Interactive Teaching with Groups of Unknown Bayesian Learners

Carla Guerra, Francisco S. Melo, and Manuel Lopes

INESC-ID & Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal
carla.guerra@gaips.inesc-id.pt, fmelo@inesc-id.pt,
manuel.lopes@tecnico.ulisboa.pt

Abstract. In this work we empirically explore the extension of an interactive approach for machine teaching from single learners to groups of learners. We use interactivity to overcome the common mismatch between the knowledge the teacher has about the students and the students themselves. With a multi-learner setting we also investigated the best way to consider the class—as a whole or divided in partitions accordingly to the students priors. The results of an user study where we teach a Bayesian estimation task have shown that, regardless of considering partitions or not, the interactive approaches significantly increase the learning performance of the class when compared to non-interactive alternatives.

Keywords: Machine Teaching · Interactivity · Group-Learning

1 Introduction

Machine teaching (MT) considers the problem of finding the smallest set of examples that allows a specific learner to acquire a given concept, explicitly considering a computational learning algorithm for the student [5, 6, 8]. Since a significant amount of teaching relies on providing examples, the learning efficiency can be greatly improved if the teacher selects the examples that are more informative for each particular learner using MT techniques. The main problem with MT is that it often assumes that the learner, or the learning algorithm, is completely known. This is a very strong assumption that does not hold in the general case, and certainly not in the case where the learner is a human. Melo et. al [4] explicitly address the unavoidable mismatch between what the machine teaching system assumes about the learner and the learner himself and propose interactivity as the means to overcome it. However, that work and most of MT research so far has focused on single learner scenarios. There are not yet many advances in machine teaching applied in a setting where the teacher must teach multiple learners although this is the reality in real-world classrooms. The students have different backgrounds and prior knowledges, which the teacher must take into consideration when delivering the same lecture to everyone.

Our goal in this work is to empirically explore the impact of that imperfect knowledge in a classroom scenario with multiple learners and how the interactive approach proposed by Melo et al. [4] can help in overcoming it. Within this problem we formulated two hypotheses:

- Hypothesis 1: Considering interactivity in the teaching process outperforms other non-interactive approaches. This hypothesis is supported by numerous pedagogical studies found in the literature suggesting that interactivity enhances students' engagement and learning [1–3];
- Hypothesis 2: Dividing the class into partitions accordingly to the students priors and giving one sample per partition makes the learning process faster than considering the class as a whole. This hypothesis was inspired in the work by Zhu et al. [7] that investigated machine teaching with multiple learners but no interactivity.

To confirm these hypotheses we conducted an user study with classes of students. We found that interactivity can be the means to close the gap between the student and teacher parameters. Also, partitioning the whole group into smaller groups, although increasing the effort of the teacher in each run, revealed to improve the overall performance when combined with interactivity.

2 User Study

In this section we present an user study created to validate the hypotheses formulated related to our work. We want to test in a real-world scenario if the interactive approach proposed by Melo et al. [4] is still faster when we extend the discussion of interactivity from single-learner to multiple-learners settings. This raises several novel questions regarding the teaching process: How to interact with multiple students? How to deal with the individual differences between students? How much does the feedback from one student inform the teacher about the state of the class? We also explored different ways of considering the class—as a whole, or partitioned.

2.1 Experimental Design

We used the same artificial problem with a Bayesian estimation task proposed by Melo et al. [4], where each participant has to estimate the mean monthly rent of an 1-bedroom apartment in a city in the US. In this case, instead of teaching only one student, we considered groups of 10 students at the same time. In each run each student gives an answer (not shared with the rest of the group). After that, we give an example (said to be real) of an 1-bedroom apartment rented in that city to each student (which is, again, not shared with the others). This is repeated for 10 runs with every group.

To compute the teaching samples to show to the group we explored 3 different approaches:

- Condition 1: Interactive Teaching with No Partitions—Figure 1 (left)—where we teach the whole group at the same time, considering in each iteration one randomly selected answer from the total group as the answer of the class. We then follow the algorithm of Melo et al. [4], where we take into account the considered answer of the class in each iteration and use it to calculate the sample to show. We show the same sample to every participant.

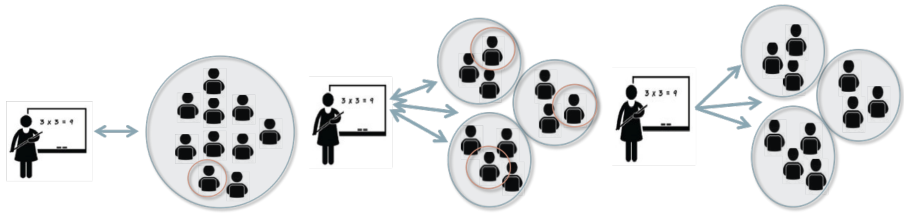

Fig. 1: Diagram with the 3 approaches considered: the interactive teaching approach with no partitions (left); the interactive teaching approach with partitions (middle); the non-interactive teaching approach with partitions (right).

- Condition 2: Interactive Teaching with Partitions—Figure 1 (middle)—instead of considering the whole-group, here we divide the group into smaller partitions. To do so we first ask each participant to select the interval where his estimate falls. We use that information to aggregate the participants into smaller groups (partitions) accordingly to this prior knowledge. In each iteration and for each partition we consider one randomly selected answer as the answer of that partition. If we have n partitions, we will consider n answers. And we will then show n samples, one per partition. Thus, not every participant sees the same sample as in condition 1—participants in different partitions see different samples.
- Condition 3: Non-Interactive Teaching with Partitions—Figure 1 (right)—the group is partitioned as in Condition 2. However, the samples are presented to the students in each partition without taking into account the answers given by the students. Instead, the system assumes perfect knowledge about the student parameters and uses the mean value of the interval of prior estimates accepted in each partition as the estimate of that partition.

2.2 Results

The results confirmed our two hypothesis regarding the use of interactivity in the teaching process and addressing the class divided into partitions instead of as a whole. The study involved a total of 239 engineering students distributed uniformly among the three conditions. The average age was 23, with 71% males.

Partitions vs. Whole group Figure 2 (left) shows that the conditions with partitions seem to have a better performance than the one with no partitions. However, when a teacher decides to partition the group, his teaching effort (the number of samples needed to teach) increases with the number of partitions considered. To take this into account, we multiplied the teaching samples of the conditions with partitions by the average number of partitions in all the groups acquired in each condition (around 4 in both of them)—Figure 2 (left). Including this extra effort, it is not so obvious that the conditions with partitions have a better performance. When comparing the interactive cases with a Mann-Whitney

U test, having partitions is indeed statistically better. However, the interactive condition with no partitions has significantly lower error rates when comparing with the non-interactive with partitions approach.

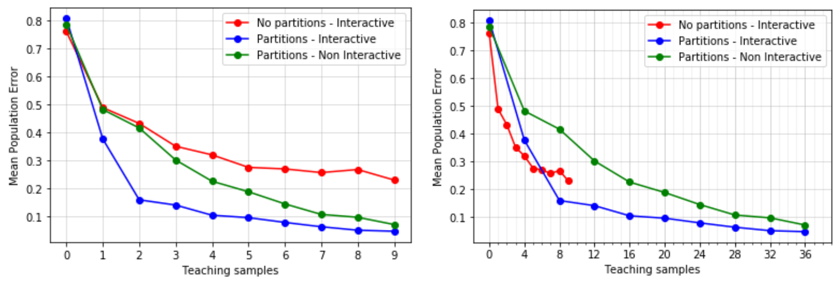

Fig. 2: Comparison of the learning performance when teaching multiple learners not considering the extra cost of teaching more partitions (left) and considering this extra effort (right).

Interactivity vs. Non-interactivity The previously mentioned results show that having partitions is not necessarily better and that one must also consider the factor of using interactivity in the teaching process. Indeed, with or without partitions, the interactive conditions showed significantly better performances when performing Mann-Whitney U tests, even when considering the extra teaching effort associated to the partitions.

3 Conclusions

In this work we empirically investigate the use of interactivity when teaching a Bayesian estimation task to groups of learners. However, we assume there is a mismatch between what the teacher knows about the learners and the learners themselves. We inspected two ways to consider the class: as a whole or partitioned accordingly to the learners priors. We could confirm that, by allowing the teacher to interactively assess the state of the class, the impact of the aforementioned mismatch is significantly mitigated. The results of an user study have shown that the interactive teaching approaches (with partitions or not) significantly outperform the non-interactive alternative. Between the interactive approaches, dividing the class into partitions leads to better learning performances.

4 Acknowledgments

This work was supported by national funds through Fundação para a Ciência e a Tecnologia (FCT) with reference UIDB/50021/2020 and the FCT PhD grant with reference SFRH/BD/118006/2016.

References

1. Alexander, R.: Developing dialogic teaching: Genesis, process, trial. Research Papers in Education **33**(5), 561–598 (2018)
2. Beauchamp, G., Kennewell, S.: Interactivity in the classroom and its impact on learning. Computers & education **54**(3), 759–766 (2010)
3. Kennewell, S., Beauchamp, G.: The features of interactive whiteboards and their influence on learning. Learning, Media and Technology **32**(3), 227–241 (2007)
4. Melo, F.S., Guerra, C., Lopes, M.: Interactive optimal teaching with unknown learners. In: Proceedings of the 27th International Joint Conference on Artificial Intelligence. pp. 2567–2573. Stockholm, Sweden (2018)
5. Zhu, X.: Machine teaching for bayesian learners in the exponential family. In: Advances in Neural Information Processing Systems. pp. 1905–1913 (2013)
6. Zhu, X.: Machine teaching: An inverse problem to machine learning and an approach toward optimal education. In: Twenty-Ninth AAAI Conference on Artificial Intelligence (2015)
7. Zhu, X., Liu, J., Lopes, M.: No learner left behind: On the complexity of teaching multiple learners simultaneously. In: Proceedings of the 26th International Joint Conference on Artificial Intelligence. pp. 3588–3594 (2017)
8. Zhu, X., Singla, A., Zilles, S., Rafferty, A.N.: An overview of machine teaching. arXiv preprint:1801.05927 (2018)