Is There ‘ONE way’ of Learning? A Data-driven Approach

Jauwairia Nasir  
CHILI Lab, EPFL  
Lausanne, Switzerland  
jauwairia.nasir@epfl.ch

Barbara Bruno  
CHILI Lab and MOBOTS group in BioRob Lab, EPFL  
Lausanne, Switzerland  
barbara.bruno@epfl.ch

Pierre Dillenbourg  
CHILI Lab, EPFL  
Lausanne, Switzerland  
pierre.dillenbourg@epfl.ch

ABSTRACT

Intelligent Tutoring Systems (ITS) are required to intervene in a learning activity while it is unfolding, to support the learner. To do so, they often rely on performance of a learner, as an approximation for engagement in the learning process. However, in learning tasks that are exploratory by design, such as constructivist learning activities, performance in the task can be misleading and may not always hint at an engagement that is conducive to learning. Using the data from a robot mediated collaborative learning task in an out-of-lab setting, tested with around 70 children, we show that data-driven clustering approaches, applied on behavioral features including interaction with the activity, speech, emotional and gaze patterns, not only are capable of discriminating between high and low learners, but can do so better than classical approaches that rely on performance alone. First experiments reveal the existence of at least two distinct multi-modal behavioral patterns that are indicative of high learning in constructivist, collaborative activities.

CCS CONCEPTS

• Human-centered computing → User studies; Collaborative and social computing design and evaluation methods.

KEYWORDS

collaborative learning, constructivist classrooms, engagement, productive engagement, learning analytics, multi-modal data

ACM Reference Format:

1 INTRODUCTION

Two of the most important components in an ITS include a student model and a pedagogical model [1, 8]. The pedagogical model is responsible for making appropriate intervention[s] in the activity (i.e., interventions that have a positive effect on the student’s learning), knowing the details of the learning activity and being informed by the student model about the student’s status. Bayesian Knowledge Tracing (BKT) [3] is one of the most widely used approaches to model student knowledge [4, 10, 13]. One of the assumptions in BKT is that at each step, the ‘student can either succeed or fail the task’, i.e. there is a straightforward, binary mapping between performance in the task and learning, which makes the approach most “relevant for tutors that use exercises and scaffolding as the main vehicle for learning” [4]. However, there is an increasing emphasis towards incorporating more open-ended/constructivist learning activities that encourage the awareness of the knowledge construction process, e.g. by promoting, among other things, Problem Based Learning that requires the learners to devise a solution to a real world problem together, and/or Cooperative Learning in which interdependence among group members is needed to solve a problem; thus, violating the requirement of a chain of binary right/wrong steps towards the goal [2, 12]. The learners rather “become engaged by applying their existing knowledge and real-world experience, learning to hypothesize, testing their theories, and ultimately drawing conclusions from their findings” [9]. As a consequence, a constructivist learning environment makes it non-trivial to accurately identify when and if a learner is productively engaged in the task, i.e., engaged in a way that is conducive to learning. Some of the latest research, for example L2TOR1 and EMOTE2 projects, has looked into advancing adaptive tutoring systems; however, not much work is found in the domain of ITS to model student knowledge in constructivists environments.

Here, we utilize data from JUSTThink: an open-ended activity in which two learners cooperate as a team to solve a semi-guided algorithmic reasoning problem [6, 7]; hence, following both the Problem Based Learning and Cooperative Learning paradigms. In

This is the author’s version of the work. It is posted here for your personal use. Not for redistribution. The definitive Version of Record was published in: 
ICMI ’20 Companion, October 25–29, 2020, Virtual event, Netherlands © 2020 Copyright held by the owner/author(s).
https://doi.org/10.1145/3395035.3425200

1http://www.l2tor.eu/
2http://www.emote-project.eu/
a previous analysis on these data, we observed that performance of the learners in the task was not directly correlated to learning gains, while it affected how the learners evaluate their competence, task engagement and mutual understanding [6]. We postulate the use of data-driven approaches on multi-modal behavioural features to make subtle learning-related behavioural cues emerge, and thus allow for the identification of high learners in a way which is independent from performance and human annotations, hence paving a way for novel ITS and, more broadly, more effective learning-supporting systems and interventions. It must be noted that here performance and learning refers to two very different things. By performance, we mean failing or succeeding at the task; while by learning we mean how the learners score at their pre- and post-tests.

2 SETUP

We utilize the dataset generated from a study done with the JUS-Think platform [6] where 68 children (in teams of two i.e., 34 teams) aged 9 to 12 years interact with a collaborative learning platform consisting of two screens and a QTrobot acting as a guide and a mediator (see Fig. 1). For post analysis of the collaborative behavior, we look at: 1) Multi-modal Behaviors and 2) Learning and Performance Measures. For the former, we look at the children’s interactions with the system, affective states, gaze and speech patterns that are extracted automatically using various sensors [6]. In total, there are 28 features where we use averages and frequencies for various behavioral features. For the latter, we take into account the team level learning measures such as absolute, relative, and joint absolute learning gain (A_LG, R_LG, J_A_LG, respectively); and the performance measure as the last error (L_E) in their solution. These measures are discussed and elaborated in an earlier work [6].

3 RESEARCH QUESTION AND OUR APPROACH

To explore whether the way in which children learn can be identified from their multi-modal behaviors in a collaborative exploratory activity, we propose the technique graphically summarized in Figure 2. Under approach A, the teams are clustered according to their learning as well as performance (A-1) and the resulting cluster labels

<table>
<thead>
<tr>
<th>Cluster no.</th>
<th>L_E</th>
<th>A_LG</th>
<th>R_LG</th>
<th>J_A_LG</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Approach A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.166</td>
<td>0.433</td>
<td>0.438</td>
<td>0.357</td>
<td>6</td>
</tr>
<tr>
<td>2</td>
<td>0.681</td>
<td>0.720</td>
<td>0.748</td>
<td>0.742</td>
<td>10</td>
</tr>
<tr>
<td>3</td>
<td>0.836</td>
<td>0.340</td>
<td>0.168</td>
<td>0.314</td>
<td>5</td>
</tr>
<tr>
<td>4</td>
<td>0.132</td>
<td>0.700</td>
<td>0.733</td>
<td>0.769</td>
<td>13</td>
</tr>
<tr>
<td>Approach B</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.461</td>
<td>0.678</td>
<td>0.693</td>
<td>0.714</td>
<td>14</td>
</tr>
<tr>
<td>2</td>
<td>0.393</td>
<td>0.616</td>
<td>0.604</td>
<td>0.607</td>
<td>12</td>
</tr>
<tr>
<td>3</td>
<td>0.393</td>
<td>0.383</td>
<td>0.348</td>
<td>0.428</td>
<td>6</td>
</tr>
</tbody>
</table>

Table 1: Clusters in each approach with mean values for learning gains (LG) as well as the last error.

4 FINDINGS

In both approaches, we rely on k-means as clustering method, using entropy analysis to properly set the parameter k, and both SVM and Random Forest as classifiers, similar to [5, 11] in steps A-3 and B-4. For both classifiers we rely on k-fold cross validation and parameters setting via grid-search.

The clustering step (A-1) in approach A identifies four different clusters, representing the four different combinations of high/low learning gains/last error (see Table 1).

As the upper half of Table 2 shows, using these Type 1 labels results in a very low accuracy and F1 score with both the SVM and Random Forest classifiers, suggesting that the Type 1 labels are not a good fit for the multi-modal features. On the other hand, the Type 2 labels, denoted Type 2 labels, could then be used as ground truth (B-3) for a classifier trained on the teams’ multi-modal behavioural features, to discriminate between high and low learners based on their behaviour (A-3).

Table 2: Classification Results

<table>
<thead>
<tr>
<th>Classifier</th>
<th>k-fold cross-validation</th>
<th>test-set</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Accuracy</td>
<td>F1-score</td>
</tr>
<tr>
<td>SVM</td>
<td>0.28</td>
<td>0.24</td>
</tr>
<tr>
<td>RF</td>
<td>0.49</td>
<td>0.29</td>
</tr>
<tr>
<td>SVM</td>
<td>0.80</td>
<td>0.74</td>
</tr>
<tr>
<td>RF</td>
<td>0.80</td>
<td>0.77</td>
</tr>
</tbody>
</table>

The symbols in the image can be attributed to the Noun Project and FAVPNG.
The initial findings highlight a number of research hypotheses (B-1), allowing for achieving very high accuracy and F1 scores both with different (p-values of \( p = 0.016, p = 0.025, \) and \( p = 0.026 \) obtained by Kruskal-Wallis test) from the learning gains exhibited by the third behavioral cluster, which is associated more with low learning teams (see Table 1). This means that condition (1) and (2), highlighted in section 3, are met.

Hence, these two conditions being met allow for a perspective shift as mentioned in the previous section: using the cluster numbers generated on multi-modal data as labels for supervised learning models where we distinguish productive from non-productive learners. As can be seen in the bottom half of table 2, Type 2 labels allow for achieving very high accuracy and F1 scores both with SVM (\( C = 1000, \) \( kernel = rbf, \) \( y = 0.001, \) \( class\_weight = balanced \)) and RF (\( n\_estimators = 100, \) \( max\_features = auto, \) \( class\_weight = balanced \)).

5 CONCLUSION
The initial findings highlight a number of research hypotheses for ITS in constructivist designs that are worth exploring: 1) performance is not always a reliable indicator of engagement in the learning process, as hinted by the analysis conducted in step B-2; 2) learners having similar learning gain and performance profiles can have multiple sets of multi-modal behaviors (as supported by the two clusters of high-learning teams identified with approach B). This could also be the reason why the Type 1 labels did not perform well for classification; and 3) the use of behavioural labels, proven solidly linked with learning, seems to allow for a better discrimination between high and low learners than the direct use of learning labels, as suggested by the classification results of the two approaches reported in Table 2.

Thorough analyses investigating the above-listed research directions will give a deeper understanding of which multi-modal behaviour data and associated processing methods allow for effectively (and autonomously) separating low learners from high learners, to lead the way to effective interventions.

ACKNOWLEDGMENTS
This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 765955 (ANIMATAS). We are also grateful to the SNSF for supporting this project through the National Centre of Competence in Research Robotics.

REFERENCES


