Data Summarization Approach to Relational Domain Learning Based on Frequent Pattern to Support the Development of Decision Making

Rayner Alfred[1, 2] and Dimitar Kazakov[1]

[1] University of York, Computer Science Department, Heslington,
YO105DD York, United Kingdom
{ralfred, kazakov}@cs.york.ac.uk
http://www-users.cs.york.ac.uk/{~ralfred,~kazakov}
[2] On Study Leave from Universiti Malaysia Sabah,
School of Engineering and Information Technology,
88999, Kota Kinabalu, Sabah, Malaysia
ralfred@ums.edu.my

Abstract. A new approach is needed to handle huge dataset stored in multiple tables in a very-large database. Data mining and Knowledge Discovery in Databases (KDD) promise to play a crucial role in the way people interact with databases, especially decision support databases where analysis and exploration operations are essential. In this paper, we present related works in Relational Data Mining, define the basic notions of data mining for decision support and the types of data aggregation as a means of categorizing or summarizing data. We then present a novel approach to relational domain learning to support the development of decision making models by introducing automated construction of hierarchical multi-attribute model for decision making. We will describe how relational dataset can naturally be handled to support the construction of hierarchical multi-attribute model by using relational aggregation based on pattern's distance. In this paper, we presents the prototype of "Dynamic Aggregation of Relational Attributes" (hence called DARA) that is capable of supporting the construction of hierarchical multi-attribute model for decision making. We experimentally show these results in a multi-relational domain that shows higher percentage of correctly classified instances and illustrate set of rules extracted from the relational domains to support decision-making.

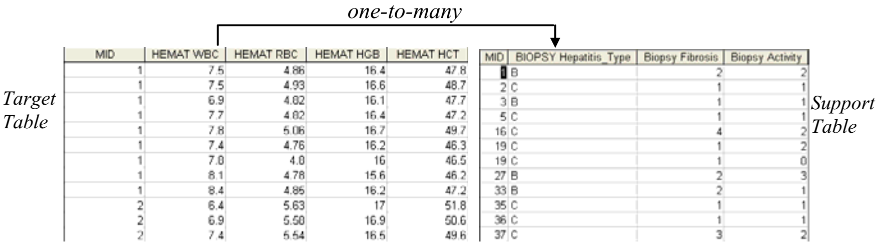
1 Introduction

The processing power to acquire and store large amount of data on documents has increased dramatically over the last few years. Despite the growing of computational power of modern computers, our abilities, to analyze these data for decision-making, are limited for data stored in relational model (multiple tables). We need to join these multiple tables in order to get more information about a specific record stored in target table that has *one-to-many* relationship with data stored in another table.

However, most traditional data mining tools cannot handle relational dataset with high-dimensional of *one-to-many* relationship, unless pre-processing task is applied to the data for data conversion. For instance, Fig 1 depicts two tables involved in the

X. Li, O.R. Zaiane, and Z. Li (Eds.): ADMA 2006, LNAI 4093, pp. 889 – 898, 2006.
© Springer-Verlag Berlin Heidelberg 2006

hepatitis database. The patients' various exams are not directly related, so joining these tables for a common analysis fails to provide a suitable dataset for discovering rules based on traditional data mining algorithm such as Apriori [32]. In other words, the results deriving from joint tables may lead to data redundancy and thence to distortions in the discovering the rules.



MID	HEMAT WBC	HEMAT RBC	HEMAT HGB	HEMAT HCT
1	7.5	4.86	16.4	47.8
1	7.5	4.93	16.6	48.7
1	6.9	4.82	16.1	47.7
1	7.7	4.82	16.4	47.2
1	7.8	5.06	16.7	49.7
1	7.4	4.76	16.2	46.3
1	7.8	4.8	16	46.5
1	8.1	4.78	15.6	46.2
1	8.4	4.85	16.2	47.2
2	6.4	5.63	17	51.8
2	6.9	5.58	16.9	50.6
2	7.4	5.54	16.5	49.6

MID	BIOPSY Hepatitis_Type	Biopsy Fibrosis	Biopsy Activity
1	B	2	2
2	C	1	1
3	B	1	1
5	C	1	1
16	C	4	2
19	C	1	2
19	C	1	0
27	B	2	3
33	B	2	1
35	C	1	1
36	C	1	1
37	C	3	2

Fig. 1. Hepatitis dataset tables (KDD CUP 2005)

Data representation stored in a relational model differs from the traditional feature-vector (single table) representation in which the relational model is more expressive than attribute-value model in capturing and describing complex structure and relationships in the business/medical domain. Mining knowledge from relational databases to support *decision-making* process has a few advantages over mining knowledge from a single table, namely: *standardized relational format for most database*, *highly expressive power in capturing complex data*, and *the ability to integrate background knowledge* [16], [17]. The integration of relational data mining and decision support is not new, however not much research done in supporting their integration. This work presents a novel approach that is capable of learning relational domain and generating automated hierarchical multi-attribute model to support the development of decision-making system. In this approach, we describe the technique of generalizing data with *one-to-many* relationships using granularity computing as a means of data summarization to automate and support the construction of hierarchical multi-attribute modeling (HMAM) for decision-making [2]. In section 2 we introduce related works in relational data mining. Section 3 covers the concept of hierarchical multi-attribute model in decision modeling. Section 4 describes the pattern-based aggregation approach to relational data mining and discusses the pre-processing procedure. Experimental results are presented in section 5 and this paper is concluded in section 6.

2 Related Works in Relational Data Mining

Relational learning research is not a new research area and it has a long history. Muggleton and DeRaedt [13] introduce the concept of Inductive Logic Programming (ILP) and its theory, methods and implementations in learning multi-relational domains. ILP methods learn a set of existentially unified first-order Horn clauses that can be applied as a classifier [5].

In a relational learner based on logic-based propositionalization [11], instead of searching the first-order hypothesis space directly, one uses a transformation module to compute a large number of propositional features and then apply a propositional learner. Both ILP and binary propositionalization lack support for numerical aggregation. In general, propositionalization approaches may outperform ILP or MRDM systems, as suggested before in the literature [4], [20]. The choices of aggregation methods and parameters also have significant effects on the results in noisy real-world domains [9]. [12] have conducted a comparative evaluation of approaches to Boolean and numeric aggregation in propositionalization; however their results are inconclusive. In contrast, [16], [9] find that logic-based relational learning and logic-based (binary) propositionalization perform poorly in a noisy domain compared to numerical propositionalization.

Distance-based methods [8], [16] are another variant of relational learning. Their central idea is that it is possible to compute the mutual distance [7] for each pair of object for clustering [3], [15]. Probabilistic Relational Models (PRMs) [6], [10] provide another approach to relational data mining that is grounded in a sound statistical framework. [6], [10] introduce a model that specifies for each attributes of an object, its (probabilistic) dependence on other attributes of that object and on attributes of related objects. Propescul et al. [18] propose a combined approach called Structural Logistic Regression (SLR) that combines relational and statistical learning. Database numeric aggregation [9] techniques propose a method in which aggregation is done by using some of the built-in functions of common relational database system such as *count*, *min*, *max*, *sum*, *avg* and *exist*. Another approach proposed by [17] uses vector distances for dimensionality reduction and is capable of aggregating high-dimensional categorical attributes that traditionally have posed a significant challenge in relational modeling.

3 Decision Support and Hierarchical Multi-Attribute Model

3.1 Decision Support

The term "*decision support*" has a variety of meanings depending on the context on how you use it. Marko [2] outlined the literature review of decision support in details. Decision support can be categorized into *human decision sciences* [24] and *machine decision-making* [23]. [24] defines *human decision sciences* as an interdisciplinary field which addresses three possibly overlapping aspects of human decision making: normative, descriptive and decision support itself. There are some other definitions of decision support that focus on specialized disciplines (Fig. 2), such as operations research and management science [25], decision analysis [26], decision support systems [23], and others including data warehousing [27], group decision support systems [23],[28] and computer-supported cooperative work. Decision analysis introduced by [26] applied decision theory. Decision analysis provides a framework for analysing decision problems by structuring and breaking them down into more manageable parts, and explicitly considering the possible alternatives, available information, uncertainties involved and relevant preferences. In [26], Clemen introduces three models in decision-making, which are *influence diagram*, *decision tree* and *multi-attribute* models.

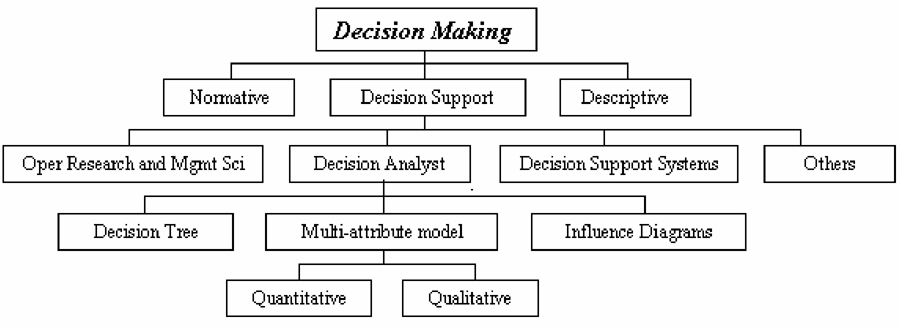


Fig. 2. Decision Making

3.2 Multi-Attribute Modeling

In principle, a multi-attribute model (MAM) [26] represents a decomposition of a decision problem into smaller and less complex sub-problems. A model consists of *attributes* and *utility functions*, as shown in Fig.3. *Attributes* are variables corresponding to decision sub-problems and all attributes at the leaf are basic attributes and attribute at the node is aggregate attribute. *Utility functions* define the relationship between the attributes at different levels in the tree and they serve for the aggregation of partial sub-problems into the overall evaluation or classification of options.

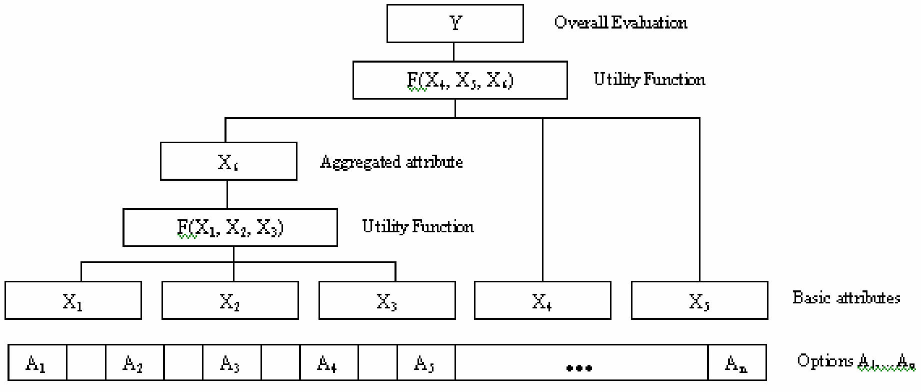


Fig. 3. Components of a Multi-Attribute model

The overall evaluation (utility) of an option is finally obtained as the value of one or more root attributes (Y) in Fig. 3. There are two types of MAM: *quantitative decision model* and *qualitative decision model*. In *Quantitative decision model*, all attributes are continuous and the utility functions are typically defined in term of attributes' weights, such as a weighted average of lower-level attributes [29], [30], [31]. In contrast, in *qualitative decision model*, all attributes are either nominal or ordinal [2], whose values are usually string values rather than numbers and the utility functions

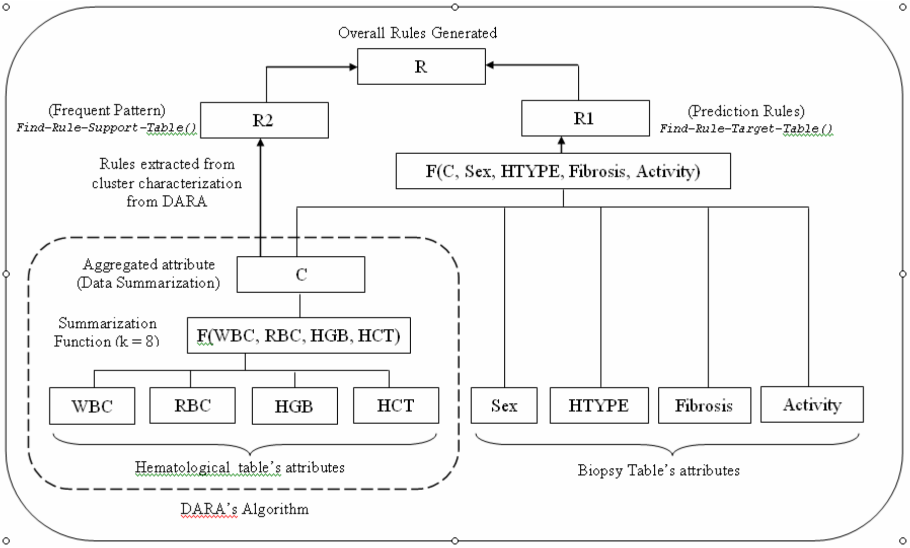


Fig. 4. Components of Multi-Attribute Decision Model for Hepatitis Datasets

use clustering functions to summarize data. This paper emphasizes the *qualitative decision model* in constructing decision making model. Fig. 4 illustrates how learning relational domains using dynamic aggregation based on patterns' distance (DARA) algorithm provides a concrete foundation for bridging relational data mining and MAM. DARA algorithm (Fig. 5) uses data summarization as the utility function to automate the construction of multi-attribute model to support decision-making.

4 Pattern-Based Aggregation

A common method to aggregate a single categorical attribute with numerous patterns is the selection of a subset of pattern that appears most often or distribution based approach. In this approach, each record (row) is viewed as a vector whose dimensions correspond to patterns occurrence in the target table stored in relational domain. Each pattern will have it's own component magnitudes. The component magnitudes are the *pf-irf* weights, as describes in (1), of the patterns which is adapted from *tf-idf* weights [12].

$$pf\text{-}irf = pf(p, r) \cdot irf(p) \tag{1}$$

$$irf(p) = \log \frac{|R|}{rf(p)} \tag{2}$$

$$sim(r_i, r_j) = \frac{r_i \cdot r_j}{||r_i|| \cdot ||r_j||} \tag{3}$$

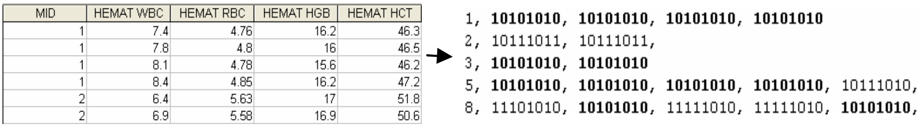
Pf-irf is the product of *pattern frequency Pf(p, r)*, and the *inverse record frequency* (2). *Pf(p, r)* refers to the number of times pattern *p* occurs in the corresponding record *r*. In (2), |*R*| is the number of records in the table and *rf(p)* is the number of records in which pattern *p* occurs at least once. Therefore, given two vectors (records) with component magnitudes described in (1), the similarity between two records is then computed in (3), where r_i and r_j are vectors with *pf-irf* coordinates as described above. Aggregation can be defined as a summarization of the underlying pattern or distribution from which the related objects were sampled. Once we compute the *pf-irf* weights, then we can compute the distance between each record and cluster them based on their weights. By grouping them into clusters or segments (validated by [1]), we are generalizing or aggregating them based on the underlying pattern or distribution from which the related objects were sample.

```
Input: A relational database
Output: a set of rules distinguish class label.
Procedure:
   Rule set R = empty
   Create-Pattern();
   Compute-Similarity-And-Transform()
   Update-Target-Table()
   Rule r1 = Find-Rule-Target-Table()
   Add r1 to the R.
   Rule r2 = Find-Rule-Support-Table()
   Add r2 to the R
   Return R
End Procedure
```

Fig. 5. Dynamic Aggregations of Relational Attributes Algorithm (DARA)

The process of data summarization is done using DARA's algorithm (Fig. 5) through the data generalization. The generalization task is done by converting each record's measurement into patterns, (in `Create-Pattern()` from Fig. 4). Then, data summarization is done for each record, by first computing the *pattern-frequency* and *inverse-record frequency* (2) and then grouping them based on the distance between records (3). This individual-centered concept, in which all rows belonging to a specific record is considered as a pattern that characterizes the individualism of each record. For instance, Fig. 6 depicts the summarization process for each record using (1) and (3). Firstly, each record is characterized by patterns of WBC, RBC, HGB and HCT measurements and they are converted into binary codes (01 = below normal, 10 = normal, 11 = above normal). Then, using (1), we compute the magnitude weight for each pattern. For example, given p = 10101010 and r = 1, pf-irf(10101010,1) = 4 x log (5/4) = 0.387. All records are then clustered (using `Compute-Similarity-And-Transform()`in Fig. 5), based on the records' component magnitudes (1). The component magnitude for each pattern is computed repeatedly for all records.

The set of overall rules, R, (in Fig. 4) is obtained from the combinated set of rules, R1 and R2. R1 is obtained by using the function *`Find-Rule-Target-Table()`*, where existing attribute-value classifiers such as C4.5, Conjunctive Rules and Naïve Bayes are applied. We use the *Weka* software [21] to extract R1. R2 is induced by using *`Find-Rule-Support-Table()`* as shown in Fig. 5 that describes each cluster/segment/



MID	HEMAT WBC	HEMAT RBC	HEMAT HGB	HEMAT HCT
1	7.4	4.76	16.2	46.3
1	7.8	4.8	16	46.5
1	8.1	4.78	15.6	46.2
1	8.4	4.85	16.2	47.2
2	6.4	5.63	17	51.8
2	6.9	5.58	16.9	50.6

Hematological Table

```
1, 10101010, 10101010, 10101010, 10101010
2, 10111011, 10111011,
3, 10101010, 10101010
5, 10101010, 10101010, 10101010, 10101010, 10111010,
8, 11101010, 10101010, 11111010, 11111010, 10101010,
```

Encoded Patterns represent a record

Fig. 6. Data Generalization for *one-to-many* relationship

group by finding pattern that has the maximum component magnitude for each cluster. The result of our experiment in Hepatitis dataset is discussed in the next section, in which rules generated from the HMAM using DARA is more efficient in terms of the percentage of correctly classified instances. In Fig. 4, we illustrate the integration of the relational learning algorithm, DARA, and HMAM in supporting the construction of decision support system.

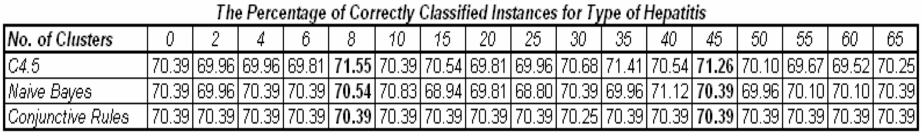
5 Experimental Results on Hepatitis Dataset

The database was collected at Chiba University hospital contains information on patients' exam dating from 1982 to 2001. Among the topics suggested by the Hepatitis dataset, we proposed to evaluate whether the level of biopsy activities and the type of hepatitis can be estimated based on laboratory tests namely WBC, RBC, HGB and HCT. These laboratory tests were chosen based on the work reported by [33]. The approach adopted here consisted of analyzing blood tests together with the biopsy results, seeking patterns that might indicate a correlation between the patients' exam results and the degree of their activities and also the type of their hepatitis.

The accuracy estimates from *10-fold cross validation* result shown in Table 1 using the k-means clustering. In Table 1, the percentage of correctly classified instances for type of hepatitis increases significantly by 1.16% using *DARA algorithm using k-means clustering* when number of clusters is 8 or 45. In contrast, in Table 2, the percentage of correctly classified instances for Biopsy Activities increases significantly by 1.45% when number of clusters is 40 and 35.

Figure 7 and 8 depicts set of rules, R1, induced using C4.5 [21] for classifying the type of hepatitis and also the Biopsy Activity. For each case, we also get the data

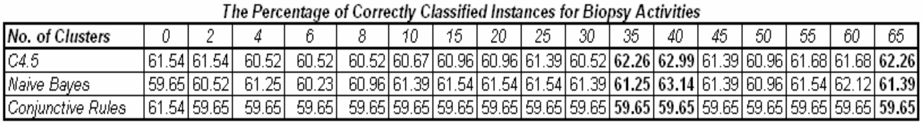
Table 1. Classifiers' performance for classifying Type of Hepatitis



The Percentage of Correctly Classified Instances for Type of Hepatitis

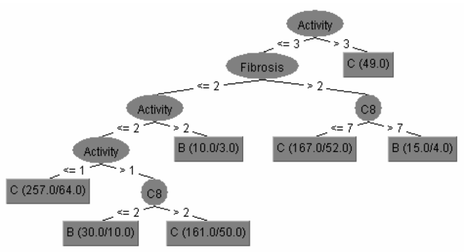
No. of Clusters	*0*	*2*	*4*	*6*	*8*	*10*	*15*	*20*	*25*	*30*	*35*	*40*	*45*	*50*	*55*	*60*	*65*
C4.5	70.39	69.96	69.96	69.81	**71.55**	70.39	70.54	69.81	69.96	70.68	71.41	70.54	**71.26**	70.10	69.67	69.52	70.25
Naive Bayes	70.39	69.96	70.39	70.39	**70.54**	70.83	68.94	69.81	68.80	70.39	69.96	71.12	**70.39**	69.96	70.10	70.10	70.39
Conjunctive Rules	70.39	70.39	70.39	70.39	**70.39**	70.39	70.39	70.39	70.39	70.25	70.39	70.39	**70.39**	70.39	70.39	70.39	70.39

Table 2. Classifiers' performance for classifying Biopsy Activities



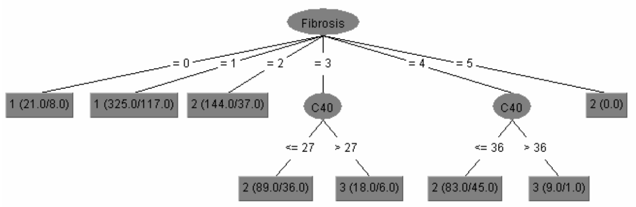
The Percentage of Correctly Classified Instances for Biopsy Activities

No. of Clusters	*0*	*2*	*4*	*6*	*8*	*10*	*15*	*20*	*25*	*30*	*35*	*40*	*45*	*50*	*55*	*60*	*65*
C4.5	61.54	61.54	60.52	60.52	60.52	60.67	60.96	60.96	61.39	60.52	**62.26**	**62.99**	61.39	60.96	61.68	61.68	**62.26**
Naive Bayes	59.65	60.52	61.25	60.23	60.96	61.39	61.54	61.54	61.54	61.39	**61.25**	**63.14**	61.39	60.96	61.54	62.12	**61.39**
Conjunctive Rules	61.54	59.65	59.65	59.65	59.65	59.65	59.65	59.65	59.65	59.65	**59.65**	**59.65**	59.65	59.65	59.65	59.65	**59.65**



```
Activity <= 3
|  Fibrosis <= 2
|  |  Activity <= 2
|  |  |  Activity <= 1: C (257.0/64.0)
|  |  |  Activity > 1
|  |  |  |  C8 <= 2: B (30.0/10.0)
|  |  |  |  C8 > 2: C (161.0/50.0)
|  |  Activity > 2: B (10.0/3.0)
|  Fibrosis > 2
|  |  C8 <= 7: C (167.0/52.0)
|  |  C8 > 7: B (15.0/4.0)
Activity > 3: C (49.0)
```

Fig. 7. R1 obtained using C4.5 for the type of Hepatitis (k = 8)

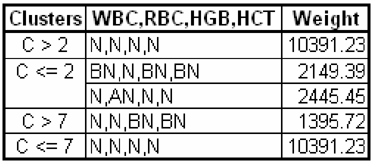


```
Fibrosis = 0: 1 (21.0/8.0)
Fibrosis = 1: 1 (325.0/117.0)
Fibrosis = 2: 2 (144.0/37.0)
Fibrosis = 3
|  C40 <= 27: 2 (89.0/36.0)
|  C40 > 27: 3 (18.0/6.0)
Fibrosis = 4
|  C40 <= 36: 2 (83.0/45.0)
|  C40 > 36: 3 (9.0/1.0)
Fibrosis = 5: 2 (0.0)
```

Fig. 8. R1 obtained using C4.5 for the Fibrosis Activity (k = 40)

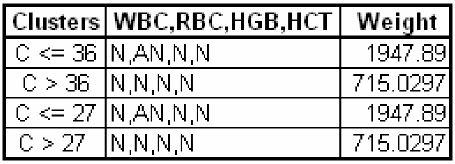
summarization (R2) for each cluster as shown in Table 3 and 4, where N = Normal, AB = Above Normal, and BN = Below Normal.

Table 3. Type of Hepatitis



Clusters	WBC,RBC,HGB,HCT	Weight
C > 2	N,N,N,N	10391.23
C <= 2	BN,N,BN,BN	2149.39
	N,AN,N,N	2445.45
C > 7	N,N,BN,BN	1395.72
C <= 7	N,N,N,N	10391.23

Table 4. Level of Biopsy Activities



Clusters	WBC,RBC,HGB,HCT	Weight
C <= 36	N,AN,N,N	1947.89
C > 36	N,N,N,N	715.0297
C <= 27	N,AN,N,N	1947.89
C > 27	N,N,N,N	715.0297

Below we summarize the findings based on R1 and R2 for classifying the type of hepatitis in Table 5 and classifying the activities of virus in Table 6.

Table 5. Finding for Classifying type of Hepatitis

TYPE	FINDINGS
Hepatitis C	a) Fibrosis level is F2 or lower and the activity of virus is A1, b) Fibrosis level is F2 or lower and the activity of virus is A2 with WBC, RBC, HGB and HCT at normal level c) Fibrosis level is greater than F2 and the activity of virus is A3 or lower with WBC, RBC, HGB and HCT at normal level d) Fibrosis level is greater than F2 with WBC, RBC, HGB and HCT at normal level
Hepatitis B	a) Fibrosis level is F2 or lower and the activity of virus is A3 or greater, b) Fibrosis level is F2 or lower and the activity of virus is A2 with WBC, HGB and HCT at below normal and RBC at normal level c) Fibrosis level is greater than F2 with WBC, RBC at normal level but HGB and HCT at below normal level

Table 6. Finding for Classifying Level of Virus Activities

LEVEL	FINDINGS
1	a) Fibrosis level is F0 and F1
2	a) Fibrosis level is F2 b) Fibrosis level is F3 or F4 with WBC, HGB and HCT at normal level and RBC at above normal level
3	a) Fibrosis level is F3 or F4 with WBC, RBC, HGB and HCT at normal level

6 Conclusion and Future Works

In this paper, we propose Dynamic Aggregation of Relational Attributes (DARA), an efficient approach to learning relational domain and integrate DARA with the HMAM to support the modeling of decision support for Hepatitis dataset. The results revealed that *DARA algorithm* generates rules that improve the performance of the classifier. There are some other techniques that can be used to perform the transformation such as Self Organizing Map (SOM) technique. SOM is very effective to be used when we have a lot of missing data and this could improve the transformation-based approach in multi-relational domain. In the future, we would proceed to validate the clinical reasonability of the results and validate the usefulness of the system on other datasets.

References

1. J.C. Bezdek. Some new indexes of cluster validiy, *IEEE Trans. Syst*., Man, Cybern. B, vol. 28, pp. 301-315, 1998
2. B. Marko. 2001. *Decision Support*. In D. Mladenic, N. Lavrač, Bohanec, M., and Moyle, S. 2003. Data Mining and Decision Support: Integration and Collaboration, Kluwer Aca. Publishers.
3. W. Dillon and M. Goldstein. *Multivariate analysis*, pages 157-208. John Wiley and Sons, Chichester, 1984.
4. S. Džeroski, H. Blockeel, B. Kompare, S. Kramer, B. Pfahringer, W. Van Laer, *Experiments in Predicting Biodegradability*, In Proceedings of ILP '99, 1999
5. S. Džeroski and N. Lavrač, editors. *Relational Data mining*. Springer-Verlag, 2001. ISBN 3540422897.
6. L. Getoor, N.Friedman, D. Koller, and A. Pfeffer. Learning Probabilistic relational models. In S. Džeroski and N. Lavrač, editors. *Relational Data mining*. Springer-Verlag, 2001.
7. T. Horvath, S. Wrobel, and U. Bohnebeck. Relational instance-based learning with lists and terms. *Machine Learning*, 43(1/2): 53-80, 2001.
8. M.Kirsten, S. Wrobel and T. Horvath. Distance based approaches to relational learning and clustering. In S. Džeroski and N. Lavrač, editors. *Relational Data mining*. Springer-Verlag, 2001.
9. A. Knobbe, M. De Haas, and A. Siebes. Propositionalization and aggregates. In *LNAI*, volume 2168, pages 277-288, 2001.
10. D. Koller and A. Pfeffer. Probabilistic frame-based systems. In *AAAI/IAAI*, pages 580-587, 1998.
11. S. Kramer, N. Lavrač and P. Flach. Propositionalization approaches to relational data mining. In S. Džeroski and N. Lavrač, editors. *Relational Data mining*. Springer-Verlag, 2001. ISBN 3540422897.

12. M.A. Krogel, S. Rawles, F. Železny, P.A. Flach, N. Lavrač, and S.Wrobel. Comparative evaluation of approaches to propositionalization. In *13th International Conference on Inductive Logic Programming* (ILP), pages 197-214, 2003.
13. S.H. Muggleton and L. DeRaedt. Inductive Logic programming: Theory and Methods. *The Journal of Logic Programming*, 19 & 20:629-680, May 1994.
14. S.H. Muggleton. Inverse Entailment and Progol. *New Generation Computing*, 13:245-286, 1995.
15. J. McQueen. Some Methods of classification and analysis of multivariate observations. In *Proceedings of Fifth Berkeley Symposium on Mathematical Statistics and Probability*, pages 281-293, 1967
16. C. Perlich and F. Provost. Aggregation-based feature invention and relational concept classes. In Proceedings *of the Ninth ACM International Conference on Knowledge Discovery and Data Mining (KDD)*, 2003.
17. C. Perlich and F. Provost. ACORA: Distribution-based aggregation for relational learning from identifier attributes. *Journal of Machine Learning*, 2005.
18. A.Propescul, L. H. Ungar, S. Lawrence, and D. M. Pennock. Structural Logistic Regression: Combining relational and statistical learning. In *Proceedings of the workshop on Multi-Relational Data Mining* (MRDM-2002), pages 130-141. University of Alberta, Edmonton, Canada, July 2002
19. A. Srinivasan and R.D. King. Feature Construction with Inductive Logic Programming: A Study of Quantitative Predictions of Biological Activity Aided by Structural Attributes. *Data Mining and Knowledge Discovery*, 3(1):37-57, 1999.
20. A. Srinivasan, R.D. King, D.W. Bristol, An Assessment of ILP-Assisted Models for Toxicology and the PTE-3 Experiment, In Proceedings of ILP '99, 1999
21. I. Witten, and E. Frank. 1999. *Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations*. Morgan Kaufman.
22. G. Salton, J. Michael, McGill, *Introduction to Modern Information Retrieval*, McGraw-Hill, Inc., New York, NY, 1986.
23. D.J. Power, 1999. *Decision Support Systems Glossary*, http://DSSResources.COM/ glossary/
24. [INSEAD, 2003] INSEAD, 2003. *Decision Sciences*. PhD Program Description, http://www.insead.edu/phd/program/decision.htm
25. F.S. Hillier and G.J. Lieberman, 2000. *Introduction to Operation Research*, McGraw Hill.
26. R. T. Clemen, 1996. Making Hard Decisions: An introduction to Decision Analysis, Duxbury Press
27. J. Han and M. Kamber, 2001. Data Mining: Concept and Techniques, Morgan Kaufman.
28. E.G. Mallach, 1994. Understanding Decision Support Systems and Expert Systems, Irwin, Burr Ridge.
29. DAS, 2001. *Decision Analysis Software*. http://faculty.fuqua.duke.edu/daweb/dasw.htm
30. H.L.S. Younes, 2001. *Current tools for assisting intelligent agents in real-time decision making*, MSc Thesis,http://www-2.cs.cmu.edu/~lorens/papers/mscthesis.html
31. G. Parmigiani, 2002. *Modelling in Medical Decision Making: A Bayesian Approach.* John Wiley & Sons, Ltd.
32. R. Agrawal and R. Srikant, Fast algorithms for mining association rules. In Proc. of the International Conference on Very Large Databases, Santiago de Chile, Chile, 1994
33. T. Watanabe, H. Suzuki, and L. Takabayashi. Application of prototypeline to chronic hepatitis data. In Working core of ECML/PKDD 2003 Discovery Challenge, p.166-177, 2003.