Avoid playing learner and system
off against each other

Ji-Ung Lee, Christian M. Meyer, Iryna Gurevych
Ubiquitous Knowledge Processing (UKP) Lab and Research Training Group AIPHES
Technische Universität Darmstadt, Germany
https://www.ukp.tu-darmstadt.de

Digital learning platforms for self-directed learning, such as Duolingo\(^1\) gained increasing popularity in recent years. A major challenge for such platforms is providing learning material that fits the needs and proficiency of a particular learner – a task that has been done by human teachers in the traditional classroom setting and that must now be fully automatic in order to scale up the platform and enable learners to learn at their own pace whilst receiving immediate feedback on their inputs. Therefore, there has been ample research on \textit{automated exercise generation} (Mitkov et al., 2006; Chinkina and Meurers, 2017) and \textit{automated difficulty prediction} (Beinborn et al., 2014; Pilán et al., 2016) using machine learning (ML).

However, introducing ML and artificial intelligence in general into the learning process raises two important ethical issues: \textit{i)} Systems may fail to recognize correctly given answers, or even worse, suggest wrong answers, and \textit{ii)} they provide learners with unsuitable (i.e. too easy or too difficult) exercises outside their \textit{Zone of Proximal Development} (Vygotsky, 1978). Both issues may severely harm the learning progress. As Hovy and Spruit (2016) point out, there is yet little work on mitigating such issues in our community.

Due to the scarcity of available training data, researchers increasingly rely on crowdsourcing (Heffernan et al., 2016) or active ML techniques (Zesch et al., 2015) to overcome the so-called cold-start problem of ML. These approaches are especially problematic, since they solely aim at improving the system and its underlying ML model – at the cost of the learning goals of the users. Learners are reduced to cheap labelers suffering from incorrect system feedback and varying task difficulty (cf. Settles et al., 2008). This demotivates learners, reduces their learning speed, and might even yield misconceptions.

In our work, we explore these issues for a language learning use case: \textit{i)} We use \textit{automatically generated C-tests} (Klein-Braley and Raatz, 1982), which have a very small solution space and thus, prevent incorrect system responses. \textit{ii)} We integrate the learner’s goal into the active ML objective of the \textit{automated difficulty prediction}. To this end, we can jointly optimize for the learner’s goal of quickly reaching their next proficiency level and the system’s goal of reliably estimating exercise difficulty at high accuracy. This will contribute to learning platforms that effectively support learners without the necessity of vastly existing training data and without treating the learners as mere data labelers.

\footnote{https://www.duolingo.com}
Acknowledgments. This work has been supported by the Hessian research excellence program “Landes-Offensive zur Entwicklung Wissenschaftlich-ökonomischer Exzellenz” (LOEWE) as part of the a! - automated language instruction project under grant No. 521/17-03 and by the German Research Foundation as part of the Research Training Group “Adaptive Preparation of Information from Heterogeneous Sources” (AIPHES) under grant No. GRK 1994/1.

References


