

Ask the Expert: NLP

Hi, my name is Victoria Yaneva and I'm Manager of NLP Research at NBME.

On-screen: [Victoria Yaneva. Manager, data scientist, growth and innovation. What are the benefits for medical education of having technologies like NLP to assess learners?]

There are multiple benefits of having technologies such as NLP in medical education to assess learners. I would broadly categorize them in two groups. The first one is scalability.

NLP and AI allow us to do things such as automated scoring as one example at a much bigger scale than we used to be able to when we were relying solely on human experts and human writers. So, what this does is it allows us to scale assessments, for example, video-based assessments in a way that was not possible before because we have an AI system or an NLP system in this case to do the scoring automatically.

The second, and probably less obvious thing that NLP and AI give us in assessment is the ability to extract new types of data and combine it in such a way as to actually be able to arrive at a new type of evidence. And that is not necessarily an extension of human capabilities, but a different way of combining information to arrive at new insights.

When we talk about responsible use of AI, here at NBME we are very cognizant that such capabilities need to be very carefully developed and we live up to that responsibility in every way that we can through doing responsible research.

On-screen: [What is your research and NBME doing to advance assessment development using NLP?]

At NBME we do a lot of research to advance assessment development using NLP and AI. A couple of examples come to mind.

We do focus a lot of our attention on automated scoring. The reason for this is that in order to develop measurements for new skills and abilities such as communications skills or clinical reasoning, frequently we would need free text responses from examinees that need to be scored in some valid and reliable way.

On-screen: [NLP allows for scalability to score free-text responses in a valid and reliable way.]

And here is a great example of how NLP allows for the scalability of such assessments where we are able to accurately score such free text responses and also provide an evidence base for the score by, for example, highlighting where exactly in the text we have found in our correspondence to a concept, of importance. So, this is to the point of scalability.

And I again want to emphasize here that a lot of research goes into the validity and the reliability and fairness of such scoring systems. So as an assessment organization, this is of paramount importance to us.

On-screen: [What excites you about the future of NLP in Medical Assessment?]

I think everyone who has witnessed the release of ChatGPT and the conversations that followed in our community and in many other communities, both within assessment, education and pretty much any industry, we have arrived at a time where we are very excited to see what NLP and AI have to offer. There is a great deal of enthusiasm and belief in these systems that we were not seeing before because their impact and capability was not quite as evident.

I am curious to see not only how our research fits into this larger narrative that other medical education, assessment and NLP communities create, but also to see how our voice around issues such as fairness, validity and reliability of such systems can be amplified. Because while we want to harness the potential of such technology, we want to do so in a responsible way.

On-screen: [Follow us on our journey. reassesssthefuture.org. NBME. Copyright © 2023 NBME. All rights reserved.]