Unveiling Hidden Dependencies with Rough Sets Methods

Sylvia Encheva

Department of EBNS, Stord/Haugesund University College,
Bjørnsonsg. 45, 5528 Haugesund, Norway

Abstract* — *Many researchers consider various ways for early detection of students who are likely to experience serious difficulties in their studies. Some of them focus anxiety related problems connected with exam and general performance, while other concentrate on particular subjects' associated ones. Mathematical subjects appear to be among the ones causing problems for engineering students. Some of these problems are related to thinking logically, communicating mathematical arguments and conclusions, understanding abstract concepts as well as overcoming technical difficulties encountered when studying new topics. In this work we apply methods from rough sets theory for drawing conclusions from inconsistent datasets obtained from students' tests results. Decision rules are visually represented with flow graphs.

Keywords — *Decision making, rough sets, inconsistent data, learning*

I. Introduction

One of the serious problems in engineering education is learning calculus in general and obtaining sufficient skills for solving integrals in particular. Early detection of learners who are likely to experience difficulties during a course in calculus will provide opportunities for timely reactions and consequently preventing them from failure or even drop out. Analysing students test results and comparing them with their final exam results reveals some useful dependences that can be act upon. In this work we apply methods from rough sets theory for drawing conclusions from inconsistent datasets. Rough sets were originally introduced in [8] and further developed in [9], [10], and [11] to name but a few. The presented approach provides exact mathematical formulation of the concept of approximatively (rough) equality of sets in a given approximation space. An approximation space is a pair (U,Θ) where U is a finite set and Θ is an equivalence relation on U.

II. Related Work

Consider an information system A=(U,A) where information about an object x∈U is given by means of some attributes from A, i.e., an object x can be identified with the so-called signature of x:

Inf(x)=a(x):a∈A.

Suppose we are given two finite, non-empty sets U and A, where U is the universe of objects, cases, and A-a set of attributes, features. The pair (U, A) is called an information table. If we distinguish in an information system two disjoint classes of attributes, called condition and decision attributes, respectively, then the system will be called a decision table and will be denoted by S=(U,C,D), where C and D are disjoint sets of condition and decision attributes, respectively, [4],[5].

The number

$$supp_x(C,D) = |A(x)| = |C(x) \cap D(x)|$$

will be called a $support$ of the decision rule $C \rightarrow_x D$ and the number $\sigma_x(C,D) = \frac{supp_x(C,D)}{|U|}$, will be referred to as the strength of the decision rule $C \rightarrow_x D$, where $|X|$ denotes the cardinality of X.

With every decision rule $C \rightarrow_x D$ we associate the $certainty\ factor$ of the decision rule, denoted $cer_x(C,D)$' and defined as follows:

$$cer_x(C,D) = \frac{|C(x) \cap D(x)|}{|C(x)|} = \frac{supp_x(C,D)}{|C(x)|} = \frac{\sigma_x(C,D)}{\pi(C(x))},$$

where $\pi(C(x)) = \frac{|C(x)|}{|U|}$.

The certainty factor may be interpreted as a conditional probability that y belongs to $D(x)$ given y belongs to $C(x)$, symbolically $\pi_x(D \mid C)$. If $cer_x(C,D) = 1$, then $C \rightarrow_x D$ will be called a $certain\ decision$ rule; if $0 < cer_x(C,D) < 1$ the decision rule will be referred to as an $uncertain\ decision$ rule. A $coverage\ factor$ of the decision rule is

$$cov_x(C,D) = \frac{|C(x) \cap D(x)|}{|D(x)|} = \frac{supp_x(C,D)}{|D(x)|} = \frac{\sigma_x(C,D)}{\pi(D(x))},$$

where $\pi(C(x)) = \frac{|D(x)|}{|U|}$, and similarly is defined as

$$cov_x(C,D) = \pi_x(C \mid D).$$

If $C \rightarrow_x D$ is a decision rule then $D \rightarrow_x C$ will be called an $inverse\ decision\ rule$. The inverse decision rules can be used to give explanations (reasons) for a decision.

With every decision table we associate a *low graph*, i.e., a directed acyclic graph defined as follows: to every decision rule $C \rightarrow_x D$ we assign a directed branch x connecting the input node $C(x)$ and the output node $D(x)$. Strength of the decision rule represents a through flow of the corresponding branch, [8], [9].

Web learners clustering model based on rough set theory is presented in [17]. The authors aim at developing personalized teaching strategies for distance learning. Another article applying clustering algorithms based on rough set theory is discussing learning features of network education learners, [1]. A concept map for each student and staff using rough sets and data mining is developed in [12]. An attempt to predict students' future performance can be seen in [13]. Rough set theory is applied in a ranking process in [14] and as an assessment method in [15], [16]. The human side of decision making is presented in [3] and [6]. Uncertain data is handled applying decision three methods in [2].

III. ESTABLISHING DEPENDENCIES

Our goal is to find dependences between students' skills to solve integrals and their abilities to solve mathematical problems requiring knowledge about fractions, trigonometrical functions, logarithms and differentiation. Students take Web-based tests containing problems from the above mentioned areas. Our goal is to find dependences between students' skills to solve integrals and their abilities to solve mathematical problems requiring knowledge about fractions, trigonometrical functions, logarithms and differentiation. Students take Web-based tests containing problems from the above mentioned areas. Results are classified as low (l) if the level of correctness is less than 40%, medium (m) if the level of correctness is more than 40% and less than 70%, and high if the level of correctness is more than 70%. The decision attribute 'Integration' should be read as 'can the student solve integrals?'. For the sake of simplicity we use only two notations yes (y) and not (n). Students are divided in groups (1,2,...,12) according to gender and results from a final exam in calculus. Students not taking tests are not included in this research due to the high level of uncertainty related to their knowledge. Rules obtained from groups in Table I are shown in Table II.

Groups	Fractions	Trig. functions	Logarithms	Differentiation	Integration	N
1	h	l	m	h	y	46
2	h	m	m	m	y	58
3	m	l	m	m	y	132
4	h	m	l	h	y	86
5	l	m	l	m	n	27
6	l	m	m	l	n	43
7	h	l	m	l	n	15
8	h	l	m	l	y	62
9	h	h	m	m	y	53
10	m	l	h	h	y	49
11	m	l	m	m	n	36
12	h	m	l	l	n	29

TABLE I
Students Result

TABLE III
Rules

Groups	Strength	Certainty	Coverage
1	0.07	1.0	0.09
2	0.09	1.0	0.12
3	0.21	0.79	0.28
4	0.14	1.0	0.18
5	0.04	1.0	0.17
6	0.07	1.0	0.29
7	0.02	0.19	0.08
8	0.09	0.81	0.12
9	0.08	1.0	0.12
10	0.08	1.0	0.11
11	0.06	0.21	0.25
12	0.05	1.0	0.21

Such results can be used for early detection of eventual failure to learn integration. Once it is established which type of knowledge is insufficient an individual help will be provided in forms of additional explanations and examples.

A decision algorithm in an information system composed of condition and decision attributes is a set of mutually exclusive and exhaustive decision rules, involving all data in the system and preserving the existing indiscernibility relation in the system, [7]. A decision algorithm extracted from Table II is presented in Table III.

TABLE IIIII
Decision Algorithms

Rules	If	Then	Certainty
1	Differentiation, h	Integr., y	1.0
2	Differentiation, m	Integr., y	0.79
3	Fractions, h and Trig. functions, h	Integr., y	1.0
4	Fractions, h and Trig. functions, m	Integr., y	0.83
5	Logarithms, m and Differentiation, m	Integr., y	0.87
6	Differentiation, l	Integr., n	0.58
7	Fractions, l	Integr., n	1.0
8	Logarithms, m and Differentiation, l	Integr., n	0.48

An inverse decision algorithm can be understood as an explanation of ability to solve integrals or as lack of such ability.

Decision tables can be used for drawing conclusions from a dataset as well as for interpreting the outcomes. Availability of sufficient skills for solving integrals is implied by: obtained high level skills in differentiation, and obtained high level skills in working with fractions and trigonometrical functions.

Lack of sufficient skills for solving integrals are implied by low level skills in working with fractions.

TABLE IVV
Inverse Decision Algorithm

Rules	If	Then	Certainty
1	Integr., y	Differentiation, h	0.37
2	Integr., y	Differentiation, m	0.50
3	Integr., y	Fractions, h and Trig. functions, h	0.11
4	Integr., y	Fractions, h and Trig. functions, m	0.30
5	Integr., y	Logarithms, m and Differentiation, m	0.50
6	Integr., n	Differentiation, l	0.41
7	Integr., n	Fractions, l	0.47
8	Integr., n	Logarithms, m and Differentiation, l	0.39

From the contents of Table 4 one can deduce that: the most probable reason for availability of sufficient skills for solving integrals is having medium or high level skills in differentiation, and the most probable reason for lack of sufficient skills for solving integrals having low level skills in differentiation.

With every decision table we associate a *flow graph*, i.e., a directed acyclic graph defined as follows: to every decision rule $C \rightarrow_x D$ we assign a directed branch x connecting the input node $C(x)$ and the output node $D(x)$. Strength of the decision rule represents a through flow of the corresponding branch, [8]

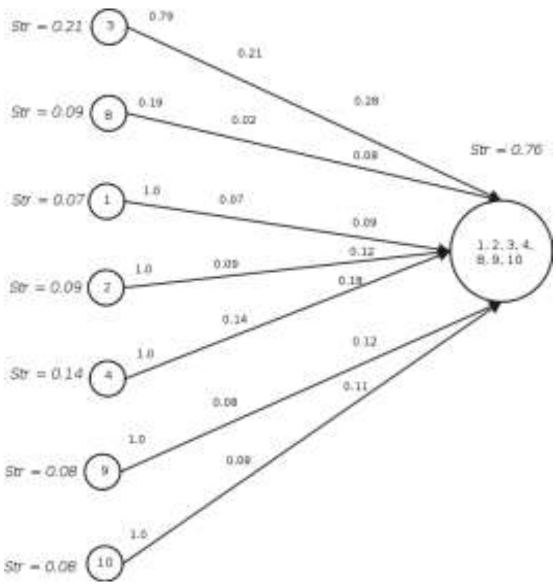

Fig. 1. Flow graph for groups of students able to solve integrals

The values for certainty, strength and coverage for decision rules 1, 2, 3, 4, 5, 7, 9, 11 and 12 are written above the connecting arrows while the corresponding values for decision rules 6, 8, 10, and 12 are written under the connecting arrows, Fig.1 and Fig.2. Strength of rules in a node is denoted by *Str*.

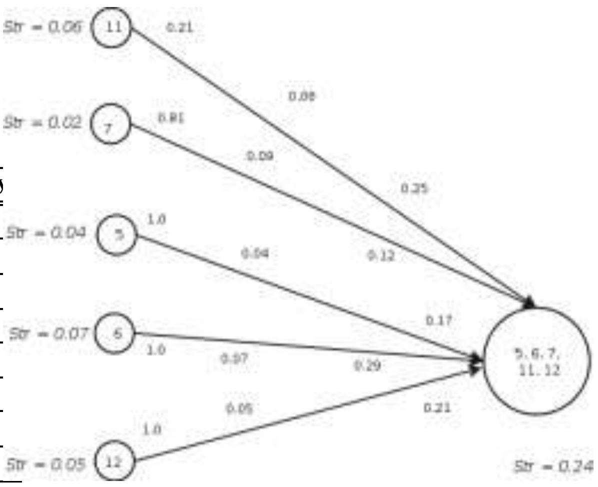

Fig. 2. Flow graph for groups of students not able to solve integrals

Classification of objects in this representation results in finding the maximal output flow in the flow graph, whereas explanation of decisions is connected with the maximal input flow associated with the given decision, [8].

IV. Conclusions

Rough set theory has been successfully applied for working with imprecise and or inconsistent data. The current work is one of the many in that line. Interesting dependences between students' results from preliminary tests and their later learning outcomes have been derived. Such findings can be used while considering which of the new students are likely to experience serious difficulties in completing a subject or a study.

References

[1] I. Duntsch and G. Gediga, *Rough set data analysis: A road to non-invasive knowledge discovery*, Methodos Publishers, 2000.

[2] A. M. Bhavitha and S. Madhuri, A Classification Method using Decision Tree for Uncertain Data, *International Journal of Computer Trends and Technology*, 3(1), pp. 102-107, 2012

[3] A. Johnstone, *Effective Practice in Objective Assessment*, www.physsci.heacademy.ac.uk/Publications/PracticeGuide/

[4] J. Kacprzyk and M. Roubens, *Non Conventional Preference Relations in Decision Making*, Springer Verlag, LNMES, 301, Berlin, 1988.

[5] G. A. Klein, Recognition-Primed Decision Making. In Klein, G. A. (Ed), *Sources of power: How people make decisions*, Cambridge, MA: MIT Press, pp. 15–30, 1998.

[6] H. Liu and R. Setiono, A probabilistic approach to feature search selection - A filter solution. In: Kumar, S. (Ed.), 13th *International Conference on Machine Learning* (ICML'96), Bari, Italy, pp. 319–327, 1996.

[7] M. Inuiguchi, S. Tsumoto, and S. Hirano, *Rough Set Theory and Granular Computing*, Springer-Verlag Berlin and Heidelberg GmbH & Co. KG, 2010.

[8] Z. Pawlak, Rough Sets, *International Journal of Computer and Information Sciences*, 11, pp. 341–356, 1982.

[9] Z. Pawlak, *Rough Sets: Theoretical Aspects of Reasoning about Data*, vol. 9 Kluwer Academic Publishers, Dordrecht, The Netherlands, 1991.

[10] Z. Pawlak, Rough sets, decision algorithms and Bayes' theorem, *European Journal of Operational Research*, 136, pp. 181–189, 2002.

[11] Z. Pawlak, In Pursuit of Patterns in Data Reasoning from Data - The Rough Set Way, in J.J. Alpigini et al. (Eds.) RSCTC 2002, *LNAI* 2475, Springer-Verlag Berlin Heidelberg, pp. 1–9, 2002.

[12] P. Ramasubramanian, K. Iyakutti, P. Thangavelu, and J. J. Winston, Teaching result analysis using rough sets and data mining, *Journal of Computing*, 1(1), pp. 168–174, 2009.

[13] P. Ramasubramanian, K. Iyakutti, and P. Thangavelu, Enhanced data mining analysis in higher educational system using rough set theory, *African Journal of Mathematics and Computer Science Research*, 2(9), pp. 184–188, 2009.

[14] P. Ramasubramanian, K. Iyakutti, and P. Thangavelu, Ranking of Professional Colleges using Rough Sets and Data Mining, *Journal of Advanced Research in Computer Engineering*, 3(2), pp. 421-424, 2009.

[15] P. Ramasubramanian, S. B. Priya, and T. Dhanalakshmi, Study on Assessment Method for Technical Institutions Using Rough Sets, *International Journal of Research and Reviews in Information Sciences* (IJRRIS), 1(4), 2011.

[16] R. W. Swiniarski and A. Skowron, Rough set methods in feature selection and recognition, *Pattern Recognition Letters*, 24, pp. 833–849, 2003.

[17] Q. Zhiming and W. Xiaoli, Application of Clustering Algorithm and Rough Set in Distance Education, *First International Workshop on Education Technology and Computer Science*, 2009. ETCS '09, Wuhan, Hubei, pp. 489–493, 2009.