure} = 2 \times (\text{Precision} \times \text{Recall})/(\text{Precision} + \text{Recall}) \quad (7)$$

Cohen's Kappa Statistics: it is the statistical analysis based on the inter-rater agreement for qualitative data Eq. (8). It based on the analysis between different classes. Higher value of Kappa Statistic is considered as better performance.

$$\kappa = (P(a) - P(e))/(1 - P(e)) \quad (8)$$

where  $P(a)$  is the observed proportionate agreement ( $\text{TR} + \text{FN}/N$ ), and  $P(e)$  is the overall probability of random agreement.

Matthews's correlation coefficient (MCC): it is a correlation coefficient between the observed and predicted binary classifications; it returns a value between  $-1$  and  $1$ . The MCC can be calculated directly using the following Eq. (9):

$$\text{MCC} = ((\text{TP} \times \text{TN} - \text{FP} \times \text{FN}))/\sqrt{((\text{TP} + \text{FP})(\text{TP} + \text{FN}) \times (\text{TN} + \text{FP})(\text{TN} + \text{FN}))} \quad (9)$$

Error Rate: it is the measure of misclassification. It can be computed as shown in Eq. (10):

$$\text{Error Rate} = ((\text{FP} + \text{FN}))/((\text{TP} + \text{TN} + \text{FP} + \text{FN})) \quad (10)$$

## 5. Experimental results and analysis

### 5.1. Medical datasets description

In this paper, historical medical data are collected from different medical research resources [17–19]. In addition, several meetings with medical experts had been attended and discussed for understanding of the medical datasets and getting the clear idea about the diseases. The computations of rules have been only done on training dataset. The computations' results of the rules were applied to the classification of the granules from the tested dataset.

#### 5.1.1. Hepatitis C Virus (HCV) dataset

These data were gathered from clinical trials of a recently developed medication for HCV developed and patented by Badria and Attia [17]. It comprised of 119 HCV cases. Each case is portrayed by 28 medical indicators: 23 numerical indicators and five categorical indicators. The intention of the dataset is to forecast the presence or absence of the hepatitis virus. The HCV attribute description exists in Table 3. For each HCV record, patient data out of 27 condition attributes- and the decision attribute describe the presence or absence of HCV-related to the proposed medication. All this information was gathered from the treatment of HCV and

**Table 3** Hepatitis C Virus data.

No.	Medical indicator	Indicators description
1	Sex	Male or female
2	Source	Source of HCV: blood transfusion, non-sterile tools by dentist or surgery
3	S.G.P.T (ALT)	The normal range between 0 and 40 U/L
4	S.G.O.T (AST)	The normal range between 0 and 45 U/L
5	Serum Bilirubin (SB)	Normal range between 0 and 1.1 mg/dL
6	Serum Albumen (SA)	Serum Albumin; normal range between 3.5 and 5.1 g/dL
7	Serum ferritin	The normal range between 22 and 300
8	Ascites	No, mild, and ascites
9	Spleen	Normal, absent, and enlarged
10	Lesions	0, 1 or 2
11	Portal vein (P.V)	Natural diameter is 12 mm
12	PCR	Quantitative analysis of the virus U/mL
13	PLT	Platelets normal range between 150 and 450/cm m
14	WBC	White blood corpuscles normal range between 4 and 11/cm m
15	HGB hemoglobin	The range for male between 12.5 and 17.5 g/dL and range for female between 11.5 and 16.5 g/dL
16	Headache	Yes or no
17	Blood pressure	Yes or no
18	Nausea	Yes or no
19	Vertigo	Yes or no
20	Vomiting	Yes or no
21	Constipation	Yes or no
22	Diarrhea	Yes or no
23	Appetite	Yes or no
24	Gasp	Yes or no
25	Fatigue	Yes or no
26	Skin color	Yes or no
27	Eye color	Yes or no
28	Decision class	−1 absent, 1 present of HCV



**Table 4** Coronary Heart Disease data.

No.	Medical indicators	Indicator description
1	Age	Continuous values between 35 and 62
2	Sex	Male or female
3	Smoking	Yes or no
4	Diabetes mellitus (Dm)	Yes or no
5	Dyslipidemias (dyslipid)	Yes or no
6	Family history (family_h)	Yes or no
7	Left main coronary artery (lmca)	Normal or diseased
8	Left anterior descending artery (lad)	Normal or diseased
9	First diagonal artery (d1)	Normal or diseased
10	Second diagonal artery (d2)	Normal or diseased
11	Left circumflex artery (lcx)	Normal or diseased
12	Obtuse marginal artery 1 (om1)	Normal or diseased
13	Obtuse marginal artery 2 (om2)	Normal or diseased
14	Right coronary artery (rca)	Normal or diseased
15	Posterior descending coronary artery (pda)	Normal or diseased
16	Decision label	Number of vessels may be injected (no vessel, single, two, and multi)

**Table 5** HCV dataset models' performance measures.

Performance measures	Rough Mereology model	Rough Genetic model	Rough Neural model	Rough model
Accuracy	96.6	96.3	95	85.7
Error Rate	3.4	3.7	5	14.3
True Positive Rate	0.97	0.963	0.949	0.857
False Positive Rate	0.36	0.23	0.1	0.102
Precision	1	0.977	0.98	0.872
Recall	0.96	0.977	0.96	0.944
Kappa Statistics	0.67	0.57	0.54	0.3103
MCC	0.88	0.877	0.74	0.69
F-measure	0.98	0.977	0.97	0.91

**Table 6** CHD dataset models' performance measures.

Performance measures	Rough Mereology model	Rough Genetic model	Rough Neural model	Rough model
Accuracy	97.4	97.3	96	91.7
Error Rate	2.6	2.7	4	8.3
True Positive Rate	0.97	0.97	0.96	0.92
False Positive Rate	0.24	0.22	0.26	0.25
Precision	1	0.97	0.98	0.9
Recall	0.97	0.97	0.96	0.92
Kappa Statistics	0.94	0.96	0.84	0.73
MCC	0.91	0.9	0.74	0.82
F-measure	0.98	0.98	0.97	0.93

divided with splitting factor 25% into HCV training set and HCV test set.

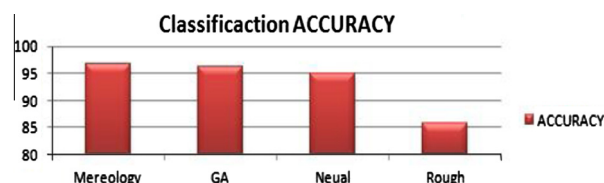
### 5.1.2. Coronary Heart Disease (CHD) dataset

CHD dataset was collected from Cardiology Department, faculty of medicine at Mansoura University. It consists of 215 Coronary patients that include condition attributes such as age, sex, family history, smoking, and other medical indicators.

In addition, it shows the decision label that shows the presence or absence of Coronary Heart Disease. Most of the attributes are binary attributes and age attribute is only numerical one, as shown in Table 4. For training and testing purposes, CHD dataset was divided into training and testing sets with the split factor 25%. It has been used to partition the datasets patented [19].

### 5.2. Model performance analysis

The performance analysis was carried on four different knowledge discovery models described in previous section. The datasets considered here are from the medical domain. The

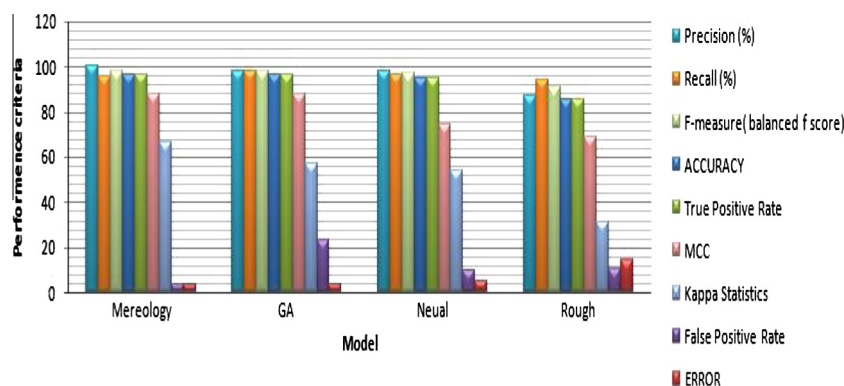
**Figure 5** HCV classification accuracy.

classification algorithms used in this work are described in previous section.

The results obtained from these different models have been defined in the form of tables that indicate the performance measures used such as accuracy, Kappa Statistics, Error Rate, True Positive Rate, False Positive Rate, Precision, Recall, F-measure and Matthews correlation coefficient (MCC) (Tables 5 and 6) as well as a graph (Figs. 5 and 6).

## 6. Conclusions

In this paper, knowledge discovery models in frame of Granular Computing methodology were proposed to facilitate the



**Figure 6** Comparative analysis performance measures of KDD models.

knowledge extraction process and classification of medical datasets.

Rough Sets – Granular Computing Model is proposed as a basic model applying the concept of Rough Sets Theory in different knowledge discovery processes (pre-processing-reduction and rule generation) to classify CAD and HCV medical datasets to induce a set of rules for making treatment decisions.

Then hybrid Rough Sets – Genetic Algorithm model is introduced that tries to use the powerful Rough Sets in data analysis and Genetic Algorithm as important tool for rule optimization to maximize the accuracy of produced decision rules.

In addition, another hybrid loosely Coupled Rough – Neural Network model is proposed. The model integrates Rough Sets in pre-processing and transformation of CAD and HCV datasets and feed-forward back-propagation neural network algorithm for classification.

Finally, Rough-Mereology model is demonstrated as an extension of Rough Sets in frame of Granular Computing that replaced the indescribability relation with similarity relation Rough inclusion relation. Rough-Mereology is dedicated to the concept of Granular Computing in constructing elementary information granules. It finds the relationships between information granules and building granules network. This model is applied also through a case study of HCV and CAD datasets.

This paper shows how the proposed models can be utilized and adapted to extract set of treatment decision rules that help medical experts in diseases diagnosing process. The results indicate the extracted rules are useful to predict unseen medical cases.

A comparative analysis among the proposed models indicates that Rough-Mereology model is the best model in knowledge extraction and in classification according to set of evaluation criterias related to classification.

Understanding the differences between knowledge discovery models can help medical experts to select the optimal model for classification suitable for medical datasets related to Hepatitis C Virus and Coronary Heart Disease.

The limitation of the proposed models can be summarized in two points. First, the complexity time of the models especially in pre-processing phase is high. Second is the data inconsistency problem. These limitations will be considered as a future work to reduce time complexity during pre-processing

phase and by using algorithms to solve the inconsistency problem.

## References

- [1] Gerhard S, Wendler T, editors. *Advances in healthcare technology: shaping the future of medical care*, vol. 6. Springer Science & Business Media; 2006. p. 403–19.
- [2] Meenu D, Dadhich P. *Applications of data mining techniques: empowering quality healthcare services*. Int J Inf, Commun Comput Technol 2013.
- [3] Raphael T, Tillement JP, Hauet T. Renal protective effect of metabolic therapy in patients with coronary artery disease and diabetes: from bench to bed side. *Curr Pharm Des* 2009;15(8): 863–82.
- [4] Fred A. *Hepatitis C: the silent epidemic*. Da Capo Press; 2007.
- [5] Dongbo Z, Wang Y, Huang H. Rough neural network modeling based on fuzzy rough model and its application to texture classification. *Neuro Comput* 2009;72(10):2433–43.
- [6] Zheng HZ, Zhan DC. Association rule algorithm based on bitmap and granular computing. *Artificial Intelligence Machine Learning (AIML) J* 2005;5:51–4.
- [7] Polkowski L, Artiemjew P. Granular computing: granular classifiers and missing values. In: *Cognitive informatics, 6th IEEE international conference*; 2007. p. 186–94.
- [8] Polkowski L, Artiemjew P. Granular computing in the frame of rough mereology. A case study: classification of data into decision categories by means of granular reflection of data. *Int J Intell Syst* 2011;26:555–71.
- [9] Zaki A, Salama MA, Hefny HA, Hassanien AE. *Rough sets-based rules generation approach: a Hepatitis C virus data sets*. *Advanced machine learning technologies and applications*. Berlin Heidelberg: Springer; 2012. p. 52–9.
- [10] Badria F, Eissa MM, Elmogy M, Hashem M. Rough based granular computing approach for making treatment decisions of Hepatitis C. In: *23rd International conference on computer theory and applications ICCTA*; 2013.
- [11] Eissa MM, Elmogy M, Hashem M, Badria F. Hybrid rough genetic algorithm model for making treatment decisions of Hepatitis C. In: *2nd International conference for engineering and technology ICET, GUC, Cairo*; 2014.
- [12] Ding S, Ding HJ, Fengxiang Jin. *Granular neural networks*. *Artificial Intelligence Rev* 2014;41(3):373–84.
- [13] Eissa MM, Elmogy M, Hashem M. Rough – granular neural network model for making treatment decisions of Hepatitis C. In: *Informatics and systems (INFOS), 2014 9th international conference on*. IEEE; 2014.
- [14] Ragab Abdul Hamid M, Noaman Amin Y, Al-Ghamdi Abdullah S, Madbouly Ayman I. A comparative analysis of classification

- algorithms for students college enrollment approval using data mining. In: Fardoun Habib M, Gallud José A, editors. Proceedings of the 2014 workshop on interaction design in educational environments (IDEE '14). New York, NY, USA: ACM; 2014. p. 106. <http://dx.doi.org/10.1145/2643604.2643631>, 8 pages.
- [15] Santos AM, Canuto AM, Neto AF. A comparative analysis of classification methods to multi-label tasks in different application domains. *Int J Comput Inf Syst Ind Manage Appl* 2011;3:218–27.
- [16] Mittal P, Gill NS. A comparative analysis of classification techniques on medical data sets. *Int J Res Eng Technol* 2014;3(6):454–60.
- [17] Badria F, Attia H. Effect of Selected natural products, thiopropine, and pegasys on hepatic platelet activation factor (PAF) in CCl4-induced hepatic fibrosis in rats. *Saudi Pharm J* 2007;15(2): 96–104.
- [18] Barakat SI, Eissa MM, EL-Henawy I. Hybrid rough sets and probabilistic flow graph model in coronary artery disease. *Egyptian Comput Sci J* 2009;33.
- [19] Amir MZ, Eissa MM, EL-Henawy I. Hybrid rough sets and decision tree model in coronary heart disease. *Zagazig Univ Med J* 2010.