



Islamic Maaref University

Scientific Journal

PAZHUHESH NAME-E AKHLAQ

Vol. 18, Autumn 2025, No. 69

Capacity Assessment and Ethical Challenges of Applying Artificial Intelligence in the Process of Students' Moral Education

Mehdi Tavakkoli¹ \ Jafar Golmohammadi²

1. PhD Candidate, Islamic Ethics Education, University of Tehran, Tehran, Iran (Corresponding Author).
tavvakolimahdi@ut.ac.ir
2. Assistant Professor, Department of Theoretical Foundations and Islamic Ethics,
University of Tehran, Tehran, Iran.
gmohamad@ut.ac.ir

Abstract Info	Abstract
Article Type: Research Article	This study provides a critical analysis of the dialectical nature of artificial intelligence in moral education and conceptualizes the fundamental conflict between the “instrumental logic of technology” and the “ethical logic of education.” Employing a qualitative approach based on conceptual analysis, the research systematically examines the opportunities and challenges of this field and seeks to answer the key question of how this tension can be addressed within an Islamic–Iranian value framework. The findings indicate that while artificial intelligence creates unprecedented capacities for the personalization of moral learning experiences, it simultaneously introduces inherent threats such as algorithmic bias, the erosion of human relationships, and the weakening of moral agency. To address these challenges, a capacity-building and responsibility-oriented framework is proposed, grounded in principles such as human sovereignty over technology, the ultimate responsibility of educators, and the integration of human supervision with algorithmic analysis. The study ultimately emphasizes the necessity of formulating a national policy framework based on the triad of “human–ethics–technology”, so that artificial intelligence may become an opportunity for the moral advancement of future generations.
Received: 2025/05/26	
Accepted: 2025/09/24	
Keywords	Capacity Assessment; Moral Education; AI Ethics; Ethical Challenges; Students; Schools; Education System.
Cite this article:	Tavakkoli, Mehdi & Jafar Golmohammadi (2025). Capacity Assessment and Ethical Challenges of Applying Artificial Intelligence in the Process of Students' Moral Education. <i>Pazhuhesh Name-E Akhlaq</i> . 18 (3). 9-24. DOI:
DOI:	
Publisher:	Islamic Maaref University, Qom, Iran.

تقدير الإمكانيات والتحديات الأخلاقية لتوظيف الذكاء الاصطناعي في عملية التربية الأخلاقية للطلبة

مهدي توکلي^١ / جعفر گل محمدی^٢

١. طالب دكتوراه، قسم تدريس الأخلاق الإسلامية، جامعة طهران، طهران، إيران (الكاتب المسؤول).

tavvakolimahdi@ut.ac.ir

٢. أستاذ مساعد، قسم الأسس النظرية والأخلاق الإسلامية، جامعة طهران، طهران، إيران.

gmohamad@ut.ac.ir

معلومات المادة	ملخص البحث
نوع المقال: بحث	يتناول هذا البحث تحليلاً نقدياً للطبيعة الجدلية للذكاء الاصطناعي في مجال التربية الأخلاقية، ويصوغ التعارض البنوي بين «المنطق الأداتي للتكنولوجيا» و«المنطق الأخلاقي للتربية». وقد أنجز هذا البحث الكيفي بمنهج التحليل المفهومي وبصورة منهجية، حيث يعمل - إلى جانب تقويم الفرص والتحديات في هذا المجال - على الإجابة عن السؤال المحوري: كيف يمكن تجاوز هذا التعارض في إطار المنظومة القيمية الإسلامية - الإيرانية؟ تشير النتائج إلى أن الذكاء الاصطناعي، على الرغم من إتاحتها إمكانيات غير مسبوقة لتخصيص التجربة الأخلاقية، يرافقه عدد من التهديدات الجوهرية، مثل الانحياز الخوارزمي، وتآكل العلاقات الإنسانية، وإضعاف الفاعلية الأخلاقية. ولتجاوز هذه التحديات يقترح البحث إطاراً «قائماً على بناء القدرات والمسؤولية»، يركز على مبادئ منها سيادة الإنسان على التكنولوجيا، وتحمل المربي المسؤولية النهائية، والجمع بين الرقابة البشرية والتحليل الخوارزمي. ويؤكد الاستنتاج النهائي ضرورة إعداد سياسة وطنية قائمة على ثلاثية «الإنسان - الأخلاق - التكنولوجيا»، بما يحول الذكاء الاصطناعي إلى فرصة لتعزيز الارتقاء الأخلاقي لدى الأجيال القادمة.
تاريخ الاستلام: ١٤٤٦/١١/٢٨	
تاريخ القبول: ١٤٤٦/٠٤/٠١	
الألفاظ المفتاحية	تقدير الإمكانيات، التربية الأخلاقية، أخلاقيات الذكاء الاصطناعي، التحديات الأخلاقية، الطلبة، المدارس، التعليم والتربية.
الاقتباس:	توکلي، مهدي و جعفر گل محمدی (١٤٤٧). تقدير الإمكانيات والتحديات الأخلاقية لتوظيف الذكاء الاصطناعي في عملية التربية الأخلاقية للطلبة. مجلة علمية النشرة الاخلاقية. ١٨ (٣). ٢٤ - ٩. DOI: ??
رمز DOI:	??
الناشر:	جامعة المعارف الإسلامية، قم، إيران.



پژوهش‌نامه اخلاق

نشریه علمی

سال ۱۸، پاییز ۱۴۰۴، شماره ۶۹

ظرفیت‌سنجی و چالش‌های اخلاقی کاربست هوش مصنوعی در فرایند تربیت اخلاقی دانش‌آموزان

مهدی توکلی^۱ / جعفر گل‌محمدی^۲

۱. دانشجوی دکتری، مدرسی اخلاق اسلامی، دانشگاه تهران، تهران، ایران (نویسنده مسئول).

tavvakolimahdi@ut.ac.ir

۲. استادیار، گروه مبانی نظری و اخلاق اسلامی، دانشگاه تهران، تهران، ایران.

gmohamad@ut.ac.ir

اطلاعات مقاله	چکیده
نوع مقاله: پژوهشی (۲۴ - ۹)	این پژوهش به تحلیل انتقادی ماهیت دیالکتیکی هوش مصنوعی در تربیت اخلاقی می‌پردازد و تعارض بنیادین میان «منطق ابزاری فناوری» و «منطق اخلاقی تربیت» را صورت‌بندی می‌کند. این تحقیق کیفی با روش تحلیل مفهومی و بصورت نظام‌مند، ضمن ظرفیت‌سنجی فرصت‌ها و چالش‌های این حوزه، به این پرسش کلیدی پاسخ می‌دهد که چگونه می‌توان از این تعارض در چهارچوب ارزشی اسلامی - ایرانی برون‌رفت. یافته‌ها نشان می‌دهد که هوش مصنوعی در عین ایجاد ظرفیت‌های بی‌بدیل برای شخصی‌سازی تجربه اخلاقی، با تهدیدهای ماهوی نظیر سوگیری الگوریتمی، فرسایش ارتباط انسانی و تضعیف عاملیت اخلاقی همراه است. برای برون‌رفت از چالش‌های مذکور، چهارچوبی «ظرفیت‌ساز و مسئولیت‌محور» که بر اصولی چون حاکمیت انسان بر فناوری، مسئولیت‌پذیری نهایی مربی و ترکیب نظارت انسانی با تحلیل الگوریتمی پیشنهاد می‌شود. نتیجه نهایی بر ضرورت تدوین یک سیاست‌نامه ملی مبتنی بر سه‌گانه «انسان - اخلاق - فناوری» تأکید می‌ورزد تا هوش مصنوعی به فرصتی برای تعالی اخلاقی نسل آینده بدل گردد.
تاریخ دریافت: ۱۴۰۴/۰۳/۰۵	
تاریخ پذیرش: ۱۴۰۴/۰۷/۰۲	
واژگان کلیدی	ظرفیت‌سنجی، تربیت اخلاقی، اخلاق هوش مصنوعی، چالش اخلاقی، دانش‌آموزان، مدارس، آموزش و پرورش.
استاد:	توکلی، مهدی و جعفر گل‌محمدی (۱۴۰۴). ظرفیت‌سنجی و چالش‌های اخلاقی کاربست هوش مصنوعی در فرایند تربیت اخلاقی دانش‌آموزان. پژوهش‌نامه اخلاق. ۱۸ (۳). ۲۴ - ۹. DOI: ??
کد DOI:	??
ناشر:	دانشگاه معارف اسلامی، قم، ایران.

طرح مسئله

نظام‌های آموزشی در سراسر جهان در میانه یک گذار پارادایمی قرار گرفته‌اند که در آن هوش مصنوعی از جایگاه یک «ابزار» کمک‌آموزشی به یک «کنشگر» فعال و تأثیرگذار بر فرایندهای شناختی، رفتاری و ارزشی دانش‌آموزان تغییر ماهیت داده است. این فناوری، دیگر صرفاً محتوا را ارائه نمی‌دهد، بلکه با تحلیل داده‌های عظیم رفتاری، شبیه‌سازی موقعیت‌های پیچیده و ارائه بازخوردهای آنی، در حال ایفای نقشی بی‌بدیل در شکل‌دهی به شخصیت و قضاوت اخلاقی نسل آینده است (Luo, 2024). این تحول بنیادین، تقابلی عمیق میان «منطق بهینه‌سازی الگوریتمی» و «حکمت پیچیده تربیتی» ایجاد کرده و پرسش‌های اساسی را درباره آینده تربیت اخلاقی مطرح ساخته است.

از یک سو، هوش مصنوعی ظرفیت‌های کم‌سابقه‌ای برای شخصی‌سازی تربیت، تعمیق یادگیری اخلاقی از طریق سناریوهای واقعیت مجازی و ارزیابی مستمر رشد فضایل اخلاقی فراهم می‌آورد، اما از سوی دیگر، همین قابلیت‌ها در صورت فقدان چهارچوب‌های نظارتی و فلسفی، مخاطرات جدی به همراه دارند. بررسی پیشینه‌ها نشان‌دهنده یک شکاف مطالعاتی معنادار است: پژوهش‌های جهانی عمدتاً بر ابعاد فنی - کاربردی (Han et al., 2025) یا اصول اخلاقی کلی و جهان‌شمول (Jobin, Ienca & Vayena, 2019) متمرکز هستند و ادبیات داخلی نیز بیشتر به سطح سیاست‌گذاری کلان (طوقدار حسینی و اسدزاده، ۱۴۰۴) یا نقش نظارتی معلم (عابدزاده و همکاران، ۱۴۰۳) پرداخته است. در این میان، جای خالی پژوهشی که به طراحی یک الگوی عملیاتی و یکپارچه برای کاربردهای هوش مصنوعی در فرایند مستقیم تربیت اخلاقی دانش‌آموزان با توجه به بافت فرهنگی و ارزشی ایران بپردازد، به شدت احساس می‌شود.

نادیده گرفتن این شکاف، نظام آموزشی را با خطر واگذاری بی‌ضابطه عاملیت اخلاقی به الگوریتم‌های «جعبه سیاه»، ترویج سوگیری‌های فرهنگی پنهان در داده‌ها (Caton & Haas, 2020) و در نهایت، جایگزینی «اخلاق ماشینی مبتنی بر داده و بهینه‌سازی» به جای «اخلاق انسانی مبتنی بر تأمل، وجدان و کرامت» مواجه می‌سازد. نگرانی از این تعارض میان کارآمدی فناورانه و استقلال انسانی، امری جدی است که در میان دانشگاهیان نیز بازتاب یافته است (Reina et al., 2025).

بنابراین، مسئله محوری این پژوهش، گذار از رویکردهای صرفاً «هشدارمحور» یا «فناوری‌زده» و حرکت به سوی یک رویکرد «ظرفیت‌ساز و مسئولیت‌محور» است. بر این اساس، پژوهش حاضر در پی پاسخ به این پرسش‌های بنیادین است:

۱. قابلیت‌ها و ظرفیت‌های آموزشی و تربیتی هوش مصنوعی برای تعمیق تربیت اخلاقی دانش‌آموزان کدام‌اند؟

۲. چالش‌های بنیادین اخلاقی، تربیتی و فنی کاربردهای هوش مصنوعی در فرایند تربیت اخلاقی چیست؟

هوش مصنوعی

هوش مصنوعی شاخه‌ای میان‌رشته‌ای از علوم شناختی، مهندسی رایانه، منطق و فلسفه ذهن است که با طراحی سامانه‌هایی سروکار دارد که قادر به انجام رفتارهایی هوشمندانه مشابه انسان، از جمله یادگیری، استدلال، پیش‌بینی، تصمیم‌گیری و حل مسئله‌اند (Russell & Norvig, 2020).

هسته اصلی هوش مصنوعی بر سه مؤلفه بنیادی استوار است: الگوریتم‌های یادگیرنده، داده‌های کلان (Big Data) و قدرت محاسباتی بالا؛ این عناصر به سامانه اجازه می‌دهد از تجربه بیاموزد، الگوهای پنهان را کشف کند و در شرایط جدید تصمیم‌اقتضایی بگیرد. هوش مصنوعی از منظر دامنه عملکرد به دو سطح تقسیم می‌شود:

۱. هوش مصنوعی محدود (Narrow or Weak AI) که بر انجام وظایف خاصی همچون ترجمه ماشینی، تشخیص تصویر یا تحلیل احساسات تمرکز دارد؛

۲. هوش مصنوعی عمومی (General or Strong AI) که در سطح نظری، واجد توانایی استدلال و حل مسئله در همه حوزه‌ها همچون ذهن انسانی فرض می‌شود.

در مرحله کنونی، توسعه‌های چشمگیر در یادگیری ماشین (Machine Learning) و به‌ویژه یادگیری عمیق (Deep Learning) موجب ایجاد نسل تازه‌ای از سامانه‌های خودآموز و تعاملی شده است که می‌توانند در عرصه‌های گسترده‌ای از پزشکی (تشخیص تصویری و پایش بالینی)، صنعت و بانکداری گرفته تا آموزش و خدمات اجتماعی به جایگزینی بخشی از فعالیت‌های انسانی بینجامند (Jordan & Mitchell, 2015; Russell & Norvig, 2020).

در حوزه تربیت و آموزش، هوش مصنوعی با قابلیت تحلیل رفتار یادگیرنده، شبیه‌سازی موقعیت‌های آموزشی، و ارائه بازخوردهای بلادرنگ، چهره تازه‌ای از یادگیری شخصی‌سازی‌شده را به نمایش گذاشته است (Luo, 2024; Holmes et al., 2022). باین‌حال، همان توانایی‌ها در صورت فقدان چهارچوب ارزشی و نظارت اخلاقی، می‌توانند به ابزارگرایی تربیتی، تقلیل تعامل انسانی و تضعیف عاملیت اخلاقی منجر گردند (Han et al., 2025). از این‌رو، فهم ماهیت و محدودیت‌های هوش مصنوعی نه صرفاً به‌عنوان فناوری، بلکه در مقام «عامل تربیتی بالقوه» برای آموزش اخلاق، اهمیت نظری ویژه‌ای دارد.

تربیت اخلاقی

واژه تربیت در لغت از ریشه «رَبَوَ» و «رَبَّ» به معنای پرورش، رشد و رساندن موجود به کمال مقتضای خویش است (فراهیدی، ۱۴۰۹: ۸ / ۲۳۸). در قلمرو تربیت اسلامی، تربیت فرایندی آگاهانه و جهت‌مند

برای فعلیت‌بخشی به استعداد‌های انسانی در مسیر قرب الهی و شکوفایی کرامت ذاتی انسان است (گروه علوم تربیتی مؤسسه امام، ۱۳۸۸: ۵۷). «اخلاق» جمع «خُلُق» و به معنای سرشت پایدار و سجیه نفسانی است (معین، ۱۳۷۸). در اصطلاح فلسفی، اخلاق به ملکات راسخی در نفس گفته می‌شود که افعال متناسب با آن بدون نیاز به تأمل از انسان صادر می‌گردد (طباطبایی، ۱۳۸۸: ۱ / ۱۱۱). علامه طباطبایی علم اخلاق را دانشی می‌داند که از چگونگی اکتساب و تثبیت فضایل و از طریق آن، تحقق سعادت انسانی سخن می‌گوید (همان: ۳۷۰). تربیت اخلاقی، در همین چهارچوب، فرایندی تدریجی، منظم و هدفمند است برای ایجاد، تقویت و تثبیت فضایل اخلاقی و زدودن رذایل در سه بعد شناختی، هیجانی و رفتاری (باقری، ۱۳۷۷؛ فیروزمهر، ۱۳۹۵). این فرایند به‌گونه‌ای سامان می‌یابد که میان معرفت اخلاقی و عمل صالح پیوندی پایدار برقرار شود.

تلفیق مفهومی اخلاق و هوش مصنوعی در بستر تعلیم و تربیت

با ظهور فناوری‌های هوشمند، مرزهای تصمیم‌گیری انسانی و سامانه‌های دگرگون شده و این پرسش اساسی پیش روی پژوهشگران اخلاق تربیتی قرار گرفته است که چگونه می‌توان فناوری را به خدمت پرورش فضایل، نه جایگزین وجدان انسانی، درآورد. هوش مصنوعی به دلیل توانش تحلیل داده‌های رفتاری و پیش‌بینی تصمیم‌های اخلاقی، می‌تواند ابزاری در خدمت آموزش ارزش‌ها باشد؛ اما در غیاب نظارت ارزشی، همان ظرفیت به منشأ تضعیف حس مسئولیت، حریم خصوصی و استقلال اخلاقی می‌انجامد (Reina et al., 2025).

از دیدگاه تعلیم و تربیت اسلامی، فناوری مادامی ارزش تربیتی دارد که در مسیر رشد عقلانی و اخلاقی انسان به کار گرفته شود و در خدمت «تسخیر در جهت تعالی انسان» قرار گیرد، نه تسلط بر او. از این منظر، ساختار نظری پژوهش بر تعامل دوسویه اخلاق و فناوری بر پایه الگوی تلفیقی «ظرفیت‌ساز و مسئولیت‌محور» بنا می‌شود؛ الگویی که هم بر ظرفیت‌سازی فناورانه بر آموزش اخلاق تأکید دارد و هم بر مسئولیت انسانی در طراحی، هدایت و کنترل پیامدهای اخلاقی فناوری هوشمند.

در پرتو مباحث فوق، مبانی نظری این پژوهش بر سه باور استوار است:

۱. اخلاق جوهره تربیت است و هر فناوری باید در خدمت تحقق آن باشد.

۲. هوش مصنوعی ابزاری دوجوهی است؛ می‌تواند موجب شکوفایی ظرفیت‌های اخلاقی شود یا در صورت غفلت، به فروگاهی وجدان اخلاقی بینجامد.

۳. ترکیب هوشمندانه این دو حوزه، نیازمند الگویی است که هم بسترساز توسعه ظرفیت‌های پرورشی هوش مصنوعی باشد و هم نگرهبان ارزش‌ها و عاملیت اخلاقی انسان.

بر همین پایه، پژوهش حاضر الگوی تلفیقی «ظرفیت‌ساز و مسئولیت‌محور» را به منزله چهارچوب نظری خود برمی‌گزیند.

ظرفیت‌سنجی هوش مصنوعی در تربیت اخلاقی دانش‌آموزان

تحول‌های فناورانه دهه اخیر، به‌ویژه در حوزه هوش مصنوعی، نه‌تنها شیوه‌های یاددهی و یادگیری، بلکه بنیان‌های مفهومی تربیت را نیز متأثر ساخته است. در نظام آموزشی امروز، فناوری صرفاً ابزار کمکی نیست؛ بلکه به‌عنوان «عامل یادگیرنده» و گاه «همراه تربیتی» عمل می‌کند و می‌تواند در شکل‌دهی ارزش‌ها، هیجانات و تصمیم‌های اخلاقی دانش‌آموزان نقش مستقیم ایفا کند. از این‌رو، بازشناسی ظرفیت‌های هوش مصنوعی در فرایند تربیت اخلاقی، ضرورتی علمی و سیاستی محسوب می‌شود. این ظرفیت‌های کلیدی هوش مصنوعی در این فرایند را می‌توان در چهار محور تحلیلی زیر طبقه‌بندی نمود:

۱. شخصی‌سازی یادگیری اخلاقی از طریق تحلیل داده‌های تربیتی

هوش مصنوعی با تحلیل الگوهای یادگیری و بازخوردهای رفتاری دانش‌آموزان، امکان ارائه مسیرهای آموزشی متناسب با سطح درک، نیازهای شناختی و زمینه فرهنگی را فراهم می‌آورد. فناوری‌های هوشمند می‌توانند محتوای تربیت اخلاقی را براساس شاخص‌های یادگیری و هیجانی دانش‌آموز تطبیق دهند و بدین‌وسیله، از آموزش یکسان به آموزش اخلاقی متناسب با شخصیت و ظرفیت فردی گذر کنند. از ثمرات این ویژگی می‌توان به تقویت عدالت در یاددهی و یادگیری ارزش‌ها اشاره کرد که از سلطه الگوهای اخلاقی یکنواخت جلوگیری می‌نماید (Luo, 2024).

۲. شبیه‌سازی و تجربه موقعیت‌های اخلاقی در محیط‌های تعاملی

یکی از ظرفیت‌های بنیادین هوش مصنوعی، توانایی بازآفرینی موقعیت‌های اخلاقی واقعی در قالب محیط‌های تعاملی و شبیه‌سازی‌شده «واقعیت مجازی»^۱ و «واقعیت افزوده»^۲ است. چنین شبیه‌سازی‌هایی به دانش‌آموزان امکان می‌دهد تا در محیطی ایمن، تصمیمات اخلاقی خود را بیازمایند و بازخورد الگوریتمی مبتنی بر اصول اخلاقی دریافت کنند. این تجربه‌ها موجب درونی‌سازی ارزش‌های اخلاقی، تقویت خودآگاهی و تکوین قضاوت اخلاقی عملی می‌شوند (Han et al., 2025).

۱. Virtual Reality: در این فناوری، کاربر به‌طور کامل وارد یک محیط دیجیتال شبیه‌سازی‌شده می‌شود و از طریق هدست یا تجهیزات خاص، تجربه‌ای همه‌جانبه و جدا از محیط واقعی دارد.

۲. Augmented Reality: در این فناوری، عناصر و اشیای دیجیتال (تصاویر، داده‌ها یا مدل‌های سه‌بعدی) روی دنیای واقعی کاربر افزوده می‌شوند، مثلاً با عینک یا تلفن همراه، کاربر ضمن مشاهده محیط واقعی، اطلاعات یا تصاویر مجازی را نیز هم‌زمان می‌بیند.

۳. ارزیابی هوشمند رشد اخلاقی و بازخورد چندبعدی

با تلفیق یادگیری ماشینی و تحلیل رفتاری، الگوریتم‌های آموزشی می‌توانند الگوهای تصمیم‌گیری اخلاقی، واکنش هیجانی، و ثبات در ارزش‌مداری را ردیابی و تحلیل کنند. چنین ارزیابی‌هایی می‌توانند به مربیان در تشخیص روند رشد اخلاقی و نیازهای تربیتی هر دانش‌آموز کمک کنند. ارزیابی اخلاقی مبتنی بر داده، اگر با نظارت انسانی همراه باشد، دقت تشخیص و مداخله تربیتی را افزایش می‌دهد و فرایند بازخورددهی را از حالت انفعالی به فرایندی مستمر و فراگیر ارتقا می‌بخشد (UNESCO, 2021).

۴. پرورش حساسیت اخلاقی از رهگذر سامانه‌های احساسی و تشخیص هیجان

هوش مصنوعی عاطفی^۱ قادر است هیجان‌های چهره، لحن صوتی، و نشانه‌های رفتاری دانش‌آموز را تحلیل کرده و الگوهای واکنش عاطفی او را به موقعیت‌های اخلاقی شناسایی کند. این فناوری امکان بازنمایی زنده رابطه هیجان و اخلاق را در محیط‌های یادگیری اخلاقی فراهم می‌سازد؛ یعنی سامانه می‌تواند تشخیص دهد که دانش‌آموز در برابر موقعیت‌های ناعادلانه یا رفتارهای غیراخلاقی چه نوع عکس‌العمل هیجانی و عاطفی‌ای از نشان می‌دهد، تا بر آن اساس بازخورد تربیتی متناسب ارائه کند. چنین فرایندی، که با نظارت انسانی همراه است، ظرفیت هوش مصنوعی را از صرف آموزش شناختی ارزش‌ها فراتر برده و آن را به ابزار مدیریت و تقویت همدلی اخلاقی و کنترل هیجانات اخلاقی بدل می‌سازد (Zhou et al., 2025).

جمع‌بندی تحلیلی

ظرفیت‌های هوش مصنوعی در تربیت اخلاقی، نویدبخش یک تحول عمیق است. با این حال، وصول به «عملکرد اخلاقی و مؤثر» مستلزم پذیرش یک چهارچوب سه‌گانه است:

۱. هوش مصنوعی باید نه به‌عنوان جایگزین، بلکه به‌عنوان ابزار توانمندساز مربی عمل کند. قضاوت نهایی، تفسیر زمینه‌ای و گفت‌وگوی همدلانه همواره باید در حیطه مسئولیت انسان باقی بماند.
۲. اصل «شفافیت و توضیح‌پذیری»: الگوریتم‌ها باید مشخص باشد که سیستم بر چه اساس و با تکیه بر کدام نظریه اخلاقی، یک تصمیم یا بازخورد را ارائه می‌دهد. این شفافیت برای ارزیابی انتقادی و اعتماد به سیستم ضروری است.
۳. اصل «حساسیت فرهنگی و بومی‌سازی ارزشی»: طراحی و پیاده‌سازی این سامانه‌ها باید با

۱. Affective AI: هوش هیجانی یا عاطفی مصنوعی یکی از شاخه‌های پیشرو هوش مصنوعی در سال‌های اخیر (به‌ویژه ۲۰۲۴ و ۲۰۲۵) است که تمرکز آن بر درک، تفسیر و پاسخ‌دهی به هیجان‌ها و احساسات انسانی است.

مشارکت فعال متخصصان علوم انسانی، جامعه‌شناسان و فلاسفه اخلاق از فرهنگ‌های مختلف صورت گیرد تا از انطباق سیستم با ارزش‌های بومی و جلوگیری از هژمونی یک دیدگاه اخلاقی خاص اطمینان حاصل شود.

بدون در نظر گرفتن این اصول، ما با خطر ساخت ابزارهایی مواجهیم که در بهترین حالت ناکارآمد و در بدترین حالت، به ابزاری برای القای یکنواختی اخلاقی و نظارت فراگیر بدل خواهند شد. بنابراین، بحث اصلی دیگر «آیا می‌توانیم؟» نیست، بلکه «چگونه باید؟» و «تحت چه شرایط اخلاقی و فلسفی؟» است. این پرسش‌ها، بنیان بخش بعدی (۵ - ۲) در تحلیل چالش‌ها را تشکیل می‌دهند.

چالش‌های محوری و تحلیل انتقادی آنها

در ادبیات پژوهش، واژه «چالش» (Challenge) به معنای صرف «مشکل» یا «مانع» نیست، بلکه به موقعیت‌هایی اشاره دارد که در آنها توانایی‌های فناوری با حدود اخلاقی و تربیتی انسان در تعامل یا تعارض قرار می‌گیرد. در این شرایط، پژوهشگر باید فراتر از توصیف پدیده، به تحلیل انتقادی ماهیت تعارض‌ها و پیامدهای اخلاقی آنها بپردازد. بنابراین، طرح «چالش‌های اخلاقی هوش مصنوعی» در آموزش به معنای بررسی حدود مسئولیت انسان و میزان نظارت اخلاقی بر رفتار الگوریتم‌ها است. در بررسی کاربست‌های تربیتی هوش مصنوعی، رویارویی با شش دسته چالش بنیادین اجتناب‌ناپذیر است. هر یک از این چالش‌ها بازتابی از تنش میان منطق فناورانه و ارزش‌های انسانی - اخلاقی‌اند؛ از این رو لازم است ماهیت هر چالش شفاف تعریف شود و سپس با نگاهی انتقادی و مبتنی بر مبانی تربیت اسلامی - ایرانی تحلیل گردد.

نخستین چالش: عدالت آموزشی به واسطه سوگیری الگوریتمی^۱ (Algorithmic Bias)

الگوریتم‌های به‌کاررفته در نظام‌های آموزشی و اخلاقی می‌توانند به‌صورت ناخواسته نگرش‌های ناعادلانه را بازتولید کنند. این امر اغلب ناشی از داده‌های آموزشی ناقص، سوگیری فرهنگی یا جنسیتی، و الگویی از نظام‌های ارزشی مغایر با زمینه بومی است. راه برون‌رفت از این چالش، توسعه مدل‌های اخلاق‌محور و آموزش داده‌ها بر پایه نمایندگی فرهنگی متنوع است. در نگاه بومی - اسلامی، عدالت یا

۱. عدالت آموزشی به معنای فراهم‌سازی فرصت‌های یادگیری، منابع آموزشی و نتایج تحصیلی براساس نیازهای واقعی هر فراگیر است، نه صرفاً یکسان‌سازی سطحی امکانات. در این تعریف، عدالت آموزشی مستلزم آن است که جنسیت، طبقه اجتماعی، زبان مادری، یا موقعیت جغرافیایی، مانع یادگیری و رشد ارزش‌های اخلاقی دانش‌آموز نشود. راه اصلاح، طراحی داده‌محور عادلانه، به‌کارگیری بازنمایی متنوع فرهنگی در داده‌های آموزشی، و نظارت انسانی اخلاق‌محور است تا «عدالت توزیعی و اخلاقی» جایگزین کارآمدی صرف گردد (UNESCO, 2021).

«قِسْطًا» نه صرفاً برابری عددی، بلکه توازن ارزش‌ها و حقوق انسانی است. قرآن کریم در آیه «إِنَّ اللَّهَ يَأْمُرُ بِالْعَدْلِ وَالْإِحْسَانِ» (نحل / ۹۰) این اصل را بنیاد همه سامان اجتماعی معرفی می‌کند. بر این اساس، طراحی فناوریانه هوش مصنوعی باید تابع عدالت توزیعی و اخلاقی باشد نه صرفاً کارآمدی آماری (Mehrabi et al., 2022؛ توکلی و همکاران، ۱۴۰۲).

نمونه چالش: آموزش و پرورش از «سامانه تصحیح هوشمند انشا» رونمایی می‌کند. این سیستم با هزاران انشای ممتاز از مدارس خاص شهرهای بزرگ آموزش دیده است. دانش‌آموزی از استان ایلام در انشای خود با موضوع «علم بهتر است یا ثروت؟»، برای بیان استدلالش از یک مثل زیبایی‌گرددی استفاده می‌کند که سرشار از حکمت است. سیستم هوشمند، این مثل را به‌عنوان «عبارت نامرتب و خارج از معیار» شناسایی کرده و نمره دانش‌آموز را کم می‌کند، زیرا این عبارت در داده‌های آموزشی «معیار» آن وجود نداشته است.

چالش دوم: هم‌سویی ارزش‌ها و جهان‌بینی اخلاقی^۱ (Educational Equity)

منظور از هم‌سویی ارزش‌ها و جهان‌بینی اخلاقی، تطابق نظام تصمیم‌گیری هوش مصنوعی با اصول فرهنگی و اخلاقی جامعه میزبان است. چون بیشتر سامانه‌های هوش مصنوعی در بستر جهان‌بینی غربی تربیت یافته‌اند، انتقال مستقیم آنها به آموزش اخلاقی بدون تنظیم ارزشی می‌تواند منجر به انتقال ناخواسته الگوهای ارزشی بیگانه شود. بر این اساس، تدوین چهارچوب ارزشی ایرانی - اسلامی برای هوش مصنوعی تربیتی به‌منظور پاسداری از استقلال فرهنگی و تضمین تربیت مبتنی بر کرامت، اختیار و مسئولیت اخلاقی انسان ضروری است. سازوکارهای تصمیم‌گیر هوش مصنوعی به دلیل تربیت الگوریتمی در بسترهای فرهنگی خاص، معمولاً در درک ارزش‌های چندفرهنگی و بومی ناتوان‌اند. اگر آموزش اخلاقی بر پایه چنین سامانه‌هایی انجام گیرد، خطر استحاله فرهنگی و انتقال ارزش‌های غیرهم‌سو افزایش می‌یابد. تحلیل انتقادی این مسئله نشان می‌دهد که راه چاره در تدوین «چهارچوب ارزشی ایرانی - اسلامی برای هوش مصنوعی تربیتی» است؛ چهارچوبی که در آن کرامت انسان، اختیار اخلاقی، و مسئولیت در برابر خالق به‌عنوان اصول راهنما وارد طراحی سامانه و محتوا شود. به‌کارگیری چنین مدلی، استقلال اخلاقی در مواجهه با جهان فناوری‌محور را تضمین می‌کند.

((Bostrom, 2014; Gabriel, 2020

۱. مقصود از هم‌سویی ارزش‌ها و جهان‌بینی اخلاقی (Value and Moral Worldview Alignment) میزان سازگاری بین نظام تصمیم‌گیری و یادگیری ماشین یک سامانه هوش مصنوعی با نظام ارزش‌ها، باورها و اصول هنجاری جامعه انسانی استفاده‌کننده از آن است (Bostrom, 2014; Gabriel, 2020).

نمونه چالش: یک بازی آموزشی اخلاقی به نام «قهرمان اخلاق» در مدارس توزیع می‌شود. در یکی از مراحل، دانش‌آموز در یک پروژه گروهی قرار می‌گیرد که یکی از اعضا کم‌کاری می‌کند. بازی دو گزینه پیش روی او می‌گذارد: الف) گزارش دادن هم‌گروهی به معلم برای گرفتن نمره کامل؛ ب) کمک به هم‌گروهی و بیدار ماندن تا صبح تا کار او هم تمام شود. الگوریتم که براساس فرهنگ فردگرایانه طراحی شده، گزینه «الف» را تشویق می‌کند؛ اما دانش‌آموز ایرانی که براساس فرهنگ «مرام» و «رفاقت» گزینه «ب» را انتخاب می‌کند، با پیام «شما با این کار به دوستان خیانت کردید و مانع رشد او شدید» مواجه شده و امتیاز از دست می‌دهد.

سومین چالش: شفافیت و پاسخگویی (Transparency and Accountability)^۱

وقتی می‌گوییم «چالش شفافیت و پاسخ‌گویی» در هوش مصنوعی، منظورمان این است که ما دقیقاً نمی‌دانیم ماشین چگونه تصمیم می‌گیرد و اگر خطا یا بی‌عدالتی پیش بیاید، چه کسی باید پاسخ‌گو باشد. ساختار موسوم به «جعبه سیاه» در بسیاری از الگوریتم‌های یادگیری عمیق مانع نظارت اخلاقی بر تصمیم‌های سیستم می‌شود. کاربر نمی‌داند چگونه سیستم به یک ارزیابی یا توصیه اخلاقی رسیده است و همین ابهام، مسئولیت‌پذیری را تضعیف می‌کند. تحلیل اخلاقی این وضعیت اقتضا می‌کند که نظامی از پاسخ‌گویی دوگانه (انسان - ماشین) برقرار شود؛ بدین معنا که همواره مسئولیت نهایی تفسیر و تصمیم‌گیری متوجه ناظر انسانی است و فناوری صرفاً نقش ابزار دارد. چنین نظامی از دیدگاه اخلاق اسلامی نیز واجد اهمیت است؛ زیرا در جهان‌بینی توحیدی، «تکلیف اخلاقی» قائم به اراده انسان عاقل است نه ابزار بی‌اختیار (Floridi & Cowls, 2019; UNESCO, 2021; Reina Marín et al., 2025).

نمونه چالش: یک «سیستم پایش هوشمند انضباط» با تحلیل دوربین‌های راهرو، به صورت خودکار دانش‌آموزان واجد شرایط برای اردوی سالانه را انتخاب می‌کند. نام یکی از دانش‌آموزان خوب کلاس در لیست اردو نیست. وقتی والدینش علت را از مدیر جویا می‌شوند، مدیر سیستم را چک کرده و می‌گوید: «سیستم به دلیل نمره انضباط پایین ایشان را حذف کرده». والدین می‌پرسند کدام رفتار باعث کسر نمره شده؟ مدیر پاسخ می‌دهد: «سیستم فقط ثبت کرده: تجمع غیرمجاز در راهرو، ساعت ۱۰:۱۵؛ اما فیلم یا جزئیات بیشتری ارائه نمی‌دهد». در واقع، دانش‌آموز فقط برای کمک به دوستی که وسایلش روی زمین ریخته بود، ایستاده بود، اما چون هیچ‌کس نمی‌داند منطق «جعبه سیاه» چیست، حق دانش‌آموز ضایع می‌شود.

۱. Transparency and Accountability: منظور از چالش شفافیت و پاسخ‌گویی، ناتوانی در درک و نظارت بر تصمیم‌های اخلاقی سامانه‌های یادگیری عمیق است که باعث تضعیف مسئولیت‌پذیری انسانی می‌شود. ریشه‌یابی فنی و اخلاقی آن اقتضا دارد که با طراحی الگوریتم‌های تبیین‌پذیر و تعیین مسئولیت دوگانه (فناور - مربی)، کرامت و تکلیف اخلاقی انسان در روند تربیت محفوظ بماند.

چهارمین چالش: حریم خصوصی و امنیت داده (Data Privacy and Security)

گردآوری و تحلیل داده‌های شخصی دانش‌آموزان - از عاطفه و گفتار تا رفتار آنلاین - تهدیدی جدی برای رضایت آگاهانه و حق کنترل داده محسوب می‌شود. بی‌توجهی به این حق می‌تواند اعتماد تربیتی را متزلزل کند. در منظومه اخلاق اسلامی، حرمت افشای اسرار و احترام به مالکیت داده‌ها از مصادیق «امانت‌داری» و «حق‌الناس» است؛ نقض این اصول نوعی تجاوز اخلاقی به حوزه فردی انسان به شمار می‌آید. بنابراین، توسعه سواد داده‌ای و تربیت در خصوص حریم اطلاعات شخصی باید به‌عنوان بخشی از تعلیم اخلاق دیجیتال در مدارس گنجانده شود (Voigt & Von dem Bussche, 2017).

نمونه چالش: در اپلیکیشن رسمی مدرسه (مانند شاد)، یک «چت‌بات مشاور هوشمند» برای دانش‌آموزان فعال می‌شود تا با آن «محرمانه» درد دل کنند. دانش‌آموزی که از اضطراب امتحان رنج می‌برد، چندین بار در مورد ترس از شکست با این چت‌بات صحبت می‌کند. در جلسه اولیا و مربیان، معلم پرورشی با نیت خیر به مادر دانش‌آموز می‌گوید: «طبق گزارش سیستم، فرزند شما هشت بار از کلمه اضطراب و ترس استفاده کرده. بهتر است بیشتر مراقبش باشید». مادر و دانش‌آموز احساس می‌کنند خصوصی‌ترین بخش زندگی‌شان بدون اجازه فاش شده و اعتمادشان برای همیشه از بین می‌رود.

پنجمین چالش: فرسایش تعامل انسانی و تضعیف تفکر

انتقادی (Erosion of Human Connection and Critical Thinking)

اگر آموزش اخلاقی صرفاً به تجارب ماشینی، آزمون‌های خودکار و توصیه‌های الگوریتمی محدود شود، احتمال افت تعاطی انسانی و گفت‌وگوی تربیتی افزایش می‌یابد. خطر آن است که یادگیری اخلاق به امری مکانیکی و انفعالی تبدیل گردد. تحلیل انتقادی این پدیده بر ضرورت تلفیق یادگیری فناورانه با فعالیت گفت‌وگومحور تأکید دارد؛ روشی که به دانش‌آموز مجال تفسیر، طرح پرسش و نقد ارزش‌ها را می‌دهد. پژوهش‌های تربیتی اخیر نیز نشان می‌دهد که رشد وجدان اخلاقی و تفکر انتقادی زمانی تحقق می‌یابد که تعامل انسانی حفظ شود (Cai et al., 2021).

نمونه چالش: مصداق عینی: معلم ادبیات برای تعمیق درس «دهقان فداکار»، دانش‌آموزان را به استفاده از یک «چت‌بات پیشرفته با قابلیت گفت‌وگوی سقراطی» تشویق می‌کند. هدف این است که دانش‌آموزان با پرسش و پاسخ، به درک عمیق‌تری از مفهوم ایثار برسند. یکی از دانش‌آموزان تیزهوش از چت‌بات می‌پرسد: «اما اگر قطار متوجه نمی‌شد و ریزعلی هم از سرما یخ می‌زد و می‌مرد، آیا کارش همچنان فداکاری بود یا یک کار بیهوده و اشتباه؟ مرز بین شجاعت و بی‌فکری کجاست؟» در این لحظه کلیدی، آن جرعه تفکر انتقادی که در ذهن دانش‌آموز زده شده بود، با یک پاسخ

اطلاعات‌محور و ناتوان از درک عمق سؤال، خاموش می‌شود. یک معلم انسانی از این سؤال استقبال کرده و آن را به یک بحث کلاسی درباره «نیت در عمل»، «نتیجه‌گرایی» و «اخلاق وظیفه‌گرا» تبدیل می‌کرد و دانش‌آموزان را به چالش می‌کشید؛ اما هوش مصنوعی، آن لحظه طلایی «تردید سازنده» را که اساس پرورش تفکر انتقادی است، از بین می‌برد. در اینجا، تکنولوژی نه تنها جایگزین تعامل انسانی نشد، بلکه فعالانه «فرصت تفکر عمیق» را نابود کرد.

ششمین چالش: کاهش هوش عاطفی و حمایت روانی

(Decline of Emotional Intelligence and Psychological Support)

سامانه‌های هوشمند هنوز فاقد توان همدلی اصیل انسانی‌اند و نمی‌توانند با ظرافتی مشابه مربی، حالات روحی و زمینه‌های فرهنگی یادگیرنده را درک کنند. اتکای افراطی به این سامانه‌ها خطر خشکی عاطفی و بی‌تفاوتی اخلاقی را در یادگیرندگان بالا می‌برد. پاسخ تربیتی، تأکید بر نقش معلم به‌عنوان میانجی هیجانی و تسهیلگر اخلاقی در محیط دیجیتال است. هوش مصنوعی باید در خدمت معلم قرار گیرد، نه جایگزین او؛ زیرا در تربیت اخلاقی، انتقال احساس و همدلی‌بخش جدایی‌ناپذیر فرایند رشد است.

نمونه چالش: مصداق عینی: یک دانش‌آموز ممتاز و باانگیزه به نام «پریسا»، ناگهان در اواسط سال تحصیلی با افت شدید نمرات در دروس ریاضی و فیزیک مواجه می‌شود. «سامانه هوشمند پیشرفت تحصیلی» مدرسه، این افت را فوراً تشخیص می‌دهد. الگوریتم، داده‌ها را تحلیل می‌کند: نمرات پایین آمده، زمان صرف‌شده در پلتفرم آموزشی آنلاین کم شده، و تمرینات کمتری حل شده است. نتیجه‌گیری منطقی سیستم این است که پریسا دچار افت انگیزه یا مدیریت زمان شده است. بنابراین، سیستم به صورت خودکار یک برنامه جبرانی برای او ارسال می‌کند: «پریسای عزیز، طبق تحلیل عملکرد، برنامه مطالعه شما بهینه نیست. پیشنهاد می‌شود روزانه ۴۵ دقیقه به تمرینات ریاضی اضافه کنید. همچنین سه ویدیوی انگیزشی با موضوع «غلبه بر افت تحصیلی» برای شما بارگذاری شد. موفق باشی!»

تحلیل چالش: واقعیت این است که والدین پریسا در آستانه جدایی هستند و او به دلیل اضطراب و آشفتگی عاطفی شدید، توان تمرکز خود را از دست داده است. در این شرایط، آخرین چیزی که او نیاز دارد، فشار درسی بیشتر و ویدیوهای انگیزشی کلیشه‌ای است. راه‌حل منطقی سیستم، از نظر انسانی «بسیار مضر» است، زیرا پریشانی او را نادیده گرفته و بر اضطرابش می‌افزاید.

اینجا دقیقاً نقطه‌ای است که هوش عاطفی شکست می‌خورد. سیستم، «علامت» (افت نمره) را به درستی تشخیص داد، اما در درک «علت ریشه‌ای» (بحران عاطفی) کاملاً ناتوان بود. یک معلم باتجربه،

با دیدن چهره پریشان پریسا در کلاس، تغییر در رفتار او و صحبت کردن با او، می‌توانست به عمق مشکل پی برده و به جای برنامه درسی فشرده‌تر، حمایت عاطفی و ارجاع به مشاور را پیشنهاد دهد. هوش مصنوعی یک «مشکل عملکردی» دید، در حالی که یک «انسان نیازمند کمک» وجود داشت.

جمع‌بندی تحلیلی

وقتی چالش‌های شش‌گانه را کنار هم قرار می‌دهیم، به جای فهرستی از مشکلات پراکنده، به یک تصویر بزرگ‌تر می‌رسیم: یک تضاد بنیادین در سه لایه اصلی که میان منطق فناوری و اهداف تربیت انسانی وجود دارد.

لایه اول: چالش هویت فرهنگی و ارزش‌های پنهان در کدها

این لایه به ریشه‌ها و ذات هوش مصنوعی می‌پردازد و دو چالش «سوگیری الگوریتم‌ها» و «عدم هم‌خوانی ارزش‌ها» را در بر می‌گیرد.

تحلیل: هر هوش مصنوعی، مانند یک موجود زنده، یک «ریشه فرهنگی» دارد که آن را از داده‌ها و جهان‌بینی طراحانش به ارث می‌برد. وقتی این با فرهنگ اسلامی - ایرانی ما هم‌خوان نباشد، سیستم هوشمند به ابزاری برای تحمیل ناخواسته یک فرهنگ خاص تبدیل می‌شود.

راهکار چیست؟ راه‌حل، صرفاً «اصلاح» یا «فیلتر کردن» سیستم‌های غربی نیست. ما باید هوش مصنوعی را از پایه و با مصالح فرهنگی و ارزشی خودمان بسازیم؛ یعنی اصول اخلاقی ما به جای پوسته، به هسته مرکزی منطق ماشین تبدیل شود.

لایه دوم: چالش قدرت نامرئی و مسئولیت گمشده

این لایه به نحوه عملکرد و کنترل سیستم‌های هوشمند مربوط است و چالش‌های «شفافیت» و «حریم خصوصی» را شامل می‌شود.

تحلیل: ساختار «جعبه سیاه» هوش مصنوعی وضعیتی را ایجاد می‌کند که در آن، ماشین قدرتمند است، اما به هیچ‌کس پاسخگو نیست. دانش‌آموز نمی‌داند چرا یک امتیاز اخلاقی را گرفته و معلم نمی‌داند در صورت بروز خطا، چه کسی مسئول است. این قدرت مبهم، حقوق اساسی انسان مانند حق دانستن و حق داشتن حریم خصوصی را تهدید می‌کند.

راهکار چیست؟ ما به سیستم‌هایی نیاز داریم که از ابتدا شفاف طراحی شده‌اند و مهم‌تر از آن، باید یک اصل اساسی وجود داشته باشد: مسئولیت نهایی همیشه با انسان است. ماشین یک ابزار مشاور است

لایه سوم: چالش فرسایش رابطه انسانی در فرایند تربیت

این لایه به تأثیر مستقیم هوش مصنوعی بر تجربه یادگیری می‌پردازد و دو چالش «کاهش تعامل انسانی» و «ضعف هوش عاطفی» را در خود جای می‌دهد.

تحلیل: تربیت اخلاقی، یک فرایند انسانی، گرم و چهره‌به‌چهره است که در بستر گفت‌وگو، همدلی و نگاه معلم به شاگرد شکل می‌گیرد. اتکای بیش از حد به ماشین، این فرایند زنده را به یک تجربه مکانیکی، سرد و فردی تبدیل می‌کند و مهم‌ترین بخش تربیت، یعنی «رابطه انسانی»، را از آن می‌گیرد. **راهکار چیست؟** باید نقش هوش مصنوعی را بازتعریف کنیم. هوش مصنوعی نباید جایگزین معلم شود، بلکه باید دستیار او باشد؛ ابزاری که کارهای تکراری و تحلیل داده‌ها را انجام می‌دهد تا معلم فرصت و انرژی بیشتری برای ارزشمندترین کار خود پیدا کند: گفت‌وگو، انتقال احساس، الگوسازی و پرورش حکمت در دانش‌آموزان.

توصیه راهبردی: انسانی کردن فناوری، نه ماشینی کردن تربیت یعنی در نهایت، تضاد اصلی نه بین «انسان و ماشین»، بلکه بین «منطق ماشینی» که به دنبال کارایی و بهینه‌سازی است و «منطق تربیتی» که به دنبال رشد و تعالی انسان است قرار دارد. برای موفقیت در این عرصه، راهکار ما باید یک تغییر نگاه بنیادین باشد که بر سه اصل استوار است:

اصالت فرهنگی: طراحی فناوری براساس ارزش‌های بومی خودمان.

حاکمیت انسانی: انسان همیشه باید کنترل نهایی و مسئولیت را در دست داشته باشد.

تقویت‌کننده تربیت: فناوری باید در خدمت تقویت رابطه انسانی باشد، نه حذف آن.

بنابراین، رسالت بزرگ ما نه ساختن یک ماشین معلم اخلاق، بلکه پرورش نسلی از انسان‌های خردمند است که بتوانند از این ابزار قدرتمند، برای انسانی‌تر کردن تربیت بهره ببرند.

نتیجه

تحلیل تطبیقی یافته‌ها نشان می‌دهد که هوش مصنوعی در تربیت اخلاقی پدیده‌ای دوسویه و دیالکتیکی است؛ فناوری‌ای که می‌تواند هم ابزار تعمیق اخلاق باشد و هم عامل تهی‌سازی آن. هوش مصنوعی با تکیه بر یادگیری عمیق، تحلیل داده‌های رفتاری و بازخورد کاربر، ظرفیت‌هایی نو برای شخصی‌سازی، تعمیق و سنجش رشد اخلاقی فراهم می‌آورد؛ اما در فقدان چهارچوب‌های هنجاری و نظارت انسانی، به فروگاهی عاملیت اخلاقی و استحاله فرهنگی منتهی می‌شود.

در بعد ظرفیت‌ها، فناوری‌های مبتنی بر یادگیری ماشینی و تحلیل هیجانی (Affective AI)، امکان

بازخورد لحظه‌ای، شبیه‌سازی موقعیت‌های اخلاقی و پرورش حساسیت عاطفی را ایجاد می‌کنند که در صورت هدایت دینی - انسانی، می‌تواند عدالت آموزشی، کنترل هیجان‌ها و تصمیم‌گیری مسئولانه را تقویت کند.

در بعد چالش‌ها، تهدیدهایی چون سوگیری الگوریتمی (Algorithmic Bias)، ابهام در پاسخ‌گویی (Transparency and Accountability)، نقض حریم خصوصی (Data Privacy)، و فرسایش ارتباط انسانی و هوش عاطفی (Erosion of Human Connection and Affective Gap)، می‌توانند تربیت اخلاقی را از محتوایی انسانی و رابطه‌ای به روندی مکانیکی و غیرشخصی تبدیل کنند.

جمع‌بندی پژوهش آشکار می‌سازد که تعارض بنیادین عصر هوش مصنوعی، نه میان انسان و فناوری، بلکه تضاد میان منطق اخلاقی تربیت و منطق ابزاری فناوری است. برای رهایی از این دوگانگی، ایجاد الگویی تلفیقی ضروری است که در این مطالعه با عنوان «الگوی ظرفیت‌ساز و مسئولیت‌محور» معرفی شده است. این الگو بر پیوند میان توانش فناورانه و ارزش‌های اخلاقی اسلامی - به‌ویژه کرامت انسان، اختیار، عدالت و امانت - تکیه دارد و پیشنهاد می‌کند که:

- هوش مصنوعی باید در خدمت رشد اخلاقی انسان به کار رود، نه جایگزین وجدان و داوری او شود؛
- مسئولیت نهایی هر تصمیم تربیتی، همچنان متعلق به انسان آگاه و مختار است؛
- ترکیب داده‌کاو فناوری با نظارت اخلاقی انسانی، زمینه‌ساز تربیتی عادلانه، شفاف و همدلانه است.
- بر این اساس، تربیت فناورانه اخلاقی آینده تنها زمانی اصیل خواهد بود که هوش مصنوعی از ابزار هدایت‌شده به عامل همکار انسان بدل گردد؛ عاملی که بتواند تحلیل الگوریتمی را در مسیر انسان‌محوری و اخلاق‌مداری هدایت کند. نتیجه نهایی پژوهش تأکید می‌کند که استفاده اخلاقی از هوش مصنوعی در مدارس ایران نیازمند تدوین سیاست‌نامه‌ای داخلی است که بر تعامل سه‌گانه انسان - اخلاق - فناوری استوار باشد تا هوش مصنوعی به جای تهدید، به فرصت تعالی اخلاقی نسل آینده تبدیل شود.

منابع و ماخذ

- باقری، خ (۱۳۷۷). *مبانی و شیوه‌های تربیت اخلاقی*. تهران: سازمان تبلیغات اسلامی.
- توکلی، مهدی و جعفر گل محمدی (۱۴۰۲). *بررسی اثر هوش مصنوعی بر سه فضیلت اخلاقی و راه‌های مواجهه با آنها*. مقاله ارائه‌شده در کنفرانس اخلاق دیجیتال. تهران: دانشکده الهیات و معارف اسلامی.
- خلیل بن احمد؛ محسن آل‌عصفور؛ مهدی مخزومی و ابراهیم سامرائی (۱۴۰۹). *العین*. قم: هجرت.
- دهخدا، علی‌اکبر (۱۳۷۳ - ۱۳۷۲). *لغت‌نامه دهخدا*. تهران: انتشارات و چاپ دانشگاه تهران.

ظرفیت‌سنجی و چالش‌های اخلاقی کاربرت هوش مصنوعی در فرایند تربیت اخلاقی دانش‌آموزان □ ۲۳

- رجیبان ده‌زیره، حسین (۱۴۰۳). شناسایی چالش‌ها و قابلیت‌های هوش مصنوعی در آموزش و یادگیری با ارائه راهکارها. *فناوری آموزشی دانشگاه علامه طباطبایی*. ۱۸ (۳). ۱۱۳ - ۹۱.
- طباطبایی، سید محمدحسین (۱۴۱۷ ق). *المیزان فی تفسیر القرآن*. قم: دفتر انتشارات اسلامی.
- طوقدار حسینی، مهدی و نسرین اسدزاده (۱۴۰۴). چالش‌ها و فرصت‌های اخلاقی هوش مصنوعی در علوم تربیتی. *مطالعات روان‌شناسی و علوم تربیتی (نگاره)*. ۱۴ (۲). ۹۷ - ۷۵.
- عابدزاده، سعید؛ حسین عابدزاده و مریم طاهری نوش آبادی (۱۴۰۳). نقش معلمان در هدایت استفاده اخلاقی و مؤثر از هوش مصنوعی در کلاس‌های درس. *پژوهش‌های نوین در تعلیم و تربیت اسلامی*. ۱۱ (۱). ۲۱ - ۱.
- فیروزمهر، محمدمهدی (۱۳۹۳). *تربیت اخلاقی با نگاهی قرآنی*. قم: پژوهشگاه علوم و فرهنگ اسلامی.
- گروه علوم تربیتی مؤسسه آموزشی و پژوهشی امام خمینی (۱۳۸۸). *مجموعه مقالات تربیتی*. قم: مؤسسه علمی - پژوهشی امام خمینی.
- معین، محمد (۱۳۷۸). *فرهنگ فارسی*. تهران: امیر کبیر.
- Baig, M., & Yadegari - Dehkordi, N. (2024). *AI applications in academic writing and ethical implications: A pedagogical review*. *Journal of Educational Integrity*. 19 (3). 245 – 261.
- Bostrom, N. (2014). *Superintelligence: Paths, dangers, strategies*. Oxford University Press.
- Cai, Z., Fan, X., & Du, J. (2021). *Human - AI interaction and critical thinking in learning environments*. *Computers & Education*. 171, 104233. <https://doi.org/10.1016/j.compedu.2021.104233>
- Caton, S., & Haas, C. (2020). *Fairness, accountability, and transparency in algorithmic systems: A literature review*. *Journal of Information, Communication and Ethics in Society*. 18 (3). 323 - 341. https://doi.org/10.1108/JICES_08_2019_0099
- Floridi, L., & Cowls, J. (2019). *A unified framework of five principles for AI in society*. *Harvard Data Science Review*. 1 (1). <https://doi.org/10.1162/99608f92.8cd550d1>
- Gabriel, I. (2020). *Artificial intelligence, values, and alignment*. *Minds and Machines*. 30 (3). 411 - 437. https://doi.org/10.1007/s11023_020_09548_2
- Han, B. Y., Luo, H. Y., & Zhou, J. (2025). *Students' perceptions: Exploring the interplay of ethical and pedagogical impacts for adopting AI in higher education*. *International Journal of Artificial Intelligence in Education*. 35 (2). 188 - 209. https://doi.org/10.1007/s40593_025_00317_7
- Holmes, W., Bialik, M., & Fadel, C. (2022). *Artificial intelligence in education: Promises and implications for teaching and learning*. Boston, MA: Center for Curriculum Redesign.

- Jobin, A., Ienca, M., & Vayena, E. (2019). *The global landscape of AI ethics guidelines*. *Nature Machine Intelligence*. 1 (9). 389 - 399. https://doi.org/10.1038/s42256_019_0088_2
- Jordan, M. I., & Mitchell, T. M. (2015). *Machine learning: Trends, perspectives, and prospects*. *Science*. 349 (6245). 255 - 260. <https://doi.org/10.1126/science.aaa8415>
- Luo, H. Y. (2024). *Reflections on school moral education in the age of artificial intelligence*. *Open Access Library Journal*. 11 (2). 1 - 10. <https://doi.org/10.4236/oalib.1101157>
- Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2022). *A survey on bias and fairness in machine learning*. *ACM Computing Surveys*. 55 (3). 1 - 35. <https://doi.org/10.1145/3494672>
- Reina Marín, Y., Cruz Caro, O., Carrasco Rituay, A. M., Guimac Llanos, K. A., Tarrillo Perez, D., Sánchez Bardales, E., & Alva Tuesta, J. N. (2025). *Ethical challenges associated with the use of artificial intelligence in university education*. *Journal of Academic Ethics*. 23 (4). 2443 - 2467. https://doi.org/10.1007/s10805_025_09660_w
- Russell, S. J., & Norvig, P. (2020). *Artificial intelligence: A modern approach* (4th ed.). Pearson Education.
- UNESCO. (2021). *Recommendation on the ethics of artificial intelligence*. Paris: UNESCO Publishing. <https://unesdoc.unesco.org/ark:/48223/pf0000379920>
- Voigt, P., & von dem Bussche, A. (2017). *The EU General Data Protection Regulation (GDPR): A practical guide*. Springer. https://doi.org/10.1007/978_3_319_57959_7
- Zhou, J., Han, B. Y., & Luo, H. Y. (2025). *Bridging the affective gap: Emotional AI and moral sensitivity in digital education*. *Computers & Education: Artificial Intelligence*. 6, 100223. <https://doi.org/10.1016/j.caeai.2025.100223>