Open Learner Models for Multi-Activity Educational Systems

Solmaz Abdi[1], Hassan Khosravi[1], Shazia Sadiq[1], and Ali Darvishi[1]

The University of Queensland, Brisbane, Australia
solmaz.abdi@uq.edu.au

Abstract. In recent years, there has been an increasing trend in the use of student-centred approaches within educational systems that engage students in various higher-order learning activities such as creating resources, creating solutions, rating the quality of resources, and giving feedback. In response to this trend, this paper proposes an interpretable and open learner model called MA-Elo that capture an abstract representation of a student's knowledge state based on their engagement with multiple types of learning activities. We apply MA-Elo to three data sets obtained from an educational system supporting multiple student activities. Results indicate that the proposed approach can provide a higher predictive performance compared with baseline and some state-of-the-art learner models.

Keywords: Learnersourcing · Open learner model · higher-order learning activity

1 Introduction

Learner models capture an abstract representation of a student's knowledge state. There are two main use cases for learner models: they are (1) employed as a key component of adaptive educational systems to provide personalised feedback or adaptivity functionalities and (2) externalised as open learner models (OLMs) [7, 8] to students with the aim of incentivising, and regulating learning. Commonly, learner models estimate a student's knowledge state only based on their performance on attempting (answering) assessment items. As a point of reference, many well-known approaches for learner modelling including Bayesian Knowledge Tracing (BKT) [11], Item Response Theory (IRT) [22], Adaptive Factor Models (AFM) [9], Performance Factor Analysis (PFA) [23], deep knowledge tracing (DKT) [25], and DAS3H [10], as well as various rating based learner models [2, 5, 21, 24] only employ students' performance on assessment items in their modelling. The reliance on only the performance of students on attempting assessment items can probably be explained by the fact that in many educational systems, students are prominently involved in just answering assessment items.

In recent years, contemporary models of learning have placed a great emphasis on the use of learner-centred approaches that involve students in higher-order learning activities. A well-recognised approach for doing so is to employ learnersourcing, which refers to a pedagogically supported form of crowdsourcing that partners with students to contribute novel content to teaching and learning while engaging in a meaningful learning experience themselves [17, 20]. Prior studies on learnersourcing, as well as evidence from the learning sciences, indicate that students have the ability to meaningfully contribute to teaching and learning activities such as creating and evaluating learning resources [3, 12, 13, 16, 29, 30] and that engaging with these activities enhances student learning [6, 14, 18, 28]

So, how can educational systems that engage students in a range of activities openly and accurately model student learning? Some of the recently proposed learner models employ data from student engagement with multiple activities towards more accurately modelling learners [1, 31]; however, they employ complex machine learning algorithms such as knowledge tracing machines [1] or tensor factorisation [31] which are not interpretable. We aim to address this limitation by proposing a multi-activity open and interpretable approach for modelling learners based on engagement with multiple types of learning activities.

2 Multi-Activity Knowledge Modelling

Problem formulation We denote students by $s_n \in \{s_1 \dots s_N\}$, learning resources (items) by $q_m \in \{q_1 \dots q_M\}$, and knowledge components (concepts) by $\delta_c \in \{\delta_1 \dots \delta_C\}$. Each item can be tagged with one or more concepts. We denote the relationship between items and concepts by $\omega_{mc} \in \Omega_{M \times C}$, where ω_{mc} is $1/f$ if item q_m is tagged with f concepts including δ_c, and 0 otherwise. Let $A = \{a_1 \dots a_k\}$ denote the different types of activities that students are allowed to perform (e.g., creating, evaluating, linking or attempting items). Finally, let's assume that the system records the interaction log for s_n on each type of activity a_k as $i_t^k = (s_n, q_m, a_k, t, r_{nmt}^k)$, where t index the timestamp of the interaction and r_{nmt}^k indicates the outcome of the interaction. If it is a graded activity and the outcome of the interaction is success then $r_{nmt}^k = 1$ and zero otherwise. For a non-graded activity, the outcome is always considered as success. Our aim is to employ interpretable methods to (1) infer a learner model for estimating s_n's knowledge state on each concept δ_c and (2) infer the difficulty of each item q_m.

Proposed approach Employing the popular method of using rating systems for modelling learners [2, 4, 5, 21, 24, 27], we present the Multi-Activity Elo-based learner model (MA-Elo), which is an extension over the multivariate Elo-based system [5], enabling interactions with multiple types of activities. To keep track of students' mastery, MA-Elo uses a two-dimensional array $\Lambda_{N \times C}$, where λ_{nc} represents student s_n's knowledge state on concept δ_c. For each item q_m, MA-Elo uses a global difficulty d_m approximating the difficulty of the item. For learning activities, MA-Elo considers two high-level categories. The first category incorporates activities in which the difficulty of learning items impacts the chance of a student's success. Examples of activities that fall into this category include attempting a learning item and creating a sample solution for an existing item. For each activity a_k in the first category, MA-Elo uses d_m of the item q_m associated in the activity to estimate the overall hardness of that activity for students. The second category consists of activities in which the chance of a student's success is independent of the difficulty level of the learning item (e.g., liking a resource). For each activity a_k in the second category, MA-Elo uses a global parameter h_k estimating the overall hardness of that activity. In practice, there are two options to calibrate the value of h_k: (1) a data-driven approach that treats h_k as a hyper-parameter and set it via cross-validation, or (2) the domain expert determines the relative difficulty of each of the learning activities. Whenever a student s_n performs a learning activity related to item q_m, MA-Elo first investigates if the activity comes from the first category or not and then uses the following equation to compute the chance of s_n's success:

$$P(r_{nmt}^k = 1) = \begin{cases} \sigma(\sum_{l=1}^{L} \lambda_{nc} \times \omega_{mc} - d_m), & \text{if the activity is from the first category} \\ \sigma(\sum_{l=1}^{L} \lambda_{nc} \times \omega_{mc} - h_k), & \text{otherwise} \end{cases}$$

where $\sigma(.)$ is the sigmoid function and $\sum_{l=1}^{L} \lambda_{nc} \times \omega_{mc}$ estimates s_n's weighted average competency on the concepts that are associated with q_m. MA-Elo then updates the student's mastery on each concept δ_l the question is tagged with based on the type of activity that is performed using $\lambda_{nl} := \lambda_{nl} + \zeta_k \cdot (r_{nmt}^k - P(r_{nmt}^k = 1))$, where r_{nmt}^k is the outcome of the interaction and ζ_k is a constant determining the sensitivity of the estimations based on the student's last interaction of the activity of type a_k. In addition, if the interaction was from the first category of activities, concurrent with updating the estimations of the student's knowledge state, the estimations of the model about the difficulty of the item q_m is also updated using $d_m := d_m + U(n) \cdot (P(r_{nmt}^k = 1) - r_{nmt}^k)$, where $U(n)$ is an uncertainty function used for stabilising the estimates of item difficulty and is computed as $U(n) = \frac{\gamma}{1+\beta * n}$, where γ and β are constant hyper-parameters determining the starting value and slope of changes, respectively, and n indicates the number of prior updates on the item difficulty [24].

3 Evaluations

To evaluate MA-Elo, we use three historical data sets obtained from an educational system called RiPPLE and compare the predictive performance of MA-Elo with five existing learner models. At its core, RiPPLE is learnersourcing adaptive educational system that recommends learning items to students based on their estimated mastery level from a pool of items learnersourced by their peers [19]. RiPPLE enables students to engage with three main types of activities within the system, namely (1) practising learning items, (2) creating new items to be added to the repository of the system, and (3) moderating learning items in which students are involved in reviewing and evaluating learning items. Please refer to [19] for the detailed information about RiPPLE, the interface used for learning item creation and learning item moderation, and the formulation of the consensus approaches used by RiPPLE for each of these tasks. The three data sets used in the experiment as outlined in Table 1 are named (1) Introduction to Information Systems (InfoSys), (2) The Brain and Behavioural Sciences (NEUR) and, (3) Artificial Intelligence (AI). For our analysis to be consistent with the prior works (e.g., [10, 26, 31]), we evaluated the predictive performance of the models using 5-fold cross-validation where each data set split was done at the student-level. We compare the predictive performance of MA-Elo to IRT, PFA, AFM, and DAS3H. For this comparison, we use the implementation of these models provided by [15]. We also compare the predictive performance of MA-Elo to Multivariate-Elo [5], which is the most similar single-activity Elo-based learner model to our proposed model. Given the three main learning activities that

Table 1: RiPPLE Data sets

Data set	Students	Items	Concepts	Practice	Create	Moderate	Interactions
InfoSys	422	2008	7	47,122	940	4,586	52,648
NEUR	519	2,836	7	26,933	2,852	628	30,413
AI	322	1,312	12	19,031	1,305	6,475	26,811

Table 2: AUC and RMSE for the RiPPLE data sets.

Model	InfoSys		NEUR		AI	
	AUC	MSE	AUC	MSE	AUC	MSE
IRT	0.688	0.203	0.740	0.189	0.726	0.197
AFM	0.571	0.222	0.533	0.225	0.550	0.229
PFA	0.619	0.216	0.610	0.218	0.592	0.224
DAS3H	0.719	0.197	0.747	0.183	0.724	0.203
Multivariate-Elo	0.722	0.199	0.741	0.187	0.726	0.205
MA-Elo	**0.730**	**0.193**	**0.758**	**0.183**	**0.737**	**0.200**

students are engaged within RiPPLE, without loss of generalisability, we implemented MA-Elo based on these three activities namely attempt (a_1), create (a_2), moderate (a_3). In addition, we only used interactions related to learning items of type MCQ. We conducted a grid search to determine the hyper-parameters of MA-Elo. Across all experiments, for MA-Elo, the value of ζ_1 (determining the sensitivity of the estimations when attempting learning items), is set to 0.4, the value of ζ_2 is set to 0.25, and the value of ζ_3 is set to 0.15. For each model, we report the area under the curve (AUC) and mean squared error (MSE).

As it is presented in Table 2, on all of the data sets, MA-Elo outperforms other learner models in terms of predictive performance. This outcome is aligned with findings from the existing literature on learnersourcing (e.g., [14]) that suggest engaging students in higher-order activities impacts their learning. MA-Elo is followed by both Multivariate-Elo and the state-of-the-art DAS3H model, which are ranked as the second best-performing models on the RiPPLE data sets. This finding shows that, in spite of simplicity, ease of implementation, and without necessitating pre-calibration on big samples of data, the models developed based on Elo rating system could perform as well as or even better than the best-performing learner models known in the literature and can be considered as practical models for the implementation of real-world educational systems.

4 Conclusion

The overarching goal of this paper is to address the problem of learner modelling in educational systems where in addition to answering assessment items, students are also engaged with multiple types of learning activities. To do so, we proposed a learner model called MA-Elo that leverages data from students engagement with different types of learning activities other than answering assessment items when modelling their learning. The results of our conducted experiment on three data sets obtained from an adaptive learnersourcing educational system suggest that MA-Elo provides higher predictive performance compared with conventional learner models. Future work aims to investigate the impact of opening MA-Elo to students and its potential impact on self regulation and student learning.

References

1. Abdi, S., Khosravi, H., Sadiq, S.: Modelling learners in crowdsourcing educational systems. In: Bittencourt, I.I., Cukurova, M., Muldner, K., Luckin, R., Millán, E. (eds.) Artificial Intelligence in Education. pp. 3–9. Springer International Publishing, Cham (2020)
2. Abdi, S., Khosravi, H., Sadiq, S.: Modelling learners in adaptive educational systems: A multivariate glicko-based approach. In: LAK21: 11th International Learning Analytics and Knowledge Conference. p. 497–503. LAK21, Association for Computing Machinery (2021)
3. Abdi, S., Khosravi, H., Sadiq, S., Demartini, G.: Evaluating the quality of learning resources: A learnersourcing approach. IEEE Transactions on Learning Technologies **14**(1), 81–92 (2021)
4. Abdi, S., Khosravi, H., Sadiq, S., Gasevic, D.: Complementing educational recommender systems with open learner models. In: Proceedings of the Tenth International Conference on Learning Analytics & Knowledge. p. 360–365 (2020)
5. Abdi, S., Khosravi, H., Sadiq, S., Gasevic, D.: A multivariate elo-based learner model for adaptive educational systems. In: Proceedings of the Educational Data Mining Conference. pp. 462–467 (2019)
6. Boud, D., Soler, R.: Sustainable assessment revisited. Assessment & Evaluation in Higher Education **41**(3), 400–413 (2016)
7. Bull, S., Ginon, B., Boscolo, C., Johnson, M.: Introduction of learning visualisations and metacognitive support in a persuadable open learner model. In: Proceedings of the Sixth International Conference on Learning Analytics & Knowledge. pp. 30–39. ACM (2016)
8. Bull, S., Kay, J.: Open learner models. In: Advances in intelligent tutoring systems, pp. 301–322. Springer (2010)
9. Cen, H., Koedinger, K., Junker, B.: Learning factors analysis–a general method for cognitive model evaluation and improvement. In: International Conference on Intelligent Tutoring Systems. pp. 164–175. Springer (2006)
10. Choffin, B., Popineau, F., Bourda, Y., Vie, J.J.: Das3h: Modeling student learning and forgetting for optimally scheduling distributed practice of skills. arXiv preprint arXiv:1905.06873 (2019)
11. Corbett, A.T., Anderson, J.R.: Knowledge tracing: Modeling the acquisition of procedural knowledge. User modeling and user-adapted interaction **4**(4), 253–278 (1994)
12. Darvishi, A., Khosravi, H., Sadiq, S.: Utilising learnersourcing to inform design loop adaptivity. In: European Conference on Technology Enhanced Learning. pp. 332–346. Springer (2020)
13. Denny, P., Hamer, J., Luxton-Reilly, A., Purchase, H.: Peerwise: students sharing their multiple choice questions. In: Proceedings of the fourth international workshop on computing education research. pp. 51–58 (2008)
14. Doroudi, S., Williams, J., Kim, J., Patikorn, T., Ostrow, K., Selent, D., Heffernan, N.T., Hills, T., Rosé, C.: Crowdsourcing and education: Towards a theory and praxis of learnersourcing. International Society of the Learning Sciences, Inc.[ISLS]. (2018)
15. Gervet, T., Koedinger, K., Schneider, J., Mitchell, T., et al.: When is deep learning the best approach to knowledge tracing? JEDM— Journal of Educational Data Mining **12**(3), 31–54 (2020)
16. Guo, P.J., Markel, J.M., Zhang, X.: Learnersourcing at scale to overcome expert blind spots for introductory programming: A three-year deployment study on the

python tutor website. In: Proceedings of the Seventh ACM Conference on Learning@ Scale. pp. 301–304 (2020)
17. Khosravi, H., Demartini, G., Sadiq, S., Gasevic, D.: Charting the design and analytics agenda of learnersourcing systems. In: LAK21: 11th International Learning Analytics and Knowledge Conference. p. 32–42. LAK21, Association for Computing Machinery, New York, NY, USA (2021)
18. Khosravi, H., Gyamfi, G., Hanna, B.E., Lodge, J.: Fostering and supporting empirical research on evaluative judgement via a crowdsourced adaptive learning system. In: Proceedings of the Tenth International Conference on Learning Analytics & Knowledge. p. 83–88 (2020)
19. Khosravi, H., Kitto, K., Williams, J.J.: Ripple: A crowdsourced adaptive platform for recommendation of learning activities. Journal of Learning Analytics **6**(3), 91–105 (2019)
20. Kim, J., et al.: Learnersourcing: improving learning with collective learner activity. Ph.D. thesis, Massachusetts Institute of Technology (2015)
21. Klinkenberg, S., Straatemeier, M., van der Maas, H.L.: Computer adaptive practice of maths ability using a new item response model for on the fly ability and difficulty estimation. Computers & Education **57**(2), 1813–1824 (2011)
22. Lord, F.M.: Applications of item response theory to practical testing problems. Routledge (2012)
23. Pavlik Jr, P.I., Cen, H., Koedinger, K.R.: Performance factors analysis–a new alternative to knowledge tracing. Online Submission (2009)
24. Pelánek, R., Papoušek, J., Řihák, J., Stanislav, V., Nižnan, J.: Elo-based learner modeling for the adaptive practice of facts. User Modeling and User-Adapted Interaction **27**(1), 89–118 (2017)
25. Piech, C., Bassen, J., Huang, J., Ganguli, S., Sahami, M., Guibas, L.J., Sohl-Dickstein, J.: Deep knowledge tracing. In: Advances in Neural Information Processing Systems. pp. 505–513 (2015)
26. Piech, C., Bassen, J., Huang, J., Ganguli, S., Sahami, M., Guibas, L.J., Sohl-Dickstein, J.: Deep knowledge tracing. Advances in neural information processing systems **28**, 505–513 (2015)
27. Reddick, R.: Using a glicko-based algorithm to measure in-course learning. In: Proceedings of the Educational Data Mining Conference. pp. 754–759. ERIC (2019)
28. Tai, J., Ajjawi, R., Boud, D., Dawson, P., Panadero, E.: Developing evaluative judgement: enabling students to make decisions about the quality of work. Higher Education **76**(3), 467–481 (2018)
29. Wang, X., Talluri, S.T., Rose, C., Koedinger, K.: Upgrade: Sourcing student open-ended solutions to create scalable learning opportunities. In: Proc. 6th (2019) ACM Conf. Learning@ Scale. pp. 1–10 (Jun 2019)
30. Zahirović Suhonjić, A., Despotović-Zrakić, M., Labus, A., Bogdanović, Z., Barać, D.: Fostering students' participation in creating educational content through crowdsourcing. Interactive Learning Environments **27**(1), 72–85 (2019)
31. Zhao, S., Wang, C., Sahebi, S.: Modeling knowledge acquisition from multiple learning resource types. arXiv preprint arXiv:2006.13390 (2020)