Analyzing Rule Evaluation Measures with Educational Datasets: A Framework to Help the Teacher

Sebastian Ventura, Cristobal Romero and Cesar Hervás
{sventura, cromero, chervas}@uco.es
Department of Computer Science and Numerical Analysis. University of Cordoba.
Campus de Rabanales, edificio "Albert Einstein", 14071, Córdoba, España.

Abstract. Rule evaluation measures play an important role in educational data mining. A lot of measures have been proposed in different fields that try to evaluate features of the rules obtained by different types of mining algorithms for association and classification tasks. This paper describes a framework for helping non-expert users such as instructors analyze rule evaluation measures and define new ones. We have carried out several experiments in order to test our framework using datasets from several Cordoba University Moodle courses.

1 Introduction

Rule discovery is one of the most popular data mining techniques, especially in EDM (Educational Data Mining), because it shows the teacher information that has been discovered and does so in an intuitive way [7]. The rules discovered by a data mining method must be of interest for end-users in order to be considered useful. Today, measuring the interest in the rules discovered is an active and important area of data mining research [2]. Its main objective is to reduce the number of mined rules by evaluating and post-pruning the rules obtained in order to locate only the most interesting rules for a specific problem. This is very important for a non-expert user in data mining, like a teacher. In order to evaluate rules there are a wide range of measures [8]. Rule evaluation measures are used for ranking and selecting rules according to their potential interest for the user [8]. However, as there is no single measure that is optimal in all application domains and some authors propose new measures, it is necessary to evaluate the performance of each measure for each different domain and/or dataset [2].

In the educational domain, the study of rule evaluation measures is a very important issue, although there are only a few papers about it. Ma et al. [4] select weak student ranking association rules by a scoring function that uses several measures. Minaei-Bidgoli et al. [6] discover interesting contrast rules using different evaluation measures. Romero et al. [7] select a group of rule evaluation measures to select the best prediction rules. Merceron and Yacef [5] revisit the measuring of degrees of interest in their strong symmetric association rules in educational data.

In this paper, we propose a framework for analyzing rule evaluation measures. The paper is organized as follows: section 2 surveys rule evaluation measures, section 3 describes the rule evaluation framework, and section 4 describes experimental tests using datasets from Moodle courses' student usage information. Finally, in section 5 the conclusions and further research are outlined.

2 Rule Evaluation Framework

Most of the general-purpose data mining tools such as Weka [9] discover (association, prediction, classification) rules and show them to the user together with traditional evaluation rule measures such as accuracy, support and confidence. But these tools do not show the values of other important rule evaluation measures [8], nor do they permit their comparison. Thus it is very hard for a user that is not an expert in data mining (e.g. a teacher) to appropriately use the great variety of existing rule evaluation measures to select only the most interesting rules. In order to try to resolve this problem, we have developed a rule evaluation framework (see Figure 1).
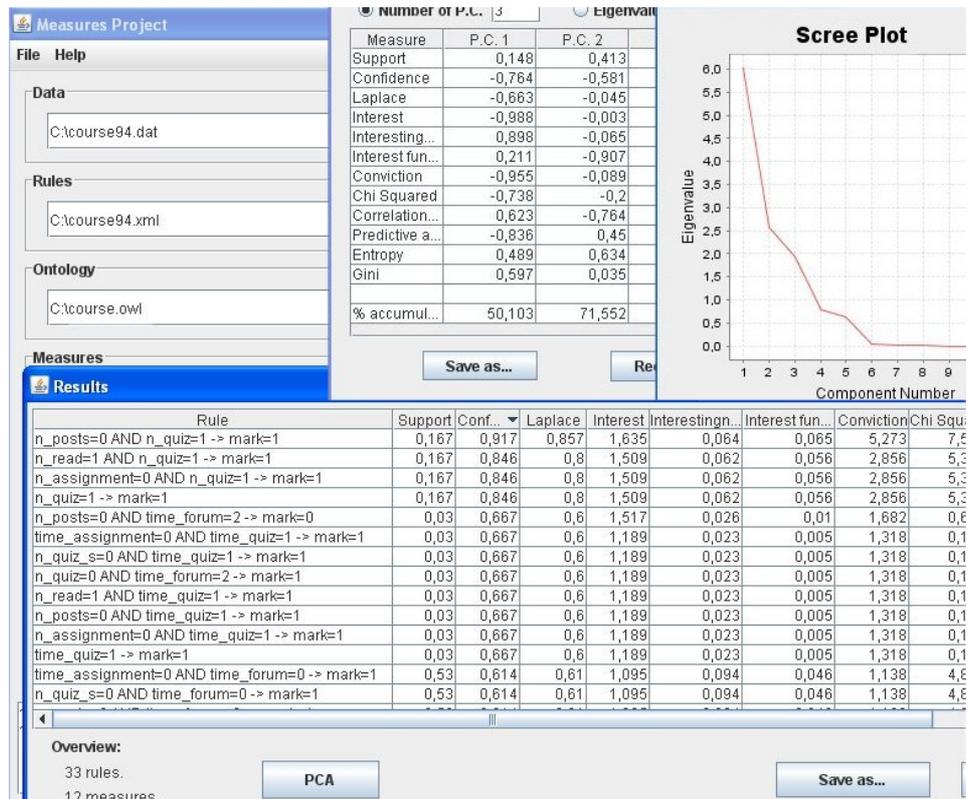
Fig. 1. Main window of the application and windows with the rules and values of the measures.

In order to use our framework, the user/teacher has to select 2 obligatory files (data and rules) and can also select 2 optional files (measures and ontology). The data file is a local file or URL (Uniform Resource Locator) that contains the dataset. This file can have CSV (Comma-Separated Values) format or Weka format [9]. The rules file is also a local file or URL that contains the rules discovered by the mining algorithm. It has to have PMML (Predictive Modeling Markup Language) format that is an XML-based language which provides a way for applications to define statistical and data mining models and to share models between compliant applications. The measure description file is an optional local file with XML format that contains the definition of the measures in Latex equation format. Thus the measures are defined using not only mathematics latex symbols and functions, but also probabilistic and contingency table symbols, such as n(A), n(C), N,

P(A), P(C/A), etc. as well as all the availables measures for defining new measures. The framework provides a default measure description file with over 40 measures already defined so that the teacher can use them directly with no need to define them. We have also developed a wizard and an equation editor using Latex equation format in order not to have to write out this XML file by hand, since this could be a difficult task for a teacher. Using this wizard, the teacher can define brand-new measures easily by only following several guided steps and the equation editor. Finally, the ontology file is an optional local file. It uses OWL (Ontology Web Language) format to define the specific domain of the data and rules used. Then, after the user/teacher has chosen the previous input files; the application calculates the values of each measure for each rule coming from the data file. Next, all the rules and all the measures are displayed for the teacher in the *results* window (see Figure 1 down). For each rule, the elements of rule antecedents and consequents are displayed as well as the calculated values of each evaluation measure. The rules can be sorted by any of the measures by simply clicking on the header of a specific column so that the teacher can compare different ranks depending on the measure used. Furthermore, if the user has selected an ontology file, then the OWL file will be visualized graphically so the teacher can interpret/understand better the meaning of the rules in that domain.

Finally, we have also developed a PCA (Principal Component Analysis) module [1] in order to help the user/teacher to group and reduce the number of measures used. Each principal component is unrelated and corresponds to orthogonal directions in the newly generated search space. In our framework, the teacher can execute PCA starting from the results windows by pressing the *PCA* button (see Figure 1 down). A new window appears (see Figure 1 up) in which the user/teacher can select the number of principal components or the maximum eigen value, and has the option of showing the scree plot and the coefficients of the measures in each PC. Then the communality values of each principal component for each measure are shown along with the scree plot. Using the scree plot and the eigenvalues, the user/teacher can select a number of principal components, normally those with eigenvalues greater than 1 or when the inclination of plot starts to decrease. Then, the teacher can group each measure into one principal component using the communality values for each measure. In order to do so, the teacher has to assign or classify each measure in the component where it shows the highest absolute value.

3   Experimental results with Moodle courses

We have carried out several experiments in order to test our framework using educational datasets. We have used 4 dataset files obtained from 4 Moodle courses with about 80 students in each course. The courses are computer science technical-engineering second-year courses in Cordoba University. We have preprocessed these students' usage data that is stored in a Moodle database. First, we have filtered only the information about the 4 courses activities that interest us, such as assignments, forums and quizzes. Next, we have created a summarization table that integrates this information at student level (number of assignments done, number of messages sent/read to/in the forum, number of quizzes taken/passed/failed, total time used on assignment/quiz/forum, and student's final mark). We have discretized all the numerical values in order to increase interpretation

and comprehensibility since categorical values (low, medium and high) are more familiar to teacher s than precise magnitudes and ranges. Finally, we have saved them in 4 data files, one for each course. Before using our framework, we applied the Apriori-C algorithm [3] to discover Class Association Rules (CARs) from previous Moodle data. Apriori-C is a well-known algorithm for discovering association rules for classification. In our experiment, the class is the mark attribute. In this way, the teacher can obtain rules that show relationships between Moodle activities that influence the mark obtained by the students [15]. Specifically, we have executed Apriori-C algorithm over the 4 course dataset files with a minimum support of 0.03 and a minimum confidence of 0.6 as parameters. Next, the class association rules obtained are saved into a PMML file. In the first column of Table 1, we can see the number of rules obtained for each dataset.

Table 1. Comparison of %Variance with the principal components using 12 measures.

Dataset	Number of rules	1 P.C.	2 P.C.	3 P.C.	4 P.C.
Course 1	33	50,103	71,552	87,693	94,235
Course 2	36	36,001	65.411	78,382	88,414
Course 3	32	34,383	60,486	78,961	87,443
Course 4	29	36,547	60,518	80,733	89,098

Next, we have used our framework to obtain values for all evaluation measures beginning with the Moodle datasets and PMML rules files. In this experiment, we have used only 12 measures: chi-squared, correlation coefficient, predictive association, entropy, support, confidence, laplace, interest, interestingness, gini, interest function and conviction. We have chosen these measures specifically because they are some of the most representatives [8]. So we have obtained the values of the 12 measures for all the rules and then we have applied PCA with 1, 2, 3 and 4 principal components in order to see how many components or groups all these measures can be grouped into. In Table 1, we can see the amount of variance obtained for each principal component from the rules discovered in each dataset. The results obtained show that we can select 3 principal components since they store between 80%-90% of the variance of the data. So, we can use these components as new evaluation measures in which almost all the information provided by the original measures is included. In this way we can reduce the number of measures used from the original 12 measures to 3 new meta-measures. The teacher could define three new measures using the editor, one for each PC using the coefficients of the measures in each PC**.** For example, the 2nd PC in course 1 could be defined as the following new measure:

$$PC2=0.565*Support-0.362*InterestFunction$$

4   Conclusions and Future Work

In this paper we have described a specific framework for analyzing rule evaluation measures. We have shown how a user/teacher can use it together with Moodle course datasets in order to: obtain the values of the measures of the rule discovered by a rule mining algorithm, sort and select the rules according to the values of any specific

measure, compare and group the measures using PCA and define new measures using a wizard and an equation editor. Currently we are working on other techniques for comparing rule evaluation measures. For example, correlation techniques in order to see what measures are correlated. In the future, we want to work with subjective and semantics-based measures. Our objective will be to add subjective restrictions to our framework user that take into account information about the domain. For example, restrictions to different granularity levels to only show rules about relationships between activities, or relations between chapters or between courses. Finally, we also want to develop brand-new semantic-based measures that can use the domain information of the OWL files. In this way, we could create new measures specifically geared toward each application domain or dataset.

References

[1] Fukugana, K.: *Introduction to Statistical Pattern Recognition*. Elsevier. 1990.

[2] Geng, L., Hamilton, H.J.: Interestingness Measures for Data Mining: A Survey. ACM Computing Surveys, Vol. 38, No. 3 (2006) 1-32.

[3] Jovanoski, V., Lavrac, N.: Classification Rule Learning with APRIORI-C. In Proc. Conf. Artificial Intelligence (2001) 44-51.

[4] Ma, Y., Liu, B., Kian, C.K. Wong, Yu, P.S., Lee, S.M.: Targeting the right students using data mining. In Proc. Conf. Knowledge discovery and data mining, Boston (2000) 457-464.

[5] Merceron, A., Yacef, K.: Revisiting interestingness of strong symmetric association rules in educational data. In Proc. of Int. Workshop on Applying Data Mining in e-Learning, Creete, Greece (2007) 3-12.

[6] Minaei-Bidgoli, B., Tan, P-N., Punch, W.F.: Mining Interesting Contrast Rules for a Web-based Educational System. In Proc. Int. Conf. on Machine Learning Applications. Louisville, USA (2004) 320- 327.

[7] Romero, C., Ventura, S. de Bra, P.: Knowledge Discovery with Genetic Programming for Providing Feedback to Courseware Author. UMUAI. Vol. 14. No. 5 (2004) 425-465.

[8] Tan, P., Kumar, V.: Interesting Measures for Association Patterns. In Proc. KDD Workshop on Postprocessing in Machine Learning and Data Mining, Boston, USA (2000) 1-9.

[9] Witten, I.H., Frank, E.: Data Mining. Practical Machine Learning Tools and Techniques with Java Implementations, Morgan Kaufmann (1999).