A Differential Approach for Identifying Important Student Learning Behavior Patterns with Evolving Usage over Time

John S. Kinnebrew[1](✉), Daniel L.C. Mack[2], Gautam Biswas[1], and Chih-Kai Chang[3]

[1] Department of EECS and ISIS, Vanderbilt University, Nashville, TN, USA
john.s.kinnebrew@vanderbilt.edu
[2] Baseball Analytics, Kansas City Royals, Kansas City, MO, USA
[3] Department of Information and Learning Technology, National University of Tainan, Tainan, Taiwan

Abstract. Effective design and improvement of dynamic feedback in computer-based learning environments requires the ability to assess the effectiveness of a variety of feedback options, not only in terms of overall performance and learning, but also in terms of more subtle effects on students' learning behavior and understanding. In this paper, we present a novel interestingness measure, and corresponding data mining and visualization approach, which aids the investigation and understanding of students' learning behaviors. The presented approach identifies sequential patterns of activity that distinguish groups of students (e.g., groups that received different feedback during extended, complex learning activities) by differences in both total behavior pattern usage and evolution of pattern usage over time. We demonstrate the utility of this technique through application to student learning activity data from a recent experiment with the Betty's Brain learning environment and four different feedback and learning scaffolding conditions.

Keywords: Interestingness measure · Sequence mining · Learning behaviors · Information gain

1 Introduction

In order to more effectively teach and promote skills required in the modern world of near-ubiquitous computing and internet connectivity, computer-based learning environments have become more complex and open-ended. This complexity also drives a need for dynamic, adaptive feedback and learning scaffolding that can support students in understanding how to employ and learn with these environments and tools. However, in order to effectively design and improve such feedback/scaffolding, we must first be able to assess the effectiveness of a variety of scaffolding options, not only in terms of overall performance and learning, but also in terms of more subtle effects on students' behavior and understanding.

W.-C. Peng et al. (Eds.): PAKDD 2014 Workshops, LNAI 8643, pp. 281–292, 2014.
DOI: 10.1007/978-3-319-13186-3_27

Identifying sequential patterns in learning activity data has provided a useful tool for discovering and better understanding such student learning behaviors. However, once these behavior patterns are mined, researchers must interpret and analyze the resulting patterns to identify a relevant subset of important patterns that provide a basis for generating actionable insights about how students learn, solve problems, and interact with the environment. Algorithms for mining sequential patterns generally associate a measure of pattern frequency in the data with the relative importance or ranking of the pattern for investigation. However, when analyzing students' learning behaviors, another important aspect of these patterns is the evolution of their usage over the course of a student's learning or problem-solving activities. Further, in order to understand how different feedback and scaffolding can affect learning behavior, the more important patterns are those that differ across different experimental groups or other categories of students.

To address this challenge, we present a data mining and visualization approach, combining traditional sequence mining and a novel information-theoretic interestingness measure for ranking behavior patterns that distinguish groups of sequences (e.g., groups of students in different experimental conditions) by differences in both total pattern usage and the evolution of pattern usage over time. This approach builds on the Temporal Interestingness of Patterns in Sequences (TIPS) analysis framework [1], but instead of basing interestingness on the extent of changes in usage over time as in TIPS, this approach incorporates the dimension of student/sequence groups to identify and visualize patterns that are used differentially across those groups (by either total frequency or relative usage patterns over time). To demonstrate the utility of this approach, we present case study results of mining student activity data from a recent experiment with the Betty's Brain learning environment. These results illustrate the effectiveness of our approach and suggest further refinements for temporal and differential analysis of sequential learning activity data.

2 Related Work

The canonical sequential pattern mining task is to discover sequential patterns of items that are found in many of the sequences in a given dataset [2,3]. Researchers have applied variations of sequence mining techniques to a wide range of educational data in order to better understand and improve learning. For example, researchers have employed sequential pattern mining to better understand specific aspects of student learning behavior, such as Nesbit *et al.* ([4]), who mined the longest common subsequences in data from the gStudy learning environment to research how students self-regulate as they learn. Others have employed sequence mining techniques to more directly scaffold and improve student learning. For example, Perera *et al.* ([5]) help student groups collaborating on software development to improve their work by observing and emulating the behaviors of the strong groups, which are determined through sequence mining. Kinnebrew *et al.* [6,7] compared sequential patterns mined from student activity sequences to identify patterns for targeted feedback that differ in frequency between student groups and between productive and unproductive periods of work. Other researchers have also employed sequential pattern mining to identify differences among student groups or generate student models to customize learning to individual students [8–10].

Sequence mining algorithms commonly rank the discovered patterns by their frequency in the dataset. Over time, researchers have developed additional measures to utilize properties other than just frequency to rank mined patterns [11]. These measures are often referred to as "interestingness measures" and have been applied to a variety of data mining tasks, such as sequence mining and association rule mining [12]. The measures can be simple calculations like accuracy, specificity, and recall, or more complicated ones like the Gini Index [13] and information gain [14], which identify patterns that possess more complex relationships to the data. In educational data mining applications, interestingness measures have been used to rank mined association rules (e.g., [15]). More generally, information gain has been employed with educational data to solve problems like identifying attributes and producing rules for predicting student performance [16].

In addition to an interestingness measure to identify potentially important patterns, our approach also employs a visualization of pattern usage over time with heat maps to more efficiently analyze identified patterns. Heat maps have commonly been employed in biomedical fields for tasks including visualization of gene expression [17] and identification of common gene locations. Heat maps have also seen use in other fields to trace dynamic changes over time, such as in the environment or in music habits, and often employ maps that contain non-regular shapes, especially those of real-world geography [18].

3 Identifying Interesting Differences in Evolving Pattern Usage

With long sequences of temporal data, such as student learning activities in a computer-based learning environment, researchers and analysts are not only interested in discovering frequent sequential patterns, but, in many cases, also need to analyze their occurrence over time and identify patterns that differ among groups of sequences. To address these needs, we present the Differential, Temporal Interestingness of Patterns in Sequences (D-TIPS) measure and approach for identifying and visualizing patterns that are employed differentially over time across groups of students (e.g., groups that receive different scaffolding in a learning environment). The first step in analyzing learning activity sequences is to define and extract the actions that make up those sequences from interaction traces logged by the environment. The definition of actions in these sequences for Betty's Brain data is discussed further in Sect. 4. Given a set of sequences corresponding to the series of actions performed by each student, the D-TIPS technique consists of four primary steps that extend the corresponding steps in TIPS [1] to identify and interpret differences among student/sequence groups instead of simply over time:

1. Generate candidate patterns that are common to many students in one or more groups through sequential pattern mining.
2. Calculate a temporal footprint for each candidate pattern by mapping it back to locations where it occurs in the activity sequences.
3. Provide a ranking of the candidate patterns using an information-theoretic interestingness measure applied to the temporal footprint of each pattern across sequence groups.
4. For the highly-ranked, differential patterns, visualize their temporal footprints using heat maps to compare trends and spikes in usage across groups.

3.1 Identifying Common Patterns

In order to identify candidate behavior patterns for investigation, we employ a sequential pattern mining algorithm to the set of student activity sequences. Standard sequential pattern mining produces a set of sequential patterns that meet a given support threshold (i.e., they occur in at least a given percentage of the sequences). Certain types of additional constraints can be imposed with some algorithms, but otherwise the choice of algorithm does not affect the resulting set of patterns. In this case, because we are interested in behavior patterns where the actions occur immediately or shortly following each other in the sequence, we employ an algorithm (from Pex-SPAM [19]) that allows constraints on the gaps between actions in the identified sequential patterns[1]. Further, to identify patterns common to the majority of the students in any group, we apply sequential pattern mining to each group of student sequences separately with a support threshold of 50 % and combine the resulting sets of patterns to identify the full set of candidate patterns.

Once the common behavior patterns are mined, researchers must interpret and analyze the resulting patterns to identify a relevant subset of important patterns that provide a basis for generating actionable insights (*e.g.*, how to scaffold user interactions with the learning environment to encourage specific, productive behaviors). A pattern's "frequency" in sequential pattern mining terms (*i.e.*, the number of sequences in which the pattern occurs) is often used to rank the importance of the identified patterns. Alternatively, the frequency of occurrence within sequences is a different frequency measure that can be more appropriate for long sequences [6]. For learning activity sequences, this occurrence frequency (*i.e.*, how often a pattern occurs *within* sequences) is generally of more interest than simply the number of students who employed a pattern at least once. D-TIPS relies on a consideration of pattern occurrence within sequences, but rather than simply ranking patterns by their occurrence frequency (e.g., average or median occurrence frequency among the students/sequences), it also incorporates information about how this frequency changes over time.

[1] In the results presented in Sect. 5, we allowed a maximum gap of one action, allowing up to one irrelevant or variable action between consecutive actions in the pattern. However, in general, other sizes of gaps or no gap at all may be appropriate depending on the data and goals of an analysis.

3.2 Calculating the Temporal Footprint

Given a set of candidate patterns from the sequential pattern mining, the next step is to map the patterns back to the activity sequences to define a temporal footprint for each pattern. Each sequence is divided into n consecutive slices, such that each contains $\frac{100}{n}$ % of the student's actions in the full sequence. Corresponding slices (e.g., the first slice from each sequence, the second slice from each, and so on) are then grouped into bins to define the temporal footprint of the sequence. Although the slices for different student sequences can be of different lengths (and there may be different numbers of students in each group), the interestingness measure described in the next section is calculated with respect to all of the actions in each bin and takes into account the proportion of total actions falling into each bin.

The number of slices/bins chosen for the temporal footprint is a parameter that determines the level of temporal granularity for the analysis. In addition to the desired granularity, an important consideration for choosing the number of bins is the quantity and variability of the data available. In particular, finer granularities (i.e., more bins) increase the likelihood that the analysis will be overwhelmed by random variation and noise. For example, spikes and other differences in pattern frequency that are identified across bins with relatively small quantities of actions are more likely to be the result of noise than with larger bins. With Betty's Brain data, initial qualitative analysis has suggested that with relatively few activity sequences (e.g., 10 to 15 students), 3 to 5 bins may be most effective, while with more sequences (e.g., 30 or more students), anything from 5 to 10 bins tends to work well.

3.3 D-TIPS for Ranking Patterns by Interestingness

In order to identify more interesting patterns by their difference in temporal usage across groups, the D-TIPS interestingness measure applies information gain (IG) with respect to pattern occurrence across the groups in each of the n corresponding bins of their temporal footprints. Information gain is defined as the difference in expected information entropy [20] between one state and another state where some additional information is known (e.g., the difference between a set of data points considered as a homogeneous group versus one split into multiple groups based on the value of some other feature or attribute). Information entropy, H, is the amount of expected uncertainty found in a random variable, X, whose value we will refer to as the *class* of the data point. In D-TIPS, each data point is an action performed by a student and its class is the corresponding student group. IG when used in classifiers, such as decision trees [13], is applied to a dataset where each data point has multiple features in addition to its class. The IG of a given feature is then the reduction in expected uncertainty about the correct class of a data point when its feature value is known. IG is calculated as the difference between the information entropy of the data without knowledge of the feature values, $H(X)$, (i.e., based solely on the probability distribution of classes over the full dataset) minus the conditional entropy of the data set when the value of the feature is known, $H(X|F)$. In D-TIPS, the features are the patterns of actions to be ranked, and a particular action's feature value is the combination of whether the action begins an occurrence of the pattern, o, and the order number of the bin in which the action occurred, b. Therefore, the D-TIPS IG metric can be represented as:

$$IG(X|F) = H(X) - \sum_{v \in Vals(F)} p(F = v)H(X|F = v)$$

where

$$Vals(F) = \{(o, b)|o \in \{true, false\}, b \in \{1 \ldots n\}\}$$

Information gain is leveraged in classifiers to determine which features are most discriminatory because they provide the least amount of uncertainty among classes in the data. In a similar fashion, D-TIPS applies information gain to determine which patterns are the most interesting because knowledge of their occurrence *and* temporal location provides the least amount of uncertainty among the student groups. This information-theoretic definition of the D-TIPS measure provides two important properties: (1) given two patterns with the same total occurrences for each corresponding group (e.g., group A has an occurrence of a and group B has an occurrence of b for both patterns, although a and b may be different values), the pattern with the greater discrimination of groups by *differences in temporal location/bin among groups* will have a higher rank, and (2) given two patterns with the same relative temporal behaviors (i.e., the same proportion of a given group's total pattern occurrence in each corresponding bin) for each corresponding group, the pattern with the greater discrimination of groups by *differences in total occurrence among groups* will have a higher rank.

The D-TIPS measure provides a way of recognizing differences among groups both by total pattern occurrence and by temporal behavior (e.g., decreasing usage versus increasing usage, or spikes in different bins). Further, when the same differences across groups (by both total pattern occurrence and temporal behavior) occur for two patterns, the pattern with higher overall frequency will have the higher rank. Thus, D-TIPS tends to emphasize patterns with large relative differences among groups (by total occurrence and/or temporal behavior) even when they are not especially frequent in the overall dataset, while also emphasizing patterns with more moderate differences among groups when the frequency of the pattern in the overall dataset is high. Conversely, D-TIPS tends to deemphasize patterns that are homogeneous across groups (by both relative occurrence and temporal behavior) or that are especially rare in all groups.

3.4 Visualizing Temporal Evolution

Given a set of highly-ranked behavior patterns, researchers will need to analyze them in further depth to understand their implications for student learning and generate additional hypotheses or changes to scaffolding in a learning environment. An important aspect of this analysis is the consideration of how a pattern's frequency changes over time in different groups. In order to more rapidly and efficiently analyze this aspect of the patterns identified by D-TIPS, we employ a heat map visualization.

Heat maps utilize the counts of data (pattern occurrence frequency in this case) in discrete cells across one or more dimensions to determine the color of the cell. Cell color is based on where the corresponding count falls between the highest and lowest count in any of the cells. The nature of the coloring scheme helps draw a user's eye to areas of contrast and thus larger changes, as well as general trends, in the data. For this analysis of D-TIPS patterns, we employ a two-dimensional heat map where the x-axis is time/temporal-bin, while the y-axis is student group. Further, rather than raw occurrence counts, each cell's count is the *percentage* of its group's total pattern occurrence that falls within that temporal bin. The use of percentages of pattern occurrence allows analysis of temporal variation normalized by the total frequency of the pattern per group. Therefore, different temporal trends in pattern usage across groups will be highlighted, even when total pattern occurrence differs significantly among groups, which would otherwise wash out any trends in the groups with lower pattern occurrence.

4 Betty's Brain Data

The data employed for the analysis in Sect. 5 consists of student interaction traces from the Betty's Brain [21,22] learning environment. In Betty's Brain, students read about a science process and teach a virtual agent about it by building a causal map. They are supported in this process by a mentor agent, who provides feedback and support for their learning activities. The data analyzed here was obtained in a recent study with 68 7^{th}-grade students taught by the same teacher in a middle Tennessee school. At the beginning of the study, students were introduced to the science topic (global climate change) during regular classroom instruction, provided an overview of causal relations and concept maps, and given hands-on training with the system. For the next four 60-minute class periods, students taught their agent about climate change and received feedback on content and learning strategies from the mentor agent.

The study tested the effectiveness of two support modules designed to scaffold students' understanding of cognitive and metacognitive processes important for success in Betty's Brain. The *knowledge construction* (KC) support module scaffolded students' understanding of, and suggested strategies for, constructing knowledge by identifying causal relations in the resources. The *monitoring* (Mon) support module scaffolded students' understanding of, and suggested strategies for, monitoring Betty's progress by using the quiz results to identify correct and incorrect causal links on Betty's map. Participants were divided into a control and three treatment groups. The knowledge construction (KC) group used a version of Betty's Brain that included the KC support module and a causal link tutorial that they could access at any time and were prompted to enter when the mentor determined they were having difficulty identifying causal links in the resources. The monitoring (Mon) group used a version of Betty's Brain that included the Mon support module and a tutorial about employing link annotations to keep track of links shown to be correct by quizzes. The full (Full) group used a version of Betty's Brain that included both support modules and tutorials. Finally, the control (Con) group used a version that included neither the tutorials nor the support modules.

In Betty's Brain, the students' learning and teaching tasks were organized around seven activities: (1) reading resource pages to gain information, (2) adding or removing causal links in the map to organize and teach causal information to Betty, (3) querying Betty to determine her understanding of the domain based on the causal map, (4) having Betty take quizzes that are generated and graded by the mentor to assess her current understanding and the correctness of links in the map, (5) asking Betty for explanations of which links she used to answer questions on the quiz or queries, (6) taking notes for later reference, and (7) annotating links to keep track of their correctness determined by quizzes and reading. Actions were further distinguished by context details, which for this analysis were the correctness of a link being edited and whether an action involved the same subtopic of the domain as at least one of the previous two actions. The definition of actions in Betty's Brain learning activity sequences are discussed further in [6].

5 Results

To illustrate and characterize the performance of the D-TIPS technique, we present selected results from its application to student learning activity data in the Betty's Brain classroom study described in Sect. 4. The first step of the D-TIPS analysis identified 560 candidate patterns that occurred in at least half of the students in one or more of the four experimental conditions. Given the limited number of students in each condition, we chose to bin pattern occurrence values into fifths of the activity sequences for a broad analysis of their usage evolution over time. Table 1 presents 3 of the top 30 most differentially-interesting patterns identified by D-TIPS across the four scaffolding conditions. For comparison, the average occurrences per student and ranking by that value is also presented. Over half (18) of the top 30 D-TIPS patterns had a rank past 50th by occurrence, with 13 of them ranking beyond 100th, indicating that they would be unlikely to have been considered and further analyzed without D-TIPS.

Table 1. Selected patterns with D-TIPS and occurrence rankings

Pattern	D-TIPS rank	Occurrence rank	Avg occurrence
[Quiz]	3	2	21.8
[Read] → [Note]	18	100	1.7
[Read] → [Read] → [Remove Link^{-}]	29	137	1.4

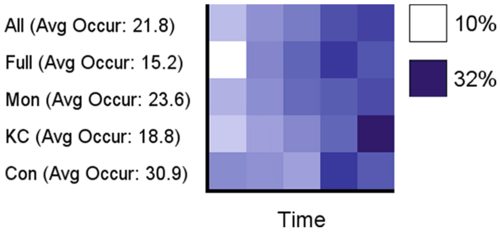

Fig. 1. [Quiz]

The first pattern in Table 1 illustrates a single action pattern that was ranked very high by both D-TIPS and overall occurrence. While individual student actions are often less interesting than longer patterns, they are still important to consider, especially when they also illustrate a tendency to be employed differentially across groups and over time. Figure 1 shows that all groups tended to use quizzes more frequently later in their work on the system. Since students' causal maps grew over time, monitoring and correction of the maps were more important later in their learning activities. There were some differences in usage trends over time among the different conditions, such as the steeper increasing trend for the KC and Full groups than the Monitoring group and the somewhat earlier peak in usage for the Full and Control groups. However, the overall occurrence by conditions differed markedly, with the Control group performing far more quiz actions than the others, and the Monitoring group performing more quiz actions than the KC and Full groups. While the Monitoring group's use of the quiz was expected to be high due to the focused monitoring support that relied heavily on the quiz, it is surprising that the Control group had the highest quiz usage. In addition, the Control group illustrated less of an increasing trend in usage compared to other groups, employing quizzes nearly as much in the first two fifths of their activity as in the final fifth. These results may indicate that without either KC or monitoring support, the Control group struggled more and fell back on strategies of guessing and checking (with the quiz).

Figure 2 illustrates a knowledge construction behavior of reading and taking notes that was ranked highly by D-TIPS. Another difference among the groups, which added to the interestingness of this pattern under the D-TIPS analysis, is that the Control group tended to perform reading followed by note-taking primarily in the last fifth of their activities, as opposed to the first two fifths for the other groups. Further analysis of the data attributed this primarily to two of the Control group students, although the reason for this aberration is still unclear.

The pattern illustrated in Fig. 3 involves a sequence of (two) reading actions followed by removing an incorrect link. While there was no consistent temporal trend in the usage of this pattern, the Monitoring and Control groups exhibited this pattern less than once per student, while the KC group averaged 2.4 times per student. Although ranked lower by D-TIPS at 45th, the sub-pattern of a

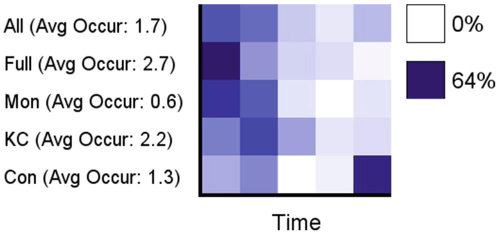

Fig. 2. [Read] → [Note]

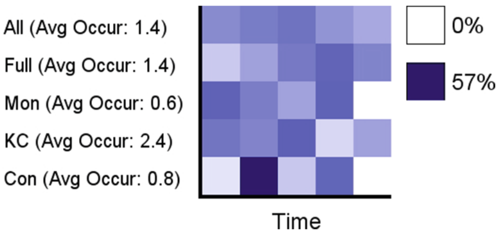

Fig. 3. [Read] → [Read] → [Remove Link^{-}]

single read action followed by removing an incorrect link illustrates the same differences. This suggests that students with the KC feedback generally relied more heavily on reading to identify incorrect links than either the Control and Monitoring groups, possibly because the Control group struggled more in general and the support in the Monitoring group focused students more on the use of quizzes to identify incorrect links.

6 Conclusion

While identification of high-frequency behavior patterns is undoubtedly useful, finding patterns that have differing occurrence over time across a set of student groups is also important for analyzing learning behaviors and the effects of scaffolding. In this paper, we presented the D-TIPS interestingness measure and mining approach, which identifies patterns that differ in their usage among student groups by either total (group) occurrence or temporal behavior, even when they are not especially frequent in the overall dataset. Results from the use of this technique to mine Betty's Brain data illustrated the potential benefits and helped characterize differences between D-TIPS and a baseline occurrence ranking. Moreover, D-TIPS identified patterns that illustrated potentially important differences in learning behavior among different scaffolding conditions that would have probably been overlooked by considering only pattern frequency. Future work will include autonomous identification of an effective number of bins for splitting a given set of activity sequences, as well as methods to individually characterize student groups by the patterns identified in D-TIPS. Further, we intend to apply the D-TIPS analysis to data in other domains to illustrate its generality and utility in areas beyond education.

Acknowledgments. This work has been supported by NSF-IIS Award #0904387 and IES Award #R305A120186.

References

1. Kinnebrew, J.S., Mack, D.L., Biswas, G.: Mining temporally-interesting learning behavior patterns. In: DMello, S.K., Calvo, R.A., Olney, A. (eds.): Proceedings of the 6th International Conference on Educational Data Mining, pp. 252–255 (2013)
2. Agrawal, R., Srikant, R.: Mining sequential patterns. In: Proceedings of the Eleventh IEEE International Conference on Data Engineering (ICDE), pp. 3–14 (1995)
3. Zaki, M.: Sequence mining in categorical domains: incorporating constraints. In: Proceedings of the Ninth International Conference on Information and Knowledge Management, pp. 422–429. ACM (2000)
4. Nesbit, J., Zhou, M., Xu, Y., Winne, P.: Advancing log analysis of student interactions with cognitive tools. In: 12th Biennial Conference of the European Association for Research on Learning and Insruction (EARLI) (2007)
5. Perera, D., Kay, J., Koprinska, I., Yacef, K., Zaïane, O.: Clustering and sequential pattern mining of online collaborative learning data. IEEE Trans. Knowl. Data Eng. **21**(6), 759–772 (2009)
6. Kinnebrew, J.S., Loretz, K.M., Biswas, G.: A contextualized, differential sequence mining method to derive students' learning behavior patterns. J. Educ. Data Min. **5**(1), 190–219 (2013)
7. Kinnebrew, J.S., Biswas, G.: Identifying learning behaviors by contextualizing differential sequence mining with action features and performance evolution. In: Proceedings of the 5th International Conference on Educational Data Mining (EDM 2012), Chania, Greece, June 2012
8. Amershi, S., Conati, C.: Combining unsupervised and supervised classification to build user models for exploratory learning environments. J. Educ. Data Min. **1**(1), 18–71 (2009)
9. Martinez, R., Yacef, K., Kay, J., Al-Qaraghuli, A., Kharrufa, A.: Analysing frequent sequential patterns of collaborative learning activity around an interactive tabletop. In: Proceedings of the Fourth International Conference on Educational Data Mining, Eindhoven, Netherlands (2011)
10. Su, J.M., Tseng, S.S., Wang, W., Weng, J.F., Yang, J., Tsai, W.N.: Learning portfolio analysis and mining for scorm compliant environment. J. Educ. Technol. Soc. **9**(1), 262–275 (2006)
11. Geng, L., Hamilton, H.J.: Interestingness measures for data mining: a survey. ACM Comput. Surv. (CSUR) **38**(3), 9 (2006)
12. Zhang, Y., Zhang, L., Nie, G., Shi, Y.: A survey of interestingness measures for association rules. In: International Conference on Business Intelligence and Financial Engineering, BIFE'09, pp. 460–463. IEEE (2009)
13. Bishop, C.M.: Pattern Recognition and Machine Learning. Springer, New York (2007)
14. Mitchell, T.: Machine Learning. McGraw Hill, New York (1997)

15. Merceron, A., Yacef, K.: Interestingness measures for association rules in educational data. In: Educational Data Mining 2008, p. 57 (2008)
16. Kotsiantis, S., Pierrakeas, C.: Efficiency of machine learning techniques in predicting students performance in distance learning systems. Technical report, Citeseer.
17. Sotiriou, C., Wirapati, P., Loi, S., Harris, A., Fox, S., Smeds, J., Nordgren, H., Farmer, P., Praz, V., Haibe-Kains, B., et al.: Gene expression profiling in breast cancer: understanding the molecular basis of histologic grade to improve prognosis. J. Natl Cancer Inst. **98**(4), 262–272 (2006)
18. Mashima, D., Kobourov, S., Hu, Y.: Visualizing dynamic data with maps. IEEE Trans. Vis. Comput. Graphics **18**(9), 1424–1437 (2012)
19. Ho, J., Lukov, L., Chawla, S.: Sequential pattern mining with constraints on large protein databases. In: Proceedings of the 12th International Conference on Management of Data (COMAD), pp. 89–100 (2005)
20. Renyi, A.: On measures of entropy and information. In: Fourth Berkeley Symposium on Mathematical Statistics and Probability, pp. 547–561 (1961)
21. Biswas, G., Leelawong, K., Schwartz, D., Vye, N., Vanderbilt, T.: Learning by teaching: a new agent paradigm for educational software. Appl. Artif. Intell. **19**(3), 363–392 (2005)
22. Leelawong, K., Biswas, G.: Designing learning by teaching agents: the Betty's Brain system. Int. J. Artif. Intell. Educ. **18**(3), 181–208 (2008)