Students’ learning style detection using tree augmented naive Bayes

Ling Xiao Li and Siti Soraya Abdul Rahman

Department of Artificial Intelligence, Faculty of Computer Science and Information Technology, University of Malaya, Kuala Lumpur, Malaysia

LXL, 0000-0002-2669-3883

Students are characterized according to their own distinct learning styles. Discovering students’ learning style is significant in the educational system in order to provide adaptivity. Past researches have proposed various approaches to detect the students’ learning styles. Among all, the Bayesian network has emerged as a widely used method to automatically detect students’ learning styles. On the other hand, tree augmented naive Bayesian network has the ability to improve the naive Bayesian network in terms of better classification accuracy. In this paper, we evaluate the performance of the tree augmented naive Bayes network in automatically detecting students’ learning style in the online learning environment. The experimental results are promising as the tree augmented naive Bayes network is shown to achieve higher detection accuracy when compared to the Bayesian network.

1. Introduction

Learning styles are a set of cognitive, emotional, characteristic and physiological factors that serve as the relatively stable indicators of how a student perceives, interacts with and responds to the learning environment [1]. For instance, most of the studies agree on the influence of learning style on the learning attitude, satisfaction level and academic achievement of students in an online education environment, and that the learning style can significantly affect the learning attitude in the education environment. By contrast, when students’ learning styles are mismatched, the learning effectiveness is reduced [2]. Past researchers have argued for consideration of learning styles in the development of e-learning systems. This is to maintain students’ motivation so that they could learn more effectively [3,4].

To detect learning style, most of the existing approaches can be divided into two types of detection methods: (i) static detection based on the learning style inventory and (ii) dynamic detection through the learning behaviour. Although the static
detection method is simple, the lengthy questionnaire on learning style often requires students to be patient in order to complete these questions, and so it often results in poor accuracy measurement and difficult to be dynamically updated [2]. Furthermore, many studies argued that learning style could vary over time and be dependent on task/learning content [5–8]. In order to solve these problems, numerous adaptive learning systems that use dynamic detection method have been developed by using the students’ learning behaviours. Several techniques have been proposed for detecting the learning behaviours such as neural networks, Bayesian networks, rule-based reasoning [9].

Many researchers have validated the effectiveness of the Bayesian network-based automatic detection method. In Feldman’s review [9], Bayesian networks are some of the most widely used detection techniques to automatically detect students’ learning styles [10–13]. Feldman et al. [9] reported the reasons for using the Bayesian network in terms of its natural representation of probabilistic information, and its ability to encode expert knowledge. However, due to some characteristics of the Bayesian network, it leads to the Bayesian network structure becoming unreasonable and affecting the classification accuracy [5].

Besides, when dynamic model is used to detect students’ learning style, a ‘cold start’ problem exists inevitably [14], because (i) there is not enough information available to build users’ profile [15] and (ii) the system is unable to infer anything for the new users. To tackle this issue, [10,16] proposed hybrid detection method that combines the static and dynamic methods. As an example, Liyanage et al. [10] used the Felder–Silverman learning style model as the basis to predict the learning style during the early stage and followed by the Bayesian network to mine the learning behaviour pattern in order to detect the learning style more accurately. In another approach, Petchboonmee et al. [16] used J48 and naive Bayes algorithm to establish the decision tree for the learning style prediction model, which is based on the Kolb learning style model. Although hybrid detection methods have been successful, we argue that simply combining the two detection methods (static and dynamic) remains a major drawback as students are still required to fill out a lengthy questionnaire.

In this paper, we propose a new approach to represent and detect students’ learning style using tree augmented naive Bayesian network to improve the precision of learning style detection. The detection process starts with preset learning styles. Then, it amends and updates the learning style according to the tree augmented naive Bayesian. Experimental results showed the effectiveness of the proposed method.

This paper is organized as follows. Section 2 briefly describes different learning styles that we are considering in this paper. Section 3 presents how the tree augmented naive Bayesian is used in comparison to the Bayesian network. Section 4 describes the detection of learning styles using the tree augmented naive Bayesian, as well as the experimental analysis and results. Section 5 discusses the experimental findings. Finally, §6 concludes the paper and provides direction for future work.

2. Learning styles revisited

Generally, a learning style model classifies students according to where they fit in a number of scales that identify the ways in which they receive and process information [17].

In this paper, we consider Felder–Silverman learning style model (FSLSM), because of the main reasons as follows:

(1) FSLSM is the most widely used learning style model. Shockley & Russell [3] analysed the use of learning style model in the adaptive learning system over the past decade and found that the usage amount of FSLSM model ranks the first (50%), much higher than the second Kolb’s model (8.6%). The findings of this study are also consistent with those of Akbulut & Cardak [18].

(2) FSLSM provides more detailed descriptions than other learning style models while its reliability and accuracy have also been proven [19].

(3) FSLSM provides a high operational index of learning style (ILS) instrument, which includes 44 questions: 11 questions for each dimension, where each question has two answers to be chosen from, in order to detect both the preference and the degree of preference.

FSLSM divides the learning styles into four dimensions: (i) procession (active/reflective), (ii) perception (sensing/intuitive), (iii) input (visual/verbal) and (iv) understanding (sequential/global) [20].

Procession: active students do not learn much in situations that require them to be passive. Whereas, reflective students do not learn much in situations that provide no opportunity to think about the
information being presented. Active students work well in groups; reflective students work better by themselves or at most, with one other person.

Perception: sensing students prefer facts, data and experimentation; whereas intuitive students prefer principles and theories. Sensing students are patient with detail but do not like complications; whereas intuitive students are bored by details and welcome complications.

Input: visual students remember best what they see and this includes pictures, diagrams, time lines, films, and demonstrations. Verbal students remember much of what they hear or read and say.

Understanding: sequential students follow linear reasoning processes when solving problems. Global students make intuitive leaps and may be unable to explain how they come up with solutions. Sequential students can work with the material when they understand it partially or superficially, while global students may have great difficulty in doing so.

3. Overview of tree augmented naive Bayesian network

3.1. Bayesian network and naive Bayesian network

Bayesian network is an uncertain relationship representation and reasoning model based on probability analysis and graph theory. It is a directed acyclic graph where nodes represent random variables and arcs represent the probabilistic correlation between variables [21].

The absence of edges in a Bayesian network denotes statements of independence. A Bayesian network encodes the following statement of independence about each random variable: a variable is independent of its non-descendants in the network given the state of its parents [22]. A Bayesian network also represents a particular probability distribution, the joint distribution over all the variables represented by nodes in the graph. This distribution is specified by a set of conditional probability tables (CPT). Each node has an associated CPT that specifies this quantitative probability information. Such table specifies the probability of each possible state of the node given each possible combination of states of its parents. For nodes without parents, probabilities are not conditioned on other nodes. These are called the prior probabilities of these variables.

Past studies [8] found that Bayesian networks are very suitable to be employed as the detection technique for adaptive education domain. However, the general Bayesian network is too complex for small datasets and easy to overfit [5]. Naive Bayesian avoids this problem. The naive Bayesian classifier is an effective classifier due to two advantages that it has over other classifiers. Firstly, it is easy to be constructed; as the structure is given a priority besides no structure learning procedure is required. Secondly, the classification process is very efficient. Both advantages are derived by assuming that all features are independent of each other. The simple structure only contains two layers, the classification node as the parent node of all other nodes. No other connection is allowed in the naive Bayesian network as shown in figure 1. The only connections link between node c and all leaf nodes $x_1, x_2, x_3, x_4$; the naive Bayes assumes that all leaf nodes are conditionally independent. But, the conditional independence assumption in the naive Bayes is rarely true in reality. In an adaptive educational domain, the naive Bayesian assumption is (nearly) always violated because the variables are often interconnected.

3.2. Tree augmented naive Bayesian network

The requirement that each node must be independent renders the naive Bayesian network structure unreasonable, resulting in the poor accuracy of the naive Bayesian classifier. Friedman et al. [23] studied tree augmented naive Bayesian, which extends the naive Bayesian by allowing tree-like structures to be used to represent the dependencies among attributes. Figure 2 shows node c and all leaf nodes $x_1, x_2, x_3, x_4$ with their respective arcs from node c, from a tree [24]. Tree augmented naive Bayesian makes a good compromise between the general Bayesian network and naive Bayesian. Also, the structure of tree augmented naive Bayesian is simple enough to avoid overfit and strong dependencies can be taken into account.

Unlike naive Bayesian networks, tree augmented naive requires a learning procedure that constructs the model structure. At present, the typical tree augmented naive learning procedure to construct the tree augmented naive classifier uses conditional mutual information.

The algorithm for learning tree augmented naive models is a variant of the Chow-Liu [25] algorithm that is used to learn tree-structured Bayesian networks. Let $C$ represent the class variable
and \( \{X_i\}_{i=1}^n \) be the features (non-class variables). The tree augmented naive learning procedure is as follows:

1. Compute the conditional mutual information:
   \[
   I(X_i; X_j | C) = \sum_{x_i, x_j, c} P(x_i, x_j, c) \log \frac{P(x_i, x_j | c)}{P(x_j | c) P(x_i | c)}. \]

   According to probability theory and information theory, the mutual information of two random variables is a quantity that measures the mutual dependence of the two random variables. Using the conditional mutual information to test the conditional independence of \( I(X_i, Y_i, Z_i) \), where \( P(\cdot) \) is the empirical distribution, computed from the training data. Intuitively, this quantity represents the gain in information by adding \( X_i \) as a parent of \( X_j \) given that \( C \) is already a parent of \( X_j \).

2. Build a complete undirected graph on the features \( \{X_1, \ldots, X_n\} \), where the weight of the edge between \( X_I \) and \( X_J \) is \( I(X_i; X_j | C) \).

3. Find a maximum weighted spanning tree of the completed undirected graph.

4. Pick an arbitrary node of the maximum weighted spanning tree as the root and set the direction of all edges to be outward from the root to build a directed graph.

5. Add a class node and an arc between the class node and attribute node to construct tree augmented naive model.

In the current Bayesian network classifiers, tree augmented naive is considered as a widely accepted Bayesian classifier with wide applicability and good comprehensiveness for performance, efficiency, and space–time complexity.

4. Learning style detection model based on tree augmented naive Bayesian network

4.1. Preset learning style

In order to avoid the inconvenience of filling out lengthy learning style questionnaires, this study suggests that the learning style of students be preset at the beginning of learning. The conclusion drawn
from the analysis of adult learning style in Shockley and Russell’s study [3] found that students are mostly reflective in the information processing dimension, intuitive in the perception dimension, visual for their input, and more sequential in the understanding dimension. In addition, researchers have also studied students’ learning style characteristics in different disciplines (biology, commerce, chemistry, finance, accounting, and many more). The results generally prove that the students’ learning styles are characterized by disciplines and specialties. Learning style is also affected by culture, background, different countries [2]. Therefore, in order to identify students’ learning style more accurately, this research gathered 46 undergraduates studying bioinformatics. They were required to fill out the ILS instrument online. This requirement was only carried out at the early stage of the experimental study. Figure 3 shows that students’ learning styles in four dimensions are more inclined to the active, intuitive, visual and sequential learning styles. In our study, only the procession dimension is different from Shockley & Russell’s result [3], but the other dimensions are the same. In fact, it was discovered that the bioinformatics students required more practical work, while most of the other courses require students to collaborate in a team-based manner. Therefore, the default preset learning styles for these students in the current study are active, intuitive, visual and sequential.

4.2. Construction of learning style detection model based on tree augmented naive Bayesian network

Students enter the system with preset learning styles. Their learning styles will be updated individually based on their learning behaviour as they interact with the system. This learning style detection model is constructed based on tree augmented naive Bayesian network to mine data from students’ learning behaviour. The learning behaviours mainly include visiting the forum, sending and receiving e-mail, watching videos, carrying out exercises, communicating, and many more. Based on the literature [12,26,27], an FSLSM-based learning style Bayesian network model was built as shown in figure 4.

Active students work well in groups; reflective students work better by themselves or at most, with one person. Also, we can assess students from their wiki, forums, online chat and e-mail usage to identify if they are active or reflective.

Sensing students prefer facts, data and experimentation whereas intuitive students prefer principles and theories. Sensing students are patient with detail but do not like complications, whereas intuitive students are bored with detail and welcome complications. If a student likes specific learning materials, learns through examples and case studies, prefers to review after exams and carefully examine the questions—these characteristics indicate that he/she is a sensing student. On the other hand, he/she is an intuitive student.
Visual students like to learn using pictures, diagrams, video and animation material. If students like to study using text and audio material, this indicates that they are verbal students.

Sequential students study in a step-by-step manner and follow linear processes according to the learning contents. On the other hand, global students make intuitive leaps and may struggle to explain how they come up with solutions. Additionally, if a student does not read or learn the relevant learning content, but he/she is able to complete the test and obtain high marks, it could be inferred that he/she is a global student.

The recommended setting of variables based on literature [19,28] is presented in Table 1. The table describes the different states of the independent variables related to students’ learning behaviour.

In this paper, we use procession (Pro) dimension node to illustrate the model construction and algorithm implementation. There are two classifications of procession (Pro) nodes: active (Pro1) and reflective (Pro2). Wiki, forum, online chat and e-mail are leaf nodes. The degree of usage is according to students’ participation as follows:

1. Wiki (W): very frequently, occasionally, never.
2. Forum (F): post, reply, read, never.
3. Online chat (C): very frequently, occasionally, never.
4. E-mail (E): very frequently, occasionally, never.

Table 2 shows the training dataset from the learning system. Rows represent all students as training data, and columns represent all relevant features of procession dimension, the values indicate the students’ behaviour and preference, respectively.

We can calculate the CPT of node Pro according to the conditional probability \( P(A/B) = P(AB)/P(B) \) (refer to Table 3). The CPT of node Pro is the prior probability. The prior probabilities of node W and F are shown in Tables 4 and 5, respectively.

4.3. Detection algorithm based on tree augmented naive Bayesian network

The steps of learning style detection by tree augmented naive Bayesian network are:

1. Establish the procession node tree augmented naive Bayesian network structure.

   Step 1. The conditional mutual information between the W, F, C, E and Pro attribute variables is calculated according to the procedure of tree augmented naive calculation as described...
Table 1. Recommended setting of variables.

| dimension      | features          | description of behaviour                                                                 |
|----------------|-------------------|------------------------------------------------------------------------------------------|
| procession     | forum             | posts messages; replies messages; reads messages; never use                                    |
|                | wiki              | very frequently use; occasionally use; never use                                            |
|                | e-mail            | very frequently use; occasionally use; never use                                            |
|                | online chat       | very frequently use; occasionally use; never use                                            |
| perception     | example           | in relation to the number of examples proposed: many (more than 75%); few (25–75%); none    |
|                | assessment        | in relation to the number of assessments proposed: more than 75%; few (25–75%); none        |
|                | exam_revision     | in relation to the time assigned to the exam: more than 20%; 10–20%; less than 10%          |
|                | exercise          | in relation to the number of exercises proposed: many (more than 75%); few (25–75%); none   |
| input          | text              | in relation to the text-based learning objects proposed: many (more than 75%); few (25–75%); none |
|                | image             | in relation to the image-based learning objects proposed: many (more than 75%); few (25–75%); none |
|                | video             | in relation to the video-based learning objects proposed: many (more than 75%); few (25–75%); none |
|                | audio             | in relation to the audio-based learning objects proposed: many (more than 75%); few (25–75%); none |
| understanding  | exam_results      | in relation to the time assigned to the exam: more than 20%; 10–20%; less than 10%          |
|                | learning path     | in fits and starts; sequential                                                             |

Table 2. Training dataset of procession dimension.

| student | procession | forum | wiki | online chat | e-mail |
|---------|------------|-------|------|-------------|--------|
| 1       | Pro1       | F1    | W1   | C1          | E1     |
| 2       | Pro1       | F1    | W1   | C2          | E1     |
| 3       | Pro1       | F1    | W1   | C1          | E2     |
| 4       | Pro2       | F2    | W1   | C3          | E1     |
| 5       | Pro1       | F2    | W1   | C1          | E1     |
| 6       | Pro1       | F2    | W1   | C2          | E3     |
| 7       | Pro2       | F3    | W1   | C3          | E3     |
| 8       | Pro1       | F3    | W1   | C1          | E1     |
| 9       | Pro2       | F4    | W1   | C3          | E2     |
| 10      | Pro1       | F4    | W2   | C3          | E1     |
| ...     | ...        | ...   | ...  | ...         | ...    |
| 35      | Pro2       | F4    | W3   | C3          | E2     |
| 36      | Pro2       | F4    | W3   | C3          | E1     |

above. The results are:

\[
I_p(W; F/Pro) = 0.228918 \quad I_p(W; C/Pro) = 0.211235 \\
I_p(W; E/Pro) = 0.168211 \quad I_p(F; C/Pro) = 0.129722 \\
I_p(F; E/Pro) = 0.121512 \quad I_p(C; E/Pro) = 0.063124
\]

Step 2. Construct the weighted undirected graph by weighting the conditional mutual information between pairs of attribute variables as shown in figure 5.
Step 3. Sort the weight $e_{ij}$ in descending order, where $e_{ij}$ represents weight between the corresponding nodes: $e_{wF}$, $e_{fC}$, $e_{wE}$, $e_{wC}$, $e_{fF}$, $e_{eC}$.

Step 4. Build the maximum weighed spanning tree: $e_{wF}$, $e_{fC}$, $e_{wE}$.

Step 5. Establish a directed tree using node C as the root node and increase the class variable node, the arcs between a class variable node and attribute node. The tree augmented naive Bayesian network structure is established with the class variable as the parent node of all attribute nodes as shown in figure 6.

(2) Establish the tree augmented naive Bayesian network parameters. According to the learning process parameter above, CPTs of nodes C, F, W and E can be calculated separately as shown in tables 6–9.
Figure 6. Tree augmented naive network structure.

Table 6. CPT of node C.

|     | Pro = Pro1 | Pro = Pro2 |
|-----|------------|------------|
| C1  | 8/16       | 4/20       |
| C2  | 5/16       | 4/20       |
| C3  | 3/16       | 12/20      |

Table 7. CPT of node F.

|     | C     | Pro1 | Pro1 | Pro1 | Pro2 | Pro2 | Pro2 |
|-----|-------|------|------|------|------|------|------|
| F   | C     |      |      |      |      |      |      |
| F1  | C1    | 3/9  | 3/5  | 1/3  | 0    | 1/2  | 2/11 |
| F2  | C2    | 2/9  | 2/5  | 1/3  | 0    | 0    | 1/11 |
| F3  | C3    | 2/9  | 0    | 1/3  | 2/4  | 1/2  | 3/11 |
| F4  | C4    | 1/9  | 0    | 0    | 2/4  | 0    | 5/11 |

Table 8. CPT of node W.

|     | F     | Pro1 | Pro1 | Pro1 | Pro1 | Pro2 | Pro2 | Pro2 | Pro2 |
|-----|-------|------|------|------|------|------|------|------|------|
| W   | F     |      |      |      |      |      |      |      |      |
| W1  | F1    | 3/7  | 3/6  | 1/2  | 1    | 0    | 0    | 1/6  | 1/8  |
| W2  | F2    | 2/7  | 2/6  | 1/2  | 0    | 2/3  | 1/2  | 2/6  | 3/8  |
| W3  | F3    | 2/7  | 1/6  | 0    | 0    | 1/3  | 1/2  | 3/6  | 4/8  |

Table 9. CPT of node E.

|     | W     | Pro1 | Pro1 | Pro1 | Pro2 | Pro2 | Pro2 | Pro2 |
|-----|-------|------|------|------|------|------|------|------|
| E   | W     |      |      |      |      |      |      |      |
| E1  | W1    | 5/7  | 3/7  | 1    | 0    | 1/8  | 0    |
| E2  | W2    | 1/7  | 3/7  | 0    | 1/3  | 2/8  | 4/9  |
| E3  | W3    | 1/7  | 1/7  | 0    | 2/3  | 5/8  | 5/9  |
(3) Tree augmented naive Bayesian network reasoning.

Assume a given student’s learning behaviour set is frequent access to wiki, reading posts, occasional chatting online, occasional e-mailing (\(X = \{W_1, F_3, C_2, E_2\}\)). Respectively, \(P(X_i | Y_i)\), \(i = 1, 2\). The prior probability \(P(Pro)\) for each class can be calculated from the training data: \(P(Pro = ‘active’) = 16/36\), \(P(Pro = ‘reflective’) = 20/36\). The prior probability can be derived as follows:

\[
P(X | Pro_1) = \prod_{i=1}^{4} p(X_i | Pa(X_i))
= p \left( x_0 = \frac{W_1}{F_3}, pro_1 \right) \times p \left( x_1 = \frac{F_3}{C_1}, pro_1 \right) \times p \left( x_2 = \frac{C_1}{pro_1} \right)
\times p \left( x_2 = \frac{E_2}{W_1}, pro_1 \right) = \frac{1}{2} \times \frac{2}{9} \times \frac{8}{16} \times \frac{1}{7} = 0.07712.
\]

Similarly,

\[
P(X | Pro_2) = \prod_{i=1}^{4} p(X_i | Pa(X_i))
= p \left( x_0 = \frac{W_1}{F_3}, pro_2 \right) \times p \left( x_1 = \frac{F_3}{C_1}, pro_2 \right) \times p \left( x_2 = \frac{C_1}{pro_2} \right)
\times p \left( x_2 = \frac{E_2}{W_1}, pro_2 \right) = \frac{1}{6} \times \frac{3}{11} \times \frac{4}{30} \times \frac{1}{3} = 0.00371.
\]

Therefore, the preliminary result of the tree augmented naive Bayesian network for \(X\) is \(Pro = ‘active’\). Then:

\[
P(Pro = Pro_1) = \frac{P(X / Pro_1)}{P(X / Pro_1) + P(X / Pro_2)} = 0.963 = 95.8\%\]
\[
P(Pro = Pro_2) = \frac{P(X / Pro_2)}{P(X / Pro_1) + P(X / Pro_2)} = 0.037 = 3.8\%\]

The scales of the index for each dimension’s learning style of FSLSM are 1,3,5,7,9,11, where 1 and 3 represent learning styles that are fairly well balanced, 5 and 7 indicate a moderate preference, and 9 and 11 indicate a very strong preference. Therefore, 50–100% is divided into three levels, corresponding to the ILS learning style preference levels. A probability of 50–66.7% indicates fairly well balanced, 66.8–83.4% indicates a moderate preference, 83.5–100% shows a strong preference. According to the above calculation, the results show a strong tendency for the ‘active’ on procession dimension.

4.4. Experiment and result analysis

4.4.1. Experimental design

We assessed 46 undergraduate bioinformatics students for 7 weeks on genomic technology topics via an online course using learning management system called Moodle. The students had no prior knowledge of the topics, and they were given same static learning materials. Overall, the online course included 396 well-balanced content objects for different learning style dimensions, covered all necessary learning materials, including examples, exercises, images, video tutorials, and many more. The students’ learning behaviours were used to detect the individuals’ learning style for comparative analysis between general Bayesian network and tree augmented naive. The learning style of the students was also obtained using the ILS instrument for the comparative analysis and to validate the preset learning styles concept as mentioned before.

4.4.2. Comparative analysis of experimental results

The results obtained by the Bayesian network, tree augmented naive and ILS instrument of four dimensions are shown in Table 10.

The precision of Bayesian network and tree augmented naive model learning style detected results can be calculated by using the following formula [12].

\[
\text{precision} = \frac{\sum_{i=1}^{n} \text{Sim}(LS_{ILS}, LS_{BNS/TANS})}{n}.
\]
The result of tree augmented naive achieves better. On the other hand, the tree augmented naive connections existed between learning objects within the same learning style dimension, such as ‘online but also from the attribute of parent node [30]. During students’ online learning process, many internal model in four dimensions of learning style are as shown in table 11.

Table 11. Results comparison.

| dimensions | precision |           |           |
|------------|-----------|-----------|-----------|
|            | BN        | TAN       |
| procession | 62.5      | 75.3      |
| perception | 67.3      | 72.8      |
| input      | 73.2      | 80.0      |
| understanding | 65.2 | 75.2 |

The precision values obtained are between [0, 1], where 1 means that the learning style obtained by the Bayesian network or tree augmented naive model is the same as that of to the ILS result; and 0 means that the learning style detected by Bayesian network or tree augmented naive model is completely opposite to the ILS result; and 0.5 if one is neutral and the other represents an extreme value; while is the number of students assessed.

The detected results for the precision of Bayesian network model and tree augmented naive Bayesian model in four dimensions of learning style are as shown in table 11.

5. Discussion

The results showed that the tree augmented naive has higher precision than the Bayesian network. This is due to the fact that the tree augmented naive algorithm loosens the conditional independence assumption which is consistent with the reality (the interconnection between variables). Comparing with naive Bayesian, tree augmented naive Bayesian allows the additional edges between the attributes of the network in order to capture correlations among them [29]. Furthermore, each attribute can have augmenting edge which encodes statistical dependencies between attributes; therefore, the joint probability of tree augmented naive Bayesian count on the probabilities conditioned not only on class but also from the attribute of parent node [30]. During students’ online learning process, many internal connections existed between learning objects within the same learning style dimension, such as ‘online chat’ often appeared together with ‘forum’ section; when the correlation of the interconnection is higher, the result of tree augmented naive achieves better. On the other hand, the tree augmented naive
algorithm takes slightly more time than the Bayesian network. This is because the tree augmented naive needs to build the tree on the basis of Bayesian network tree.

In this study, 36 students’ data were used for training the classifier. Nevertheless, when looking at the possible variables, for example, in perception dimension, which has 4 different features, 81 (3⁴) possible different states exist because each feature can have 3 states. Using only 36 students as input data might affect the precision of the detected results. Meanwhile, from another point of view, the precision of the results using the proposed approach could be further improved when running in big dataset environment. Another limitation of the current study is that the results of the experiment were only tested on Moodle platform with a specific subject. The consistency of performance needs to be tested when it runs with different learning management platforms or other online courses. Our future work will involve exploring further the performance in different environments.

6. Conclusion

We have evaluated the capability of tree augment naive Bayesian to model and detect students’ learning styles. The results obtained are positive. Since the tree augmented naive Bayesian network retains the structural features of naive Bayes and relaxes its independence assumption, we could make classifications with higher accuracy. Experimental results prove that the proposed method is more accurate than the results obtained using the Bayesian network.

Although the experiment only assessed restricted numbers of students, the results obtained provide valuable data about students’ learning behaviours with regards to online courses. These data will be used in future to enhance students’ learning style modelling. For future work, the experiment will be carried out on a larger scale in order to validate the results obtained so far and to test the performance consistency.

In summary, provided that we take into account issues on learning style detection, the proposed tree augmented naive Bayesian model enables us to discover students’ learning styles in a highly precise manner.

Ethics. The experimental procedures were approved by the University Committee for the Protection of Human Subjects, and the research was conducted according to the principles expressed in the Declaration of Helsinki. Informed consent was obtained from each participant.

Data accessibility. Data are included as the electronic supplementary material.

Authors’ contributions. L.L.X and S.S.A.R conceived and design the study; L.L.X performed the experiments, contributed to data acquisition, carried out the data analyses and drafted the manuscript; S.S.A.R supervised the research.

Competing interests. The authors declare no conflict of interests.

Funding. The authors received no funding for the study.

References

1. Keefe JW. 1987 Learning style. Theory and practice. Estados Unidos: National Association of Secondary School Principals.
2. Graf S, Liu TC. 2010 Analysis of learners’ navigational behaviour and their learning styles in an online course. J. Comput. Assist. Learn. 26, 116–131. (doi:10.1111/j.1365-2729.2009.00336.x)
3. Shockley DR. 2005 Learning styles and students’ perceptions of satisfaction in community college Web-based learning environments. Doctoral dissertation, North Carolina State University, Raleigh, NC, USA.
4. Ahmad N, Tasei Z, Kasim J, Sahat H. 2011 Automatic detection of learning styles in learning management systems by using literature-based method. Procedia Social Behav. Sci. 103, 181–189. (doi:10.1016/j.apsbs.2013.12.024)
5. Kozma RB. 1999 Learning with media. Rev. Educ. Res. 61, 179–211. (doi:10.3102/00346543061002179)
6. Snow RE. 1991 Attitude-treatment interaction as a framework for research on individual differences in psychotherapy. J. Consult. Clin. Psychol. 59, 205. (doi:10.1037/0022-006X.59.3.205)
7. Hecht C. 2003 iWeaver: towards’ learning style’-based e-learning in computer science education. In Proc. 5th Australasian Conf. on Computing Education, Adelaide, Australia, 4–7 February 2003, pp. 273–279. Darlinghurst, Australia: Australian Computer Society, Inc.
8. Beech N, McGill I, Brockbank I. 2017 Reflective learning in practice. In Reflective learning in practice (eds A Brockbank, I McGill), pp. 30–40. Abingdon, UK: Routledge.
9. Feldman J, Montsenies A, Amandie A. 2015 Automatic detection of learning styles: state of the art. Artif. Intell. Rev. 44, 157–186. (doi:10.1007/s10462-014-9422-6)
10. Usanjan MPP, KS LG, Hirakawa M. 2016 Detecting learning styles in learning management systems using data mining. J. Inf. Process. 24, 740–749. (doi:10.2197/ipsjji.24.740)
11. Saberi N, Montazer GA. 2012 The extraction of learner’s ability in e-learning environment using Bayesian network. Educ. Technol. Lett. 2, 51–56.
12. Garcia P, Amadie A, Schaffino S, Campo M. 2007 Evaluating Bayesian networks’ precision for detecting students’ learning styles. Comput. Educ. 49, 794–808. (doi:10.1016/j.compedu.2005.11.017)
13. Millán E, Lobado T, Pérez-de-la-Cruz JL. 2010 Bayesian networks for student model engineering. Comput. Educ. 55, 1663–1683. (doi:10.1016/j.compedu.2010.07.010)
14. Ramírez-Anellano A, Bory-Reyes I, Hernández-Simon LM. 2016 Learning object assembly based on learning styles. In Smart education and e-learning 2016 (eds V Uskov, R Howlett, L Jain), pp. 447–462. Cham, Switzerland: Springer. (doi:10.1007/978-3-319-39690-3_40)
15. Souali K, El Afiö A, Faci R, Chiheb R. 2011 A new recommender system for e-learning environments. In 2011 Int. Conf. on Multimedia Computing and Systems, Ouarzazate, Morocco, 7–9 April 2011. (doi:10.1109/ICMCS.2011.5945630)
16. Petchboonme P, Phonak D, Tiantong M. 2015 A Comparative data mining technique for David Kolb’s experiential learning style classification. Int. J. Inf. Educ. Technol. 5, 672. (doi:10.7763/IJ.IET.2015.05.590)
17. Coffield F, Moseley D, Hall E, Ecclestone K. 2004 Should we be using learning styles? What research has to say to practice. London, UK: Learning and Skills Research Centre.

18. Akbulut Y, Cardak CS. 2012 Adaptive educational hypermedia accommodating learning styles: a content analysis of publications from 2000 to 2011. Comput. Educ. 58, 835–842. (doi:10.1016/j.compedu.2011.10.008)

19. García P, Amandi A, Schiaffino S, Campo M. 2005 Using Bayesian networks to detect students’ learning styles in a web-based education system. In Proc. Argentine Symp. on Artificial Intelligence, Rosario, Argentina, 29–30 August 2005, pp. 115–126.

20. Felder RM, Silverman LK. 1988 Learning and teaching styles in engineering education. Eng. Educ. 78, 674–681.

21. Jensen FV. 1997 An introduction to Bayesian networks. New York, NY: Springer.

22. Pearl J. 2014 Probabilistic reasoning in intelligent systems: networks of plausible inference. Burlington, MA: Morgan Kaufmann.

23. Friedman N, Geiger D, Goldszmidt M. 1997 Bayesian network classifiers. Mach. Learn. 29, 131–163. (doi:10.1023/A:1007465528199)

24. Khor IC, Ting CY, Amnuaisuk SP. 2009 From feature selection to building of Bayesian classifiers: a network intrusion detection perspective. Am. J. Appl. Sci. 6, 1948. (doi:10.3844/ajassp.2009.1948.1959)

25. Chow C, Liu C. 1968 Approximating discrete probability distributions with dependence trees. IEEE Trans. Inf. Theory 14, 462–467. (doi:10.1109/TIT.1968.1054142)

26. Carmona C, Castillo G, Millán E. 2007 Discovering student preferences in e-learning. In Proc. Int. Workshop on Applying Data Mining in e-Learning, Crete, Greece, 17–20 September 2007, pp. 33–42.

27. Romero C, Ventura S. 2010 Educational data mining: a review of the state of the art. IEEE Trans. Syst. Man Cybern. C 40, 601–618. (doi:10.1109/SMCC.2010.205332)

28. Graf S, Liu TC. 2009 Supporting teachers in identifying students’ learning styles in learning management systems: an automatic student modelling approach. J. Educ. Technol. Soc. 12, 3.

29. Carvalho AM, Oliveira AL, Sagot MF. 2007 Efficient learning of Bayesian network classifiers. In AI 2007: advances in artificial intelligence (eds MA Orgun, J Thornton). Lecture Notes in Computer Science, vol. 4830, pp. 16–25. Berlin, Germany: Springer. (doi:10.1007/978-3-540-76928-6_4)

30. Dhakar M, Tiwari A. 2014 A novel data mining based hybrid intrusion detection framework. J. Inf. Comput. Sci. 9, 037–048.