

## نامه به سردبیر

# هوش مصنوعی و آزمون بالینی ساختارمند عینی: استفاده از ChatGPT برای ارزیابی مهارت‌های

## بالینی

حمید مهدوی فرد<sup>۱\*</sup>

از پزشکی از راه دور در طول دهه گذشته و به طور خاص با همه‌گیری COVID-۱۹، مربیان تمرکز خود را بر آموزش و ارزیابی مجازی از طریق آسکی افزایش داده‌اند (۴).

برگزاری آزمون آسکی نیاز به تدارکات قابل توجهی برای ایجاد ارزیابی‌های با کیفیت بالا، موفق و معتبر از دانشجویان و دستیاران دارد. دستیابی به استاندارد در آزمون آسکی با توجه به تنوع در بیماران استاندارد و روحیات دانشجو، دانش یا تجربه قبلی دانشجو، پیچیدگی موارد و سطح آمادگی بیمار استاندارد دشوار است. با این حال، استفاده از موارد متعدد می‌تواند مطلوبیت و اعتبار آزمون آسکی را بهبود بخشد (۱). بنابراین، مربیان اغلب برای بهبود ارزیابی تکوینی و تراکمی دانشجویان خود، به دنبال آسکی‌های معتبر قبلی هستند. اگرچه کارآموزان در طول یک دوره یا برنامه برای عملکرد خوب در آزمون آسکی آماده می‌شوند، بسیاری از کارآموزان به دلیل ماهیت پراهمیت امتحانات آسکی احساس استرس و اضطراب می‌کنند (۲). با توجه به طیف گسترده‌ای از تظاهرات بالینی که ممکن است در طول امتحان ارزیابی شوند، دانشجویان اغلب در مورد آمادگی امتحان ابراز نگرانی می‌کنند. پلتفرم‌هایی که امکان تمرین زنده

آزمون بالینی ساختارمند عینی<sup>۲</sup> برای ارزیابی مهارت‌های استاندارد شده و مشخص است که در سراسر جهان و برای تمامی تخصص‌های پزشکی توسط دانشکده‌های پزشکی برای کمک به برای بهبود توانمندی و مهارت‌های بالینی دانشجویان پزشکی و دستیاران استفاده می‌شود (۱). این آزمون در کمک به کارآموزان برای تبدیل شدن به پزشکان موفق ضروری است (۲). آزمون‌های آسکی که در ابتدا در اواسط دهه ۱۹۷۰ ابداع شد، در ارائه بازخورد تکوینی و نمره‌دهی تراکمی در زمان واقعی نقش اساسی دارند و ممکن است به صورت حضوری، مجازی و پلتفرم‌های ترکیبی برگزار شود (۳). کارآموزان در آزمون‌های آسکی مواجهه‌های با دنیای واقعی را در یک محیط کنترل‌شده تقلید می‌کنند، بنابراین مربیان می‌توانند شکاف‌های عملکردی کارآموزان را شناسایی کنند و بازخورد مؤثر و ویژه‌ای ارائه دهند. آسکی‌ها می‌توانند پویا/قابل مشاهده باشند که در آن یک بیمار استاندارد شده<sup>۳</sup> با کارآموزی که خود را به عنوان یک پزشک یا مراقب بیمار معرفی می‌کند تعامل داشته باشد، یا می‌تواند ایستا/شبه-آسکی باشد که در آن دانش کارآموز در ایستگاه‌هایی با سوالات کتبی یا آنلاین ارزیابی می‌شود (۲). با افزایش استفاده

۱- مرکز مطالعات و توسعه آموزش علوم پزشکی، دانشگاه علوم پزشکی نیشابور، نیشابور، ایران.

\* (نویسنده مسئول): تلفن: +۹۸۵۱۴۳۳۰۶۲۵۵، پست الکترونیکی: [mahdavifh1@mums.ac.ir](mailto:mahdavifh1@mums.ac.ir)

تاریخ پذیرش: ۱۴۰۴ / ۰۳ / ۱۰

تاریخ بازبینی: ۱۴۰۴ / ۰۲ / ۳۱

تاریخ دریافت: ۱۴۰۴ / ۰۲ / ۰۲

<sup>2</sup> Objective Structured Clinical Examinations (OSCE)

<sup>3</sup> standardized patient (SP)

تقویت چشمگیری برای توانایی مربیان برای ایجاد محتوا و ارزیابی سریع آسکی فراهم کند (۲). همچنین باید توجه داشت که آسکی معمولاً برای مدیریت، نیاز به بودجه و نیروی اجرایی دارد و استفاده از ChatGPT ممکن است با کاهش زمان و پرسنل مورد نیاز برای برگزاری امتحانات به کاهش هزینه‌های دانشکده کمک کند. مربیان باید با این فناوری جدید آشنا شوند، به خصوص در عصر حاضر که استفاده از هوش مصنوعی پذیرش گسترده‌ای پیدا کرده است (۵).

اولین گام در توسعه آسکی مشخص کردن شکاف‌ها در ارزیابی مهارت‌های بالینی دانشجویان است که با اهداف برنامه و روش‌های ارزیابی موجود هماهنگ است. در حالی که ChatGPT، می‌تواند یک برنامه را برای شناسایی مهارت‌های بالینی ارائه نموده و آن را تجزیه و تحلیل کند، فقدان دوره‌های آموزش پزشکی استاندارد شده و چالش اجرای همه اقدامات در ChatGPT اثربخشی آن را محدود می‌کند. علاوه بر این، آپلود داده‌های عملکرد دانشجو برای شناسایی شکاف‌های دانش در مهارت‌های بالینی به دلیل قوانین حقوقی آموزشی و حریم خصوصی و مسائل مربوط به قالب‌بندی داده‌ها با موانعی مواجه است. این محدودیت‌ها، کاربرد محدود ChatGPT را در مراحل اولیه آماده سازی آسکی نشان می‌دهد (۲).

پس از شناسایی نیاز برای ارزیابی مهارت‌های بالینی، یک مربی پزشکی یک آزمون را طراحی می‌کند که شامل ارائه، تاریخچه، یافته‌های معاینه فیزیکی، آزمایشگاه/تصویربرداری، استدلال بالینی، تشخیص و ارزیابی است. ChatGPT به طور قابل توجهی به این فرآیند کمک می‌کند (۲، ۴). هنگامی که مشخصاتی مانند تشخیص و سن بیمار درخواست شود، به سرعت موارد خواسته شده را مشخص می‌کند و از طریق گفت‌وگوی تعاملی، موارد مطابق با اهداف آموزشی اصلاح می‌شوند. ظرفیت ChatGPT برای ایجاد فوری موارد متنوع از طریق اعلان‌های ساده بسیار گسترده است، اما مربیان باید از دقت بالینی و تناسب سطح کارآموز اطمینان حاصل کنند (۴).

آسکی را فراهم می‌کنند می‌توانند آمادگی برای آزمون را بهبود بخشند و اضطراب امتحان را کاهش دهند (۲).

ChatGPT یک مدل زبانی بزرگ<sup>۴</sup> مبتنی بر هوش مصنوعی<sup>۵</sup> (AI) است که پاسخ‌هایی را ارائه می‌کند که مشابه پاسخ‌های انسانی است. پاسخ‌ها، عمدتاً در قالب متن است و تنها محدود به بخش مراقبت‌های بهداشتی نیستند. قابلیت‌های GPT-۵ که جدیدترین مدل ChatGPT است، شامل مجموعه‌ای از وظایف مانند ترجمه فوری زبان و تحلیل لحن احساسی است. ChatGPT بر روی یک چارچوب گسترده از ۱۷۵ میلیارد پارامتر ساخته شده است، تعدادی که همچنان در حال افزایش است (۵)، و آن را قادر می‌سازد تا طیف وسیعی از موضوعات را پوشش دهد و با توجه به ورودی‌های کاربر در گفتگوهای دقیق شرکت کند. برای مدل زبانی مانند ChatGPT، توانایی تولید متن محاوره‌ای اساسی است (۲). مدل‌های زبانی با تجزیه و تحلیل الگوها در مجموعه عظیمی از داده‌های متنی، با استفاده از معماری‌های یادگیری عمیق که ساختار پیچیده مغز انسان را تقلید می‌کنند، به این امر دست می‌یابند. از دسامبر ۲۰۲۲، ChatGPT توسط مربیان و کارآموزان پزشکی در سراسر جهان استفاده شده است. جالب اینکه عملکرد ChatGPT در یک آسکی واقعی در اوایل سال ۲۰۲۳ مورد آزمون قرار گرفت (۶).

آسکی برای آماده‌سازی به زمان قابل توجهی نیاز دارد تا چالش‌های ارزیابی دانشجو در یک دوره/برنامه شناسایی شود، سناریوهای موردی مناسبی برای پر کردن این چالش‌ها ایجاد کند، تمام جزئیات مورد نیاز را طراحی کند، ایستگاه‌های ثابت ایجاد کند، راهنماهای بیمار استاندارد شده را تدوین نماید، ابزارهای ارزیابی را برای ارزیابی دقیق عملکرد دانشجویان ایجاد کند، و بازخورد معنی‌داری برای کارآموزان ارائه دهد. آموزش پزشکی به سرعت فناوری‌های در حال تحول را مورد استفاده قرار می‌دهد و به کارگیری این فناوری‌ها برای بهبود کارایی آسکی به نفع مربیان پزشکی است. هوش مصنوعی می‌تواند

<sup>4</sup> large language model (LLM)

<sup>5</sup> Artificial Intelligence (AI)

می‌توانند ضبط‌های ویدیویی از مراحل جراحی شبیه‌سازی شده را تجزیه و تحلیل کنند، حرکات ابزار، مراحل جراحی و حتی خطاها را شناسایی کنند (۹، ۱۰). الگوریتم‌های هوش مصنوعی می‌توانند معیارهایی مانند لرزش، طول مسیر، نیروی اعمال شده و کارایی حرکات را کمی کنند و بازخورد دقیقی ارائه دهند که مشاهده سنتی انسان ممکن است از دست بدهد (۹).

در آموزش پزشکی، استفاده از هوش مصنوعی، به‌ویژه ChatGPT، به دلیل عدم دقت احتمالی در محتوای تولید شده توسط هوش مصنوعی نیازمند احتیاط است. عملکرد ChatGPT در آموزش پزشکی در آزمون‌های مقطع پزشکی عمومی و تخصصی پزشکی مستند شده است، و اگرچه چشمگیر است، اما هنوز کامل نیست (۶). کارآموزان نباید برای کسب دانش و آمادگی برای امتحان صرفاً به دقت پزشکی ChatGPT اعتماد کنند. با این حال، هوش مصنوعی می‌تواند ابزار ارزشمندی برای تقویت آموزش پزشکی باشد زمانی که توسط مربیان آموزش‌دیده و متخصص پزشکی هدایت شود. استفاده از هوش مصنوعی در آماده سازی آسکی، اگرچه امیدوارکننده است، اما در مراحل اولیه توسعه باقی مانده است (۲).

فراتر از ایجاد مورد بالینی کامل، ChatGPT به تدوین استراتژی سوالات ایستگاه ایستا همسو با اهداف دوره کمک می‌کند. به عنوان مثال، ChatGPT می‌تواند سوالاتی را برای ارزیابی دانش دستیاران از درمان نارسایی حاد قلبی یا درک یک دانشجوی پزشکی از مدیریت وضعیت آسم ایجاد کند (۲).

یکی از مهم‌ترین کاربردهای هوش مصنوعی در آزمون‌های OSCE، خودکارسازی نمره‌دهی و بازخورد است. به‌طور سنتی، ممتحنین به‌طور دقیق عملکرد دانشجویان را بر اساس یک چک‌لیست نمره می‌دهند که اغلب منجر به خستگی و ناهماهنگی ارزیاب می‌شود (۷). توسعه بیماران مجازی مبتنی بر هوش مصنوعی یکی دیگر از کاربردهای هوش مصنوعی در آسکی است. این عوامل هوشمند می‌توانند مواجهه واقعی با بیمار را شبیه‌سازی کنند و به دانشجویان این امکان را می‌دهند که گرفتن شرح حال، معاینه فیزیکی و مهارت‌های ارتباطی را در یک محیط امن و کم‌خطر تمرین کنند (۸). آموزش جراحی به شدت به شبیه‌سازی متکی است تا مهارت فنی و مهارت‌های تصمیم‌گیری را بدون به خطر انداختن ایمنی بیمار توسعه دهد. هوش مصنوعی با ارائه ارزیابی عینی و فوری از عملکرد جراحی، این حوزه را متحول می‌کند. تکنیک‌های بینایی کامپیوتر

## References

1. Onwudiegwu U. *OSCE: DESIGN, DEVELOPMENT AND DEPLOYMENT*. JWACS. 2018; 8(1): 1–22.
2. Misra SM, Suresh S. *Artificial intelligence and objective structured clinical examinations: Using ChatGPT to revolutionize clinical skills assessment in medical education*. J Med Educ. 2024; 11: 23821205241263475.
3. Rm H. *Assessment of clinical competence using objective structured examination*. BMJ. 1975; 22: 447–51.
4. Ali M. *What Now and What Next? The New Era of OSCE: Innovation in Learning Assessment*. Pharm Educ. 2020; 20(2): 56–8.
5. Waisberg E, Ong J, Masalkhi M, Kamran SA, Zaman N, Sarker P, et al. *GPT-4: a new era of artificial intelligence in medicine*. Ir J Med Sci. 2023; 192(6): 3197–200.
6. Giannos P. *Evaluating the limits of AI in medical specialisation: ChatGPT's performance on the UK Neurology Specialty Certificate Examination*. BMJ neurology open. 2023; 5(1): e000451.
7. Luordo D, Torres Arrese M, Tristán Calvo C, Shani Shani KD, Rodríguez Cruz LM, et al. *Application of Artificial Intelligence as an Aid for the Correction of the Objective Structured Clinical Examination (OSCE)*. Appl Sci. 2025; 15(3).
8. Lavigne E, Lopez A, Frandon J, Blaizot G, Gabellier L, Adham S, et al. *AI-Standardized Clinical Examination training on OSCE performance*. NEJM AI. 2025; 2(8): AIoa2500066.
9. Power D, Burke C, Madden MG, Ullah I. *Automated assessment of simulated laparoscopic surgical skill performance using deep learning*. Sci Rep. 2025; 15(1): 13591.
10. Khan RA, Jawaid M, Khan AR, Sajjad M. *ChatGPT-Reshaping medical education and clinical management*. PaK J Med Sci. 2023; 39(2): 605.