CHAPTER 15

Online outlier detection of learners’ irregular learning processes

M. Ueno
Faculty of Engineering, Nagaoka University of Technology, Japan.

Abstract

Distance education using e-learning has become popular in educational situations. One problem is that the instruction strategy tends to be one way, so sometimes learners find it more boring than conventional instruction methods. This chapter proposes a method of online outlier detection of learners’ irregular learning processes using the learners’ response time to e-learning content. The unique features of this method are as follows: (1) It uses the Bayesian predictive distribution, (2) It can be used for small samples, (3) It unifies the methods of various statistical tests using a hyper-parameter and provides more accurate test results than one of the traditional methods alone, (4) It assists two-way instruction using data mining results of learners learning processes, (5) Outlier statistics are estimated by considering both students’ abilities and the difficulty of content. In addition, this chapter proposes an animated agent which provides adaptive messages to the learners using the data mining. Moreover, the system was evaluated and the results showed the effectiveness of the system.

1 Introduction

Motivation is essential to learning and performance, particularly in e-learning environments where learners must take an active role and be self-directed in their learning [1]. Despite the importance of motivation to learning, between 1988 and 2000, less than one percent of papers at the international conferences concerned with distance education focused on motivational issues. Keller (1999) argues that although motivation is idiosyncratic, learner motivation can also be affected by
external aspects [2]. Visser reported that motivational messages can reduce dropout rates [3] and later attempted to improve motivation in e-learning situations using such messages [4]. Gabrielle applied technology-mediated instructional strategies to Gagnes events of instruction and showed the effects of these strategies on motivation [5]. These studies emphasize the effects of the teacher’s motivational messages adapted to a learner’s status. However, when the number of learners is large, it becomes difficult for a teacher to individualize and personalize messages to students.

On the other hand, acquiring huge amounts of learning history data using e-learning is easy because it is automatically saved as log-data in learning history data-base. In this situation, it is important how we save or use this data. In this sense, the data mining technologies has become one of notable techniques to discover useful knowledge from the huge amounts of data [6, 7].

The main idea of this work in this chapter is to develop a method of detecting students who requires the previous mentioned motivational messages using the data mining technologies. The main the proposal in this chapter is an online outlier detection of learners’ irregular learning processes using response times to e-learning content. Although many outlier detection techniques have been proposed, the techniques can be classified into the following two. One is a detection technique using statistical test methods [8, 9] and another is one using neural network techniques [10–13]. However, applying the conventional outlier techniques to the problem of detecting irregular learning processes in e-learning has the following problems.

- If a learner starts learning using irregular learning processes, then regular learning processes will be regarded as irregular.
- The conventional techniques assume that all data in a time series are for the same task. However, in educational situations, tasks (= content) in a time series differ in the sense of difficulty.
- In the conventional techniques, the criteria that specify outliers are not statistically well defined.

Considering these problems, this chapter proposes a new outlier detection technique to detect learners’ irregular e-learning processes. The unique advantages of this method are as follows:

- It can use prior knowledge reflecting the response time characteristics of each content using Bayesian approach. This prior knowledge can avoid regarding regular processes as irregular at the beginning of the learning processes.
- The proposed method uses a model that incorporates task difficulties parameters and learner’s ability parameters. Outlier statistics are estimated considering both students’ abilities and the difficulty of the content, so the method efficiently detects irregular learning.
- The proposed method uses a unified statistical test derived from Bayesian approach. By changing the value of the hyper-parameter in the model, the proposed method can express various statistical test methods. This test has a clear mean and criterion in the sense of a statistical predictive distribution.
The author also developed a learning management system (LMS) called ‘Samurai’ that uses the online outlier detection method. This system supports online outlier detection of learners’ irregular learning processes and enables two-way instruction by mining data on learners’ learning processes. The system was used for actual classes and shown to be efficient.

2 Learning management system ‘Samurai’

In this section, an outline of an LMS I previously developed [13, 14] is introduced. The author has developed a LMS called ‘Samurai’ (See Ueno 2003), and has provided many e-learning courses (now the LMS provides 78 e-learning courses from my university). The LMS consists of a Contents Presentation System (CPS), a Contents Database (CD), a Learning Histories Database (LHD), and a Data Mining System (DMS). The CPS integrates various kinds of content and presents the integrated information on a web page.

Figure 1 shows a typical e-learning content presentation by Samurai. The contents are presented by clicking on the menu button. A sound track of the teacher’s narration is also presented according to [15], and the red pointer automatically moves as the narration proceeds.

This lesson corresponds to a 90-minute lecture at university and includes 42 topics. Although the content in Fig. 1 is text, the system also provides illustrations, animation or computer graphics, and video clips. In this lesson, there are 11 text contents, eleven illustrations, ten animations, and ten video clips. The system also

![Figure 1: Example of e-learning instruction.](image-url)
presents some test items to assess the learners’ degree of comprehension as soon as lessons have been completed (Fig. 2).

The CD consists of various kinds of media, text, jpeg, mpeg, and so on. The proposed platform monitors learners’ learning processes and saves them as log data in the LHD. The teacher makes prepares a lecture, and saves its contents in the CD. Then, the CPS automatically integrates the contents, and presents them to the learners.

The learners can learn them through the Internet. Learning history log data are saved in the LHD and analyzed in the DMS. The DMS provides feedback to the learners and teacher.

The LMS monitors learners’ learning processes and stores them as log data in the LHD. The stored data consists a Contents ID, a Learner ID, the number of topics which the learner has completed, a Test Item ID, an Operation Order ID (which indicates what operation was done), a Date and Time ID (which indicates the time and date of an operation started), and a Time ID (which indicates the time it took to complete operation). This data enables the system to recount the learner’s behaviors in e-learning.

3 Online outlier detection

3.1 Data

Response time data is used to detect learner’s irregular learning processes, as shown in Fig. 3. The horizontal axis indicates the number of content items a learner has accessed, and the vertical axis indicates the response time for the content.
The horizontal axis indicates the number of contents which a learner has learned, and the vertical axis indicates the response time for the content. Using this data, we can discover irregular learning processes. To do this, this chapter proposes a new method to detect learner’s irregular learning processes. The main idea is to derive a Bayesian predictive distribution of a new data item, $x_{n+1}$, given a learner’s learning processes data, $x_1, \ldots, x_n$, and provides a test for outlier detection of the new data item.

### 3.2 Model

This section derives a Bayesian predictive distribution of a new data item, $x_{n+1}$, given a learner’s learning processes data, $x_1, \ldots, x_n$. Let $w_{ij}$ be a learner $j$’s response time for the $i$th content item, and consider the following linear equation

$$x_{ij} = \frac{w_{ij} - \overline{w}_i}{s_i} = \mu_j + e_j,$$

where $\overline{w}_i$ indicates the average response time and $s_i$ indicates the variance of response time of $i$th content item.

A Bayesian predictive distribution of a new data $x_{n+1}$ given the learner’s learning processes data item, $x_1, \ldots, x_n$, can be derived as follows:

$$p(x_{n+1}|X) = \int \int p(x_{n+1} | \mu, \sigma^2)p(\mu, \sigma^2 | x_1, \ldots, x_n) d\mu d\sigma^2$$

$$= \left(1 + \left[\frac{(x_{n+1} - \mu_*)}{\sqrt{\frac{n_0 + n + 1}{n_0 + n} \lambda_*^2}}\right]^2 \right)^{-\frac{v+1}{2}},$$

where

$$t = \frac{(x_{n+1} - \mu_*)}{\sqrt{\frac{(n_0 + n + 1)\lambda_*^2}{(n_0 + n)v}}}.$$  

Then, $t$ has a $t$ distribution with degrees of freedom $v = n_0 + n - 1$. 

---

**Figure 3:** Learning time data for content items.
Here, $\mu_*$ indicates the hyper parameter of the prior distribution, which is the mean parameter of the normal distribution. The value of $\mu_*$ is determined by the mean of data item $x$. In this case, the data $x$ is standardized with mean zero and standard deviation one, so $\mu_*$ is zero. This prior can combine the prior knowledge with objective data analysis. For examples, even if a learner begins by using an irregular learning processes, the proposed method does not regard the processes as regular processes.

### 3.3 Model

Using eqn (3), we can detect outliers of learning processes. The procedure is as follows:

1. Get a new data item $x_{ij}$ during a learner is learning
2. Calculate the value of $t$ using eqn (3)
3. If the value of $t$ is greater than the value of $t$ in a $t$ distribution with $\alpha$ or the value of $t$ is less than the value of negative $t$ in the $t$ distribution with $\alpha$, then the new data item is detected as an outlier.

One unique feature of this method is that it unifies various statistical test methods. Changing the value of the hyper-parameter, $n_0$, has the following effect:

- When $n_0$ is large enough, the method is equivalent to the Z test.
- When $n_0$ is equivalent to zero (called a non-information prior distribution), the method is equivalent to the traditional $t$ test.

Thus, the proposed method unifies two major traditional test methods. In addition, the proposed method expresses various statistical methods which have different characteristics from the traditional test methods.

### 3.4 Outlier detection curves and examples

To detect irregular learning processes, this study proposes a method called outlier detection curve. The curve corresponds to a learner’s learning processes to a lecture. For example, Fig. 4 shows learner 4’s (outlier) detection curve, corresponding to Fig. 3. The four parallel lines in Fig. 4 indicate the outlier detection line. For example, if the $t$ value corresponding to a learning process exceeds the top detection curve, it means that the learning process was too long. If the $t$ value exceeds the bottom detection curve, it means that the learning process was too short.

In Fig. 4, outlier processes were used to learn content items 129–145. We should note that the response times, which may seem comparatively very long or short in Fig. 3, are not always judged as outlier processes. The reason is that the statistical value of $t$ for the outlier detection is estimated by considering both the student’s ability to learn and the difficulty of content items. For example, even if we find...
responses that took longer than others, we cannot say that they are outlier processes because the content may require more time than other content.

These features are quite different from the traditional data mining methods using the outlier system; for example, discovering robberies using depositing processes in banks and discovering invasions into computer networks do not require such complex methods because they need not consider respectively different tasks in a series.

An example of raw learning time data and detection curve when outlier processes are rare is shown in Fig. 5. From the raw data, it seems that learning times for content items 72–92 are irregularly long, but only the learning processes for items 77 and 78 are detected as outliers. An example of raw learning time data and detection curve when outlier processes are common is shown in Fig. 6. Note that the shape of the curve that shows the raw learning time data in Fig. 6 is very similar to one in Fig. 5. However, the outlier detection curve in Fig. 6 shows many outliers in the learning processes. This learner may not have studied hard. Thus, the proposed method can detect outlier learning processes that we cannot notice by only analyzing raw learning time data. In practical use for distance education, a teacher cannot track all students’ learning processes. The teacher can detect learners who should be taken care of and send them e-mails with messages like ‘Did you have trouble understanding some of the content?’, ‘Are you studying hard enough?’, and ‘Are you bored?’. These messages will show students who have some learning problems that the teacher notices them, and I expect that it will motivate their learning.

4 Simulation experiments

Although the proposed method represents various statistical test methods using a hyper-parameter, \( n_0 \), the way to determine the value of \( n_0 \) is unknown. In this
Figure 5: Example of detection curve with few outlier processes.

Figure 6: Example of a detection curve with many outlier processes.
section, I describe some simulation experiments to determine the optimum value of the hyper-parameter. The flow of the simulation experiments is as follows:

Fix a learner, \(j\), and generate random data using

\[
x_{ij} = \frac{t_{ij} - \bar{t}_i}{s_i} = \mu_j + e_j.
\]

- Apply my method to the generated data.
- Repeat these procedures 1000 times.
- Calculate the probabilities that the method correctly detects irregular processes and incorrectly detects regular processes by changing the value of the hyper-parameter.

The results are shown in Table 1. Column \(n\) indicates which random data sequences were used for the outlier detection. For example, 1–10 indicates that the outlier detection procedure used the first through tenth items in the random data sequence. The calculated probabilities are the averages of the probabilities that the method incorrectly detects regular processes and the probabilities that the method correctly detects irregular processes using different values of the hyper-parameter. The results show that the probability that the method correctly detects irregular processes in each range of \(n\) increases as the hyper-parameter increases. As the hyper-parameter increases, the probabilities get closer to the probabilities of the \(Z\) test. The results also show that the probability that the method incorrectly detects regular processes in each range of \(n\) increases as the value of the hyper-parameter decreases. The probabilities when \(n_0 = 0\) are equivalent to the probabilities of the \(t\) test. Thus, we have to consider the balance between the two probabilities when choose the value of the hyper-parameter. Here, this considers that minimizing the probability that the method incorrectly detects regular processes is important, so this study used \(n_0 = 1\).

5 System

The author developed a LMS including the outlier detection system. The outlier detection system is shown in Fig. 7. The system presents a graph called ‘Outlier detection curve’ as shown in Fig. 7. Figure 7 corresponds to a learner’s learning processes to a lecture.

The horizontal axis of the graph indicates the number of content items a learner has accessed, and the vertical axis indicates the value of \(t\) in eqn (3). Furthermore, there are two lines which indicate the value of \(t\) in a \(t\) distribution with \(\alpha\) in the graph. If the value of \(t\) is greater than the value of \(t\) in a \(t\) distribution with \(\alpha\) or the value of \(t\) is less than the value of negative \(t\) in the \(t\) distribution with \(\alpha\), then the new data item is detected as an outlier. For example, Fig. 4 shows that the number of irregular processes increases as the learner proceeds to learn the contents items. Thus, the learners’ irregular learning processes are detected using an online system. If an irregular process is detected, the teacher investigates the learner’s learning processes and sends an e-mail with some comments later.
Table 1: The outlier detection results changing the value of the hyperparameter.

<table>
<thead>
<tr>
<th>n</th>
<th>( Z_{\text{test}} )</th>
<th>( n_0 = 0 )</th>
<th>( n_0 = 1 )</th>
<th>( n_0 = 5 )</th>
<th>( n_0 = 10 )</th>
<th>( n_0 = 15 )</th>
<th>( n_0 = 20 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1–10</td>
<td>0.37</td>
<td>0.07</td>
<td>0.057</td>
<td>0.10</td>
<td>0.15</td>
<td>0.21</td>
<td>0.25</td>
</tr>
<tr>
<td>11–20</td>
<td>0.47</td>
<td>0.12</td>
<td>0.11</td>
<td>0.19</td>
<td>0.28</td>
<td>0.36</td>
<td>0.42</td>
</tr>
<tr>
<td>21–30</td>
<td>0.44</td>
<td>0.11</td>
<td>0.10</td>
<td>0.16</td>
<td>0.24</td>
<td>0.31</td>
<td>0.37</td>
</tr>
<tr>
<td>31–40</td>
<td>0.47</td>
<td>0.14</td>
<td>0.12</td>
<td>0.20</td>
<td>0.28</td>
<td>0.37</td>
<td>0.43</td>
</tr>
<tr>
<td>41–50</td>
<td>0.46</td>
<td>0.15</td>
<td>0.14</td>
<td>0.20</td>
<td>0.29</td>
<td>0.37</td>
<td>0.43</td>
</tr>
<tr>
<td>51–60</td>
<td>0.46</td>
<td>0.15</td>
<td>0.14</td>
<td>0.21</td>
<td>0.29</td>
<td>0.37</td>
<td>0.43</td>
</tr>
<tr>
<td>61–70</td>
<td>0.46</td>
<td>0.15</td>
<td>0.14</td>
<td>0.21</td>
<td>0.29</td>
<td>0.37</td>
<td>0.43</td>
</tr>
<tr>
<td>71–80</td>
<td>0.45</td>
<td>0.15</td>
<td>0.14</td>
<td>0.21</td>
<td>0.28</td>
<td>0.37</td>
<td>0.43</td>
</tr>
<tr>
<td>81–90</td>
<td>0.44</td>
<td>0.16</td>
<td>0.15</td>
<td>0.21</td>
<td>0.29</td>
<td>0.37</td>
<td>0.43</td>
</tr>
<tr>
<td>91–100</td>
<td>0.45</td>
<td>0.16</td>
<td>0.15</td>
<td>0.22</td>
<td>0.29</td>
<td>0.37</td>
<td>0.43</td>
</tr>
</tbody>
</table>

Probabilities that method correctly detects irregular processes:

<table>
<thead>
<tr>
<th>n</th>
<th>( Z_{\text{test}} )</th>
<th>( n_0 = 0 )</th>
<th>( n_0 = 1 )</th>
<th>( n_0 = 5 )</th>
<th>( n_0 = 10 )</th>
<th>( n_0 = 15 )</th>
<th>( n_0 = 20 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1–10</td>
<td>0.89</td>
<td>0.72</td>
<td>0.74</td>
<td>0.82</td>
<td>0.86</td>
<td>0.87</td>
<td>0.88</td>
</tr>
<tr>
<td>11–20</td>
<td>0.98</td>
<td>0.88</td>
<td>0.90</td>
<td>0.97</td>
<td>0.99</td>
<td>0.99</td>
<td>1.00</td>
</tr>
<tr>
<td>21–30</td>
<td>0.95</td>
<td>0.84</td>
<td>0.85</td>
<td>0.92</td>
<td>0.95</td>
<td>0.95</td>
<td>0.96</td>
</tr>
<tr>
<td>31–40</td>
<td>0.98</td>
<td>0.91</td>
<td>0.92</td>
<td>0.97</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
</tr>
<tr>
<td>41–50</td>
<td>0.99</td>
<td>0.97</td>
<td>0.97</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>1.00</td>
</tr>
<tr>
<td>51–60</td>
<td>0.99</td>
<td>0.97</td>
<td>0.98</td>
<td>0.95</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>61–70</td>
<td>0.99</td>
<td>0.98</td>
<td>0.98</td>
<td>0.99</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>71–80</td>
<td>0.99</td>
<td>0.98</td>
<td>0.99</td>
<td>0.99</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>81–90</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>91–100</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>

6 Evaluation

In this section, some experiments to evaluate the effects of this system are described. The author gave the learners the following instructions during their learning.

‘If you feel that the learning process you used in this section had some problems, please click the ‘No’ button. If you understand the content, please click the ‘Yes’ button.’

The learners’ responses are divided into three categories: yes, no, and no response. We can consider that students who answer ‘Yes’ have no problems, but those who answer ‘No’ or give no response might have some problems. We can consider that a ‘No’ response indicates that the learner did not understand the content and that no response indicates that the learner might not have completed the content.

Table 2 shows the conditional probabilities of the responses to the above question given the learners’ learning processes are or are not detected as irregular.
Figure 7: Online outlier detection system.

Did you understand this topic?

☐ Yes  ☐ No

Figure 8: Yes–No button used in experiment.

Table 2: Comparisons of detected irregular processes and learners’ statements.

<table>
<thead>
<tr>
<th>Learning process detected as irregular</th>
<th>Learning process not detected as irregular</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. ‘Yes’</td>
<td>0.24</td>
</tr>
<tr>
<td>2. ‘No’</td>
<td>0.31</td>
</tr>
<tr>
<td>3. No response</td>
<td>0.45</td>
</tr>
</tbody>
</table>

Table 2 indicates that the detection system efficiently detected learners’ irregular processes. The probability of a ‘Yes’ response when a learning process is detected as an irregular process is comparatively large; this indicates that the detection criterion is somewhat strict.
7 Animated agent to enhance learning

Although the feedback of the outlier detection to the teacher is presented as shown Fig. 7, it is very difficult for a teacher to understand and interpret the outlier detection curve. In addition, even if the teacher can understand the curve, it will overload the teachers much to send motivational messages to the detected students by e-mail. To solve this problem, the author installed an animated agent to provide some motivated messages generated from the data mining results into the LMS shown in Figs 1, 9 and 10. In this system, the agent provides some messages to the detected students, e.g. ‘Didn’t you skip the content 21, 22, 23, and 29? Try again’, ‘You took a long time for the content 41, 42, and 51. If you are having trouble with the contents, please consult your teacher.’, and so on.

Furthermore, there is evidence that the agent system is effective. Ueno [16] analyzed that learners begin to be bored of their learning every 18 min, so he reported that the animated agent system is very effective in attracting the learners’ motivation.

In addition, the LMS ‘Samurai’ provides various motivational messages besides the previous messages using the outlier detection system. The message generation system is as follows [17].

The method is to apply a data mining method to the huge amount of stored data and construct a learner model to predict each learner’s final status: (1) Failed (Final examination score below 60); (2) Abandon (The learner withdraws before the final examination); (3) Successful (Final examination score is more than 60 but less than 80); and (4) Excellent (Final examination mark is more than 80). For this purpose, the well-known data mining method, Decision Tree [18], is employed using the following variables reflecting each learner’s status each week:

- Probability is 30% now.
- The progress of your lesson is slow. Please take a lecture more.

Figure 9: An intelligent agent system.

Figure 10: Various actions of the agent.
1. The number of topics which the learner has learned.
2. The number of times the learner accessed the e-learning system.
3. The average number of times the learner has completed each topic. (This implies the time the learner repeated each topic.)
4. The average learning time for each lecture, which consists of several types of contents and runs 90 min.
5. The average of the degree of understanding of each topic. (This is measured by the response to the question which is corresponding to each topic.)
6. The average learning time for each course which consists of fifteen lectures.
7. The average number of times the learner has changed the answer to the questions in the e-Learning.
8. The number of times which the learner has posted opinions or comments to the discussion board.
9. The average learning time for each topic.

Because all courses run for 15 weeks, fifteen decision trees are estimated corresponding to learners’ learning histories data for the fifteen weeks. We use the ID3 algorithm [18] as a learning algorithm for the decision trees. The program was developed using Java and installed in Samurai. The decision trees are always estimated using updated learning histories. Therefore, the decision trees structures for predicting the learner’s final status always change. In this algorithm, all variables are always used. A decision tree learned from 1,344 learners’ data is shown in Fig. 11. This tree was estimated using 14 weeks of learning history data. The two values in parentheses indicate the number of cases in which the inference is correct and incorrect. For example, (408/18) indicates that the probability of the correct inference is 408/426. In this system, decision trees corresponding to the weekly learner’s status are always being constructed.

The agent provides adaptive messages to the learner using the learner model. The agent system also performs various actions based on the learner’s current status as shown in Fig. 10. The instructional messages to a learner are generated as follows:

1. The system predicts the target learner’s future status and it’s probability using the constructed decision tree.
2. If the predicted status is ’Excellent’, then the agent provides messages like ’Looking great!’, ‘Continually do your best.’, and ’Probability of success is xx%’. If the predicted status is not ’Excellent’, the system then searches for the closest ‘Excellent’ node from the current predicted status node. For example, let us consider a part of the decision tree in Fig. 11 (see Fig. 12). If the predicted status is ’Failed’, the nearest node ‘Excellent’ is the gray node in the figure. The system finds the nearest node ‘Excellent’ and determines the operations that will change the learner’s predicted status to ’Excellent’. In this case, ’the average learning time for each topic’ is detected. The system provides the messages with the predicted future status, the probability of success estimated by the decision tree, and the instructional messages according to Table 3.
Ueno [17] reported that this agent system reduced the dropout rate of e-learning to about one-third. This is a remarkable effect of the agent system.

8 Conclusions

This chapter proposed a method of online outlier detection of learners’ irregular learning processes using learner response times to e-learning content. The unique
Table 3: Instructional messages corresponding to the detected variables.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Instructional messages</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. The number of topics which the learner has learned.</td>
<td>1. Progress in your lesson is behind. Please take more lectures.</td>
</tr>
<tr>
<td></td>
<td>2. Progress in the lesson is liable to be slow. Let’s take more lectures.</td>
</tr>
<tr>
<td>2. The number of times the learner accessed the e-learning system.</td>
<td>3. You have not engaged the lesson well. Let’s access the system more.</td>
</tr>
<tr>
<td>3. The average number of times the learner has completed each topic.</td>
<td>4. Don’t forget the previous contents! Let’s confirm the previous contents again.</td>
</tr>
<tr>
<td>4. The average learning time for each lecture, which consists of several</td>
<td>5. It seems that you are working through the lecture too quickly. Please spend more</td>
</tr>
<tr>
<td>types of contents and runs 90 minutes.</td>
<td>time on each lecture.</td>
</tr>
<tr>
<td>5. The average of the degree of understanding of each topic (This is</td>
<td>6. Were the contents of the lesson difficult? Let’s take the lecture from the</td>
</tr>
<tr>
<td>measured by the response to the question which is corresponding to</td>
<td>beginning once again.</td>
</tr>
<tr>
<td>each topic).</td>
<td>7. When there is something you don’t understand, let’s ask a question on a discussion</td>
</tr>
<tr>
<td></td>
<td>board.</td>
</tr>
<tr>
<td>6. The average learning time for each course which consists of fifteen</td>
<td>8. You have not engaged in the lesson well. Let’s access the system and study more</td>
</tr>
<tr>
<td>lectures.</td>
<td>slowly and carefully.</td>
</tr>
<tr>
<td>7. The average number of times the learner has changed the answer to</td>
<td>9. It looks as if your knowledge is not so robust. Let’s take a lecture from the</td>
</tr>
<tr>
<td>the questions in the e-Learning.</td>
<td>beginning once again.</td>
</tr>
<tr>
<td>8. The number of times which the learner has posted opinions or comments</td>
<td>10. Learning is better done between learners. Let’s participate in and contribute to</td>
</tr>
<tr>
<td>to the discussion board.</td>
<td>the discussion board.</td>
</tr>
<tr>
<td>9. The average learning time for each topic.</td>
<td>11. Did you do the lecture correctly? Ordinarily, a lesson should take more time.</td>
</tr>
</tbody>
</table>

Features of this method are as follows: (1) It uses an outlier detection method using a Bayesian predictive distribution. (2) It can be used for small samples. (3) It can conveniently calculate predictive distributions. (4) It assists two-way instruction using data mining results on learners’ learning processes. (5) The outlier statistics are estimated by considering both students’ abilities and the difficulty of content.
The system was used for actual classes, and the results show the efficiency of the system.

The advantages of this outlier detection method are that it represents various statistical test methods by using a changing hyper-parameter and that the model, which depends on learners’ learning ability and the difficulty of items, provides results that cannot be seen by only looking at raw learning time data.

In addition, the author proposed an animated agent which provides motivational messages to the learners using the data mining methods. The evaluation results show that these methods are effective enough.

References


