Explanations at Scale via Machine Learning

For the technology review I took the time to review an Improvement System(AXIS) with the core focus of applying Machine Learning to influence previously used data on explanations for any given problem to show to someone who may be attempting to solve a problem in the future. Explanations can be essential and hammer home points that go beyond what facts alone can do, rather explanations directly identify core concepts that will give a further understanding of new creative problems. One of the biggest examples would be online learning, in that students can use a method of trial and error to reach solutions without gaining a high level understanding of why the answer is what it is. The purpose of this system is to dynamically improve over time as learners' interaction with said material increases.

The core components used for this project consist of an interface and an explanation selection policy. The interface collects data in the form of explanations from learners, and then takes this data and elicits the generation of new explanations from future learners. The explanation selection policy is fairly self-explanatory in that this system will decipher which explanations are most relevant to return to future learners; this policy continues to be dynamically updated based on submitted explanations from learners. This system is multi-layered in that it requires users to rate explanations given to them, and from there a multi-armed bandit algorithm is applied to statistically analyze ratings done by learners and update the explanation selection policy. By doing so the system will continue to include new explanations while dynamically learning which explanations are most relevant to represent to future learners.

The biggest challenge faced when attempting to return relevant explanations to problems to future learners is figuring out which explanations are deemed relevant autonomously rather than manually. This is where the multi-armed bandit algorithm, a system that requires repeated select action to learn which action is most effective, comes into play. After an explanation is given to the user measuring the effectiveness would be considered the observed reward in the bandit formulation. Thompson sampling, which is a Bayesian algorithm can also be used especially when dealing with exploitation versus exploration, and like most bandit algorithms Thompson sampling gives a dynamic policy for choosing which explanation to give a new user by observing the reward just after an explanation is selected. Parameter tuning is done for each distribution but set prior to beliefs about an explanation, then updated based on likelihood of the observed evidence. Once an explanation is chosen and displayed to the user, the user's ratings on that explanation's effectiveness is taken in using successes and failures as the metric on a 10 point scale with 10 being most helpful.

Using this implementation a case study was done for solving math problems, the target being an online instructional designer overseeing online math problems (somewhat similar to Khan Academy). As a method of improving AXIS' validity the implementation and deployment was conducted in two phases: the first phase was how data is collected and policy algorithm changes with time, and the second stage would report results from a randomized experiment. 150 learners were set to participate and generated between 60 and 72 explanations for each of the four problems outlined. As a way of testing the validity of the two phases, a randomized comparison to explanations was taken from a range of sources. It was important to keep in mind how good each explanation is, and how likely it is to be shown in that current policy. The learnersourced explanations AXIS presented were rated as significantly more helpful for learning than the explanations removed by the filtering rule ($M = 6.83$ vs. 6.03, $SE = 0.28$, $p < 0.01$) (AXIS: Generating Explanations at Scale with Learnersourcing and Machine Learning 7).

All in all generating explanations for a large number of online learning materials require a significant amount of time and effort from instructors. With AXIS the Machine Learning model combines techniques to tackle these limitations. The impact of this model can be applied to more than just Math based problems as the study had conducted, and some limitations of the current structure include adding further functionality, giving how-to instructions for students who will be writing explanations that teach procedural skills, add more illustrative and visual based examples, and or clarifying task instructions on online workspaces to improve clarity. Improving the reward based system could also be improved for effectiveness metric, as I don't think rating out of 10 may always be the best. Overall this model was quite interesting and outlined some interesting qualitative analysis on how explanations can be quite helpful for information retention.