The Leopard Framework: Towards Understanding Educational Technology Interventions with a Pareto Efficiency Perspective

José P. González-Brenes JOSE.GONZALEZ-BRENES@PEARSON.COM
School Research, Pearson, Philadelphia, PA, USA

Yun Huang YUH43@PITT.EDU
Intelligent Systems Program, University of Pittsburgh, Pittsburgh, PA, USA

Abstract

Adaptive systems teach and adapt to humans; their promise is to improve education by minimizing the subset of items presented to students while maximizing student outcomes (Cen et al., 2007). In this context, items are questions, problems, or tasks that can be graded individually. The adaptive tutoring community has tacitly adopted conventions for evaluating tutoring systems (Dhanani et al., 2014) by using classification evaluation metrics that assess the student model component— student models are the subsystems that forecast whether a learner will answer the next item correctly.

Unfortunately, it is not clear how intuitive classification metrics are for practitioners with little machine learning background. Moreover, our experiments on real and synthetic data reveal that it is possible to have student models that are very predictive (as measured by traditional classification metrics), yet provide little to no value to the learner. Additionally, when we compare alternative tutoring systems with classification metrics, we discover that they may favor tutoring systems that require higher student effort with no evidence that students are learning more. That is, when comparing two alternative systems, classification metrics may prefer a suboptimal system.

We recently proposed Learner Effort-Outcomes Paradigm (Leopard) for automatic evaluation of adaptive tutoring (González-Brenes & Huang, 2015). Leopard extends on prior work on alternatives to classification evaluation metrics (Lee & Brunskill, 2012). At its core, Leopard quantifies both the *effort* and *outcomes* of students in adaptive tutoring. Even though these metrics are novel by itself, our contribution is approximating both without a randomized control trial.

In this talk, we will describe our recently published results on meta-evaluating Leopard and conventional classification metrics. Additionally, we will present preliminary results of framing the value of an educational intervention as multi-objective programming. We argue that human-propelled machine learning, and educational technology in particular, aim to optimize the Pareto boundary of effort and outcomes of humans.

Proceedings of the 32^{nd} International Conference on Machine Learning, Lille, France, 2015. JMLR: W&CP volume 37.

References

Cen, Hao, Koedinger, Kenneth R., and Junker, Brian. Is Over Practice Necessary?—Improving Learning Efficiency with the Cognitive Tutor Through Educational Data Mining. In *Proceedings of the 2007 Conference on Artificial Intelligence in Education: Building Technology Rich Learning Contexts That Work*, pp. 511–518, Amsterdam, The Netherlands, 2007. IOS Press. ISBN 978-1-58603-764-2. URL http://dl.acm.org/citation.cfm?id=1563601.1563681.

Dhanani, Asif, Lee, Seung Yeon, Phothilimthana, Phitchaya, and Pardos, Zachary. A comparison of error metrics for learning model parameters in bayesian knowledge tracing. Technical Report UCB/EECS-2014-131, EECS Department, University of California, Berkeley, May 2014. URL http://www.eecs.berkeley.edu/Pubs/TechRpts/2014/EECS-2014-131.html.

González-Brenes, José P. and Huang, Yun. Your model is predictive— but is it useful? theoretical and empirical considerations of a new paradigm for adaptive tutoring evaluation. In Boticario, Jesus G., Santos, Olga C., Romero, Cristöbal, and Pechenizkiy, Mykola (eds.), *Proceedings of the 8th International Conference on Educational Data Mining*, Madrid, Spain, 2015.

Lee, Jung In and Brunskill, Emma. The impact on individualizing student models on necessary practice opportunities. In Yacef, Kalina, Zaïane, Osmar R., Hershkovitz, Arnon, Yudelson, Michael, and Stamper, John C. (eds.), *Proceedings of the 5th International Conference on Educational Data Mining*, pp. 118–125, Chania, Greece, 2012. URL http://educationaldatamining.org/EDM2012/uploads/procs/Full_Papers/edm2012_full_11.pdf.