



تصویرسازی: صدیقه هاشمی

سنجش الکترونیکی

همتاسنجی و خودارزیابی در عصر دیجیتال

((فرخ فیضی / دانشجوی دکتری تکنولوژی آموزشی))

اشاره

در چند سال گذشته دوره‌های گسترده آزاد آنلاین (MOOC) به شکل فزاینده‌ای محبوب شده‌اند. «گسترده» عنوان شدن این دوره‌ها در ارتباط با تعداد یادگیرندگان ثبت‌نام‌شده‌شان است. محتوای این دوره‌ها معمولاً واحدهایی صوتی، تصویری و متنی هستند که از طریق اینترنت ارسال می‌شوند و رایگان‌اند. ثبت‌نام در این دوره‌ها از هزاران تا صدها هزار نفر متغیر است. البته در سال ۲۰۱۴، با گسترش موبک‌ها انتظار می‌رفت ثبت‌نام یادگیرندگان تا ۱۰۰ درصد افزایش یابد. این پیش‌بینی دو سال پس از آن صورت گرفت که نیویورک تایمز سال ۲۰۱۲ را به‌عنوان سال موبک‌ها در حال رشد است، در سنجش و ارزشیابی از آنجا که تعداد موبک‌ها در حال رشد است، در سنجش و ارزشیابی در موبک، به‌عنوان بخشی مهم از این سیستم، تمرکز فزاینده‌ای وجود دارد. به دلیل ثبت‌نام‌های زیاد در هر موبک، برای مربی غیرممکن است بتواند تکالیف را ارزیابی و فعالیت‌های هر دانش‌آموز را ارزشیابی کند. در نتیجه برای ردیابی و پیش‌بینی رفتار یادگیرنده به هنگام تکمیل ارزشیابی موبک، سیستم‌های هوشمندی طراحی شده‌اند. این ارزشیابی‌های خودکار شرایط لازم برای نمره دهی به عملکرد دانش‌آموز و ارائه بازخورد لازم به آن‌ها را فراهم آورده است. این شیوه از ارزشیابی به معلمان کمک می‌کند دانش‌آموزان را رتبه‌بندی و در فرایند یادگیری از آن‌ها حمایت کنند.

کلیدواژه‌ها: سنجش الکترونیکی، همتاسنجی، موبک

ارزشیابی در موبک

ارزشیابی موضوع مهمی است که باید بررسی شود، زیرا می‌تواند در یادگیری ابزار قدرتمندی باشد؛ حتی در مورد موبک. ارزشیابی‌ها می‌توانند نتایج مثبت یا منفی داشته باشند، چراکه بر یادگیری و انگیزه یادگیرندگان تأثیر می‌گذارند. آن‌ها می‌توانند یادگیری را، هم بسنجند و هم از آن پشتیبانی کنند. از آنجا که هزاران موبک وجود دارد، ارزشیابی هر دوره به‌صورت جداگانه امکان‌پذیر نیست. بنابراین، سیستم‌عامل‌های ارائه‌دهنده این دوره‌ها، ارزشیابی را با تست‌هایی چندگزینه‌ای یا فعالیت‌هایی محاسباتی ارزشیابی می‌کنند که امکان طبقه‌بندی و تصحیح خودکار آن‌ها توسط رایانه وجود داشته باشد.

برخی از موبک‌ها از سؤالات باز پاسخ و انشایی به‌عنوان روش ارزشیابی مطلوب استفاده می‌کنند، چون شیوه چندگزینه‌ای نمی‌تواند مهارت‌های سطوح بالای یادگیری را به اندازه سؤالات باز پاسخ ارزشیابی کند. اما

شود، اما هر شیوه از ارزیابی محدودیت‌ها و مشکلات خاص خود را دارد. تا به امروز کاربردی‌ترین روش برای همهٔ موک‌ها، استفاده از همتاسنجی برای ارائهٔ بازخورد لازم است. هر چند که این شیوهٔ ارزیابی نیز از کمبود اعتبار رنج می‌برد.

در آموزش سنتی کلاسی، از همتاسنجی برای تسهیل مباحث کلاس استفاده می‌شود، اما در موک همتاسنجی در یک محیط بسیار متفاوت صورت می‌گیرد. اولین و بدیهی‌ترین موضوع در این مقایسه، مقیاس است. برای یک تکلیف واحد در یک موک، ده‌ها تا صدها هزار همکار بالقوه وجود دارد که بیش از صدها هزار مورد ارسال را ارزیابی می‌کنند. تفاوت دوم در این است که به دلیل مقیاس، وساطت، نظارت و راهنمایی مربی بسیار اندک است یا اصلاً وجود ندارد. تفاوت سوم در این است که در موک‌ها، شرکت‌کنندگان بین‌المللی هستند و در زبان، فرهنگ، جهان‌بینی و ارزش‌های آن‌ها تفاوت‌های زیادی وجود دارد. ضمن اینکه بدون نظارت مربی، یادگیرندگان نیز احساس مسئولیت و انگیزهٔ کمی برای جدی گرفتن روند ارزیابی همتایان خود دارند. فرا تحلیل لی و همکاران (۲۰۱۶)، با بررسی مقالات ارزیابی همتایی که در ۱۵ سال گذشته منتشر شده‌اند، نشان داد همتاسنجی وقتی بر کاغذ مبتنی است، با ارزیابی معلم همبستگی بیشتری دارد تا زمانی که مبتنی بر رایانه است. همچنین، وقتی ارزیابی‌ها داوطلبانه باشند، همبستگی بالاتری با ارزیابی معلم خواهند داشت. نکتهٔ آخر اینکه وقتی یادگیرندگان در ایجاد معیارهای رتبه‌بندی شرکت دارند، ارزیابی همتایان دقیق‌تر است.

منابع

1. Admiraal, W., Huisman, B., & Pilli, O. (2015). Assessment in Massive Open Online Courses. *Electronic Journal of E-learning*, 13(4), 207-216.
2. Balfour, S. P. (2013). Assessing Writing in MOOCs: Automated Essay Scoring and Calibrated Peer Review™. *Research & Practice in Assessment*, 8, 40-48.
3. Chauhan, A. (2014). Massive open online courses (MOOCs): Emerging trends in assessment and accreditation. *Digital Education Review*, (25), 7-17.
4. Costello, E., Holland, J., & Kirwan, C. (2018). The future of online testing and assessment: question quality in MOOCs. *International Journal of Educational Technology in Higher Education*, 15(1), 42.
5. Li, H., Xiong, Y., Zang, X., L. Kornhaber, M., Lyu, Y., Chung, K. S., & K. Suen, H. (2016). Peer assessment in the digital age: a meta-analysis comparing peer and teacher ratings. *Assessment & Evaluation in Higher Education*, 41(2), 245-264.
6. Suen, H. K. (2014). Peer assessment for massive open online courses (MOOCs). *The International Review of Research in Open and Distributed Learning*, 15(3).
7. Timmis, S., Broadfoot, P., Sutherland, R., & Oldfield, A. (2016). Rethinking assessment in a digital age: Opportunities, challenges and risks. *British Educational Research Journal*, 42(3), 454-476.
8. Ventista, O. M. (2018). Self-assessment in Massive Open Online Courses. *E-Learning and Digital Media*, 15(4), 165-175.

تصحیح و نمره‌گذاری سوالات باز پاسخ نیز با ماشین به‌دقت انسان امکان‌پذیر نیست. از سوی دیگر، بدون شک مشارکت مدرسان در ارزیابی هزاران تکلیف باز پاسخ غیرممکن است. به همین دلیل، موک‌ها به دنبال شیوه‌های جدیدی از ارزیابی بودند. برای مثال ای دی ایکس از سیستمی به نام «Automated Essay Scoring» استفاده می‌کند که با وجود عملکرد مثبت در شناسایی و گزارش خطاهای دستوری، خطاهای استفادهٔ کلمات، تنوع جملات، سبک، پیچیدگی متن، واژگان، تراز کردن محتوا با متن، ایده‌های حمایتی، نتیجه‌گیری و بخش‌های بی‌ربط، به دلیل محدود کردن رتبه‌بندی بر اساس عوامل سطحی، و نداشتن توانایی امتیازدهی به خلاقیت، طنز یا دریافت استعاره‌ها، مورد انتقاد قرار گرفته است و قادر به ارزیابی متون به روش ارزیابی‌کننده‌های انسانی نیست. این موضوع، این شبهه را ایجاد کرد که سیستم رایانه‌ای قادر به ارزشیابی فعالیت‌های دانش‌آموزان نیست و به ارزیابی‌های انسانی نیاز است.

از آنجا که ارزیابی تکلیف هزاران یادگیرنده توسط مربیان غیرممکن است، دو موک بزرگ کور سرا و ای دی ایکس (۲۰۱۶) از ارزیابی همتا و خودارزیابی استفاده می‌کنند. به‌طور خاص، کور سرا از یک نظام ارزیابی «همتای واسنجی شده» استفاده می‌کند. در این نظام یادگیرندگان ابتدا برخی تکلیف را که قبلاً مربی تصحیح کرده است، نمره‌گذاری می‌کنند. سپس سیستم میزان توافق بین نمرهٔ مربی و یادگیرنده را بررسی و نمرات داده‌شده توسط یادگیرندگان را واسنجی می‌کند و با توجه به میزان توافق، به یادگیرندگان شاخص صلاحیت داوری داده می‌شود. سپس هنگامی که یادگیرندگان تکلیف همتایان خود را ارزیابی می‌کنند، این درجه را می‌توان مطابق با شاخص صلاحیت داوری، کم‌وبیش به‌عنوان نمرهٔ نهایی آنان در نظر گرفت.

در شیوهٔ ارزیابی همتایان، یادگیرندگان تکلیف چهار یا پنج نفر از همتایان خود را با استفاده از یک روال ارزیابی می‌کنند و بقیه نیز به همان ترتیب تکلیف اینان را ارزیابی می‌کنند. بعضی از دوره‌ها نیز شرایط خودارزیابی را ارائه می‌کنند.

اما در ای دی ایکس (۲۰۱۶) برای تکلیف باز پاسخ دانش‌آموزان، وقتی که توسط همتایان خودارزیابی و ارزشیابی می‌شوند، فقط نمرهٔ همتایان به آن‌ها داده می‌شود و خودارزیابی در نظر گرفته نمی‌شود. این بدان معناست که در این نظام، خودارزیابی تا حدی دست‌کم گرفته شده است. ارزیابی در کورسرا و ای دی ایکس می‌تواند پایانی یا تکوینی باشد. ارزیابی برای یادگیری یا تکوینی قصد دارد از یادگیری پشتیبانی کند، درحالی‌که ارزیابی پایانی با درجه‌بندی، گواهی‌نامه و مدرک همراه است. توضیح آنکه یادگیرندگان برای به دست آوردن گواهی‌نامه در هر دو نظام، ملزم به پرداخت هزینه هستند.

نتیجه‌گیری

از آنجا که چرخهٔ آموزش-یادگیری-ارزیابی در موک شکسته شده است، بدون ارزیابی سازنده و بازخورد، موک به نمایشگر مقداری اطلاعات نه‌چندان ارزشمند تبدیل خواهد شد تا تجربه‌های آموزشی مفید! اگرچه در راه‌حل‌های ارائه‌شده سعی شده ارزیابی‌های سازنده به موک برگردانده