



Epistemological Foundations of Trust in Artificial Intelligence within the Islamic Systematization: An Analysis of the Hallucination Phenomenon Based on the Views of Allameh Tabataba'i

Emameenassab Saeed¹ | mohamad jabarpoor² | Mohamad Rabiee³

1. Specialized Doctorate, Humanities, Theology, Shahid Motahari, Tehran, Iran. saeedemameen@gmail.com
2. Assistant Professor, Department of Thought and Cultural Studies, Imam Hossein Comprehensive University, Tehran, Iran. fadak4381@gmail.com
3. Researcher, Media Think Tank, Shahid Sadr Research Institute, Imam Hossein Comprehensive University, Tehran, Iran. sepehrtry@gmail.com

Abstract

The phenomenon of "hallucination" in generative artificial intelligence models and large language models is one of the most serious technical and epistemological challenges in the path of developing contemporary intelligent technologies. The present research, with the aim of fundamentally explaining this phenomenon from the perspective of complementary philosophy, undertakes an investigation of the computational mechanism of the machine in confrontation with the epistemological system of Allameh Tabatabai. The main problem of the research is how the phenomenon of hallucination is explained on the basis of the confrontation between the realist theory of truth as correspondence and the theory of coherence, and what limitations it creates for civilizational rationality. This research has been organized with a descriptive-analytical method and the tools of library data collection and note-taking. The findings of the research show that artificial intelligence, due to its reliance on probabilistic processes of word prediction, defines truth solely within the framework of "statistical coherence theory" and syntactic relations internal to the structure. In contrast, from the perspective of Allameh Tabatabai, the criterion of truth is "correspondence with the reality-in-itself (nafs al-amr)," and this matter necessitates the dependence of acquired knowledge (knowledge by presence) upon the intuition of "knowledge by presence" and the immaterial knowing agent. This article demonstrates that hallucination is not a temporary technical bug, but rather an inherent requirement of acquired knowledge severed from presential intuition in an architecture devoid of a machine self. Ultimately, this strategic implication is drawn that the delegation of epistemological authority and macro-level decision-making in the governance systems of modern Islamic civilization to systems based on statistical coherence is, due to the structural error of hallucination, devoid of rational and epistemological legitimacy.

Keywords: Constructed perceptions (i'tibariyyat), artificial intelligence hallucination, civilizational rationality, knowledge by presence, Allameh Tabatabai (may God have mercy upon him).

Cite this article: Saeed, E., jabarpoor., M and Rabiee, M. (2026). Epistemological Foundations of Trust in Artificial Intelligence within the Islamic Systematization: An Analysis of the Hallucination Phenomenon Based on the Views of Allameh Tabataba'i. , 1(1), 48-71.

Publisher: Imam Hossein University.

"Authors Retain The Copyright"

Volume info

Vol. 1
Series: 1
Spring 2026
P.P: 48-71

Article Type

Research Paper

Article History

Received:
23 June 2026
Revised:
28 August 2026
Accepted:
25 September 2026
Published:
23 September 2026

ISSN – E-ISSN

ISSN: 0000-0000
E-ISSN: 0000-0000



OPEN ACCESS





مبانی معرفت شناختی اعتماد به هوش مصنوعی در نظام سازی اسلامی؛ تحلیلی بر پدیده توهم بر اساس آرای علامه طباطبائی

سعید امامی نسب^۱ | محمد جبارپور | محمد ربیعی^۳

۱. نویسنده مسئول: دکترای تخصصی، علوم انسانی، الهیات، شهید مطهری، تهران، ایران. saeedemameen@gmail.com

۲. استادیار گروه اندیشه و مطالعات فرهنگی، دانشگاه جامع امام حسین (ع)، تهران، ایران. fadak4381@gmail.com

۳. پژوهشگر اندیشکده رسانه پژوهشکده شهید صدر (ره)، دانشگاه جامع امام حسین (ع)، تهران، ایران. sepehtrry@gmail.com

چکیده

پدیده «توهم» در مدل های هوش مصنوعی مولد و مدل های زبانی بزرگ، یکی از جدی ترین چالش های فنی و معرفت شناختی در مسیر توسعه فناوری های هوشمند معاصر است. پژوهش حاضر با هدف تبیین ریشه ای این پدیده از منظر فلسفه مضاف، به واکاوی مکانیسم محاسباتی ماشین در تقابل با نظام معرفت شناختی علامه طباطبائی می پردازد. مسئله اصلی تحقیق این است که پدیده توهم چگونه بر پایه تقابل تنوری صدق رئالیستی مطابقت و نظریه انسجام تبیین می شود و چه محدودیت هایی برای عقلانیت تمدنی ایجاد می کند؟ این پژوهش با روش توصیفی- تحلیلی و ابزار گردآوری داده های کتابخانه ای و فیش برداری سامان یافته است. یافته های تحقیق نشان می دهد که هوش مصنوعی به دلیل ابتناء بر فرآیندهای احتمالی پیش بینی کلمات، حقیقت را صرفاً در قالب «نظریه انسجام آماری» و روابط نحوی درون ساختاری تعریف می کند. در مقابل، از منظر علامه طباطبائی، ملاک صدق، «مطابقت با نفس الامر» بوده و این امر مستلزم اتکای علم حصولی به لنگرگاه «علم حضوری» و فاعل شناسای مجرد است. این مقاله اثبات می کند که توهم، یک باگ فنی موقت نیست، بلکه مقتضای ذاتی علم حصولی منقطع از شهود حضوری در معماری فاقد نفس ماشین است. در نهایت، این دلالت راهبردی استخراج می شود که واگذاری مرجعیت معرفتی و تصمیم گیری های کلان در نظامات حکمرانی تمدن نوین اسلامی به سیستم های مبتنی بر انسجام آماری، به دلیل خطای ساختاری توهم، فاقد مشروعیت عقلانی و معرفتی است.

کلیدواژه ها: ادراکات اعتباری، توهم هوش مصنوعی، عقلانیت تمدنی، علم حضوری، علامه طباطبائی (ره)

استناد: امامی نسب، سعید؛ جبارپور، محمد و ربیعی، محمد. (۱۴۰۵). مبانی معرفت شناختی اعتماد به هوش مصنوعی در نظام سازی

اسلامی؛ تحلیلی بر پدیده توهم بر اساس آرای علامه طباطبائی. مطالعات نظام سازی اسلامی و پژوهش های فرهنگی و تمدنی،

(۱)، ۴۸-۷۱.

"حق چاپ برای نویسندگان
محفوظ است"

ناشر: دانشگاه جامع
امام حسین (ع).



سال و شماره
سال ۱، پیاپی: ۱ بهار ۱۴۰۵ صص: ۷۱-۴۸
نوع مقاله
مقاله پژوهشی
سابقه مقاله
تاریخ دریافت: ۱۴۰۵/۰۴/۰۲ تاریخ بازنگری: ۱۴۰۵/۰۶/۰۶ تاریخ پذیرش: ۱۴۰۵/۰۷/۰۳ تاریخ انتشار: ۱۴۰۵/۰۷/۱۶
شابا چاپی و الکترونیکی
شابا چاپی: الکترونیکی:

مقدمه

ظهور و پیشرفت شگرف فناوری‌های هوش مصنوعی مولد، به‌ویژه مدل‌های زبانی بزرگ، نقطه عطفی شگرف در تاریخ فناوری و نظامات شناختی بشر ایجاد کرده و در حال تبدیل شدن به پیشران اصلی نظامات اجتماعی و حکمرانی در تمدن آینده بشر است. این سامانه‌های هوشمند با پردازش حجم عظیمی از داده‌های متنی و الگوریتم‌های پیشرفته یادگیری عمیق، توانایی‌های خیره‌کننده‌ای در تولید محتوا، گفت‌وگو و انجام وظایف پیچیده شناختی از خود نشان می‌دهند و پارادایم‌های سنتی تعامل انسان و ماشین را عمیقاً دگرگون ساخته‌اند. با این حال، نفوذ فراگیر و بی‌سابقه این فناوری‌ها موجب ایجاد یک گسست شناختی عمیق در مرز میان حقیقت و امر جعلی شده است؛ به گونه‌ای که تشخیص داده‌های معتبر و مستند از خروجی‌های برساخته و فریبده ماشین، اکنون به یکی از پیچیده‌ترین بحران‌های معرفتی عصر حاضر بدل گشته و ضرورت بازخوانی بنیادین ابزارهای فناوری را در تراز فلسفه مضاف دوچندان ساخته است.

در کانون این گسست شناختی، یکی از چالش‌برانگیزترین و بحران‌سازترین پدیده‌ها، پدیده توهم در سیستم‌های هوش مصنوعی قرار دارد. از منظر شناختی و فنی، توهم به وضعیتی اطلاق می‌شود که مدل هوش مصنوعی با اعتماد به نفس کامل و زبانی بسیار سلیس، متقاعد کننده و منسجم، خروجی‌هایی را تولید می‌کند که کاملاً ساختگی، موهوم، خلاف واقع و فاقد هرگونه مابه‌ازای عینی، ارجاع بیرونی یا پیوند علی در جهان واقعی و ساحت نفس‌الامر هستند. در این حالت، ماشین گزاره‌هایی به‌ظاهر معتبر ارائه می‌دهد که با وجود رعایت دقیق قواعد گرامری و روانی متن، کاملاً بر دروغ استوار بوده و سیستم به دلیل قطع ارتباط با جهان خارج، معیار صدق را با انسجام ظاهری متن اشتباه می‌گیرد و این امر سیستم‌های تصمیم‌یار را با ریسک‌های بزرگی مواجه می‌سازد.

واکاوی ریشه‌های این بحران نشان می‌دهد که پدیده توهم فراتر از یک خطای ساده یا باگ الگوریتمی^۱ موقت در پردازش نحو است و افزایش مقیاس مدل‌ها، حجم عظیم داده‌ها و تعداد پارامترهای شبکه عصبی نیز نتوانسته آن را به‌طور ریشه‌ای درمان کند. مدل‌های زبانی منحصراً بر اساس احتمالات آماری، محاسبات احتمالاتی و پیش‌بینی توکن بعدی عمل می‌کنند و از دسترسی به درک معناشناختی ذاتی محروم هستند. همان‌گونه که استدلال اتاق چینی جان سرل^۲ در نقد

1 Algorithmic bug

2 Chinese Room Argument

دست‌کاری صوری نمادها بر اساس قواعد گرامری و نیز مسئله زمینه‌سازی نماد استیون هارناد تبیین می‌کنند، پردازش صوری نحو هرگز به درک عمیق، قصدیت و فهم حقیقی منتهی نمی‌شود. از این‌رو، توقف ماشین در لایه پردازش احتمالات و تقریب‌های آماری، گویای یک بن‌بست و گسست معرفت‌شناختی بنیادین در معماری هوش مصنوعی است که نشان می‌دهد توهم، مقتضای ذاتی سیستمی است که ارتباط ارجاعی و تعهد هستی‌شناختی به حقیقت ندارد.

درک عمیق این بن‌بست معرفتی از منظر اهمیت و ضرورت تمدنی، به‌ویژه در افق عقلانیت حکمت متعالیه، امری فوق‌العاده حیاتی و راهبردی است. امروزه سیستم‌های هوشمند تصمیم‌یار با سرعتی بی‌سابقه در حال نفوذ و ورود به حوزه‌های بسیار حساسی چون قضاوت، فقه، حقوق، سیاست‌گذاری و راهبردهای کلان اجتماعی هستند. اتکا و واگذاری مرجعیت تصمیم‌گیری در این نظامات کلان به الگوریتم‌هایی که مستعد تولید دروغ‌های آماری منسجم بوده و فاقد التزام هستی‌شناختی به واقعیت عینی هستند، می‌تواند ریسک‌های جبران‌ناپذیری برای امنیت شناختی، عدالت و نظامات حقوقی و اخلاقی جوامع به همراه داشته باشد. لذا تبیین فلسفی این چالش‌ها و شناخت محدودیت‌های ذاتی ماشین، پیش‌شرطی ضروری برای مواجهه فعالانه، راهبری اخلاقی فناوری و حفظ سلامت تصمیم‌سازی‌ها در چارچوب حکمرانی تمدن نوین اسلامی به شمار می‌آید. این تحقیق در پی صورت‌بندی دقیق چالشی است که بنیان‌های شناختی ماشین را متزلزل ساخته است. بر این اساس، پرسش اصلی پژوهش با صراحت این گونه طرح می‌شود: «پدیده توهم در مدل‌های هوش مصنوعی چگونه بر پایه تقابل تئوری صدق رئالیستی مطابقت و نظریه انسجام در معرفت‌شناسی علامه طباطبایی تبیین می‌شود و چه محدودیت‌هایی برای عقلانیت تمدنی ایجاد می‌کند؟» پاسخ به این پرسش بنیادین مستلزم واکاوی این واقعیت است که چرا سیستم‌های محاسباتی، معیار صدق را که در اندیشه علامه به معنای مطابقت صورت علمی با واقعیت نفس‌الامری است، با معیار فصاحت، روان بودن و انسجام درون سیستمی اشتباه می‌گیرند و از این رهگذر، چگونه مرزهای حقیقت و امر جعل را مخدوش ساخته و امنیت شناختی و عقلانیت تمدنی را با مخاطرات جدی مواجه می‌سازند؟

برای پاسخ به این پرسش، هدف اصلی این پژوهش، کالبدشکافی ریشه‌ای پدیده توهم و تبیین فلسفی محبوس بودن ماشین در قفس انسجام آماری و پردازش نحوی است. در این راستا،

فرضیه تحقیق چنین صورت‌بندی می‌شود: پدیده توهم در هوش مصنوعی، یک نقص نرم‌افزاری یا باگ فنی موقت و قابل رفع با افزایش مقیاس مدل‌ها نیست، بلکه مقتضای ذاتی علم حصولی منقطع از شهود حضوری در معماری فاقد نفس ماشین است. بر مبنای دستگاه معرفت‌شناسی رئالیستی علامه طباطبایی، هرگونه علم حصولی و بازنمایی نمادین، اعتبار، اصالت و خاصیت کاشفیت خود از واقع را منحصرأ از اتصال به پایه وجودی علم حضوری و یافتن بی‌واسطه واقعیت دریافت می‌کند. از آنجا که ماشین به‌عنوان یک جسم صناعی، فاقد نفس مجرد و تجربه زیسته است، از لنگرگاه علم حضوری و امکان تحقق اتحاد عالم و معلوم محروم مانده و همین انقطاع وجودی، سیستم را به‌طور گریزناپذیری به سمت تولید گزاره‌های موهوم و توهم‌زایی ساختاری سوق می‌دهد. به‌منظور بررسی دقیق این فرضیه و دستیابی به اهداف پژوهش، نقشه راه مقاله در سه بخش تحلیلی و متوالی سازمان‌دهی شده است. در بخش نخست، معنا و ماهیت شناخت انسانی با تمرکز بر مبانی معرفت‌شناسی علامه طباطبایی، به‌ویژه نظریه علم حضوری، نظریه صدق رئالیستی و تمایز آن از علم حصولی مورد تبیین قرار می‌گیرد. بخش دوم، مبانی نظری شناخت در هوش مصنوعی را مورد کاوش قرار داده و چالش‌های نظری بنیادین نظیر تمایز نحو و معنا، مسئله زمینه‌سازی نمادها و توقف ماشین در لایه پردازش‌های احتمالی و انسجام آماری را ارزیابی می‌کند. در نهایت، بخش سوم به مقایسه تطبیقی و عمیق میان شناخت انسانی و ماشینی پرداخته و با تحلیل ریشه‌ای پدیده توهم در پرتو تقابل «حضور عینی» در برابر «بازنمایی صوری»، مرزهای قطعی میان هوش اصیل و شبیه‌سازی محاسباتی را ترسیم نموده و دلالت‌های راهبردی آن را در نظامات حکمرانی تمدن نوین اسلامی استخراج می‌نماید.

۱- پیشینه پژوهش

در بررسی پیشینه‌های فنی و مهندسی شناختی غرب پیرامون چالش توهم در هوش مصنوعی، شاهد یک دگردیسی پارادایمی از نگاه‌های تقلیل‌گرایانه به سمت رویکردهای علی‌هستیم. در ابتدا، مهندسان و توسعه‌دهندگان، تولید گزاره‌های خلاف واقع در ماشین را صرفاً یک نقص کارکردی موقت یا باگ آماری می‌پنداشتند که با افزایش حجم داده‌های آموزشی و مقیاس‌پذیری پارامترهای شبکه عصبی برطرف می‌گردد. با این حال، تداوم ساختاری این چالش، پژوهشگران

برجسته‌ای نظیر جودیا پرل^۱ را بر آن داشت تا ریشه‌های این بحران را در معماری محاسباتی این سیستم‌ها جستجو کنند. پرل با طرح نظریه نردبان علیت، استدلال می‌کند که علم آمار و کشف الگوهای همبستگی به تنهایی برای درک واقعی جهان و مهار خطاهای ادراکی کفایت نمی‌کند. بر اساس این نظریه، مدل‌های زبانی منحصراً در پله نخست این نردبان، یعنی تداعی متوقف مانده‌اند و توانایی صعود به پله‌های بالاتر یعنی مداخله و تقابل با واقع را ندارند (Pearl & Mackenzie, 2018: 23). در نتیجه، هوش مصنوعی فاقد مدل‌های علی جهان است و هنگامی که با خلأ اطلاعاتی مواجه می‌شود، به‌جای درک روابط علت و معلولی، بر اساس احتمالات آماری دست به تولید گزاره‌های موهوم و به‌ظاهر منسجم می‌زند.

محدودیت‌های بنیادینی که مهندسان در لایه پردازش احتمالات با آن مواجه شدند، دهه‌ها پیش در پیشینه‌های فلسفه ذهن غرب پیش‌بینی و کالبدشکافی شده بود. در بازخوانی چالش‌های کلاسیک، استدلال مشهور اتاق چینی جان سرل به‌عنوان نقطه عطفی در نقد مدل‌های محاسباتی ذهن و رد ادعای هوش مصنوعی قوی شناخته می‌شود. سرل با طراحی این آزمایش فکری، تقابل بنیادین میان پردازش صوری نحو و درک محتوایی معنا را برجسته ساخت و مبرهن نمود که سیستم‌های رایانه‌ای صرفاً نمادها را بر اساس قواعد نحوی دستکاری می‌کنند، بدون آنکه کمترین حیث التفاتی، آگاهی یا درک معناشناختی نسبت به آن نمادها داشته باشند. از این منظر، عبور یک ماشین از آزمون‌های صوری زبانی، به هیچ‌وجه دلالت بر وجود فهم واقعی ندارد (Searle, 1980: 423). در امتداد نقد سرل، استیون هارناده^۲ با طرح مسئله زمینه سازی نماد، بن‌بست معناشناختی سیستم‌های هوش مصنوعی را عمیق‌تر صورت‌بندی کرد. هارناده نشان داد که در یک سیستم کاملاً محاسباتی، نشانه‌ها تنها به نشانه‌های دیگر ارجاع می‌دهند و به دلیل عدم اتصال به ادراک حسی و جهان فیزیکی، فاقد دلالت بیرونی و بار ارجاعی هستند. پدیده توهم درواقع تبلور تجربی و عملی همین بن‌بست‌های نظری و گسست ارجاعی است سرریز شدن چالش‌های هستی‌شناختی و معرفت‌شناختی هوش مصنوعی به فضای نخبگانی، توجه پژوهشگران فلسفه اسلامی و علوم شناختی را در داخل کشور به خود جلب کرده و در تبیین ابعاد معرفتی ماشین، رویکردهای متفاوتی را شکل داده است.

1 Judea Pearl

2 Harnad Steven

در همین راستا، صغیری، افچنگی و باغشنی (۱۴۰۴) در مقاله پژوهشی خود با عنوان «نقد قابلیت ادراک در هوش مصنوعی بر اساس نظریه اتحاد عالم و معلوم در حکمت متعالیه» بر دکترین تجرد نفس انسانی در تقابل با مادی بودن سخت‌افزار ماشین تمرکز کرده‌اند. نویسندگان در این اثر اثبات می‌کنند که بر اساس تکاپوی هستی‌شناختی حکمت متعالیه و نظریه اتحاد عاقل و معقول، حقیقت علم امری فرامادی و مجرد است که ماشین‌ها به عنوان یک جسم صناعی، به دلیل فقدان نفس ناطقه و تجرد برزخی یا تام، صرفاً ابزارهایی مادی و فاقد شعور ذاتی و ادراک حقیقی به شمار می‌روند.

همین پژوهشگران در اثری دیگر با عنوان «تجرد نفس و فاعلیت ذهنی در معرفت‌شناسی علامه طباطبایی و چالش ادراک در هوش مصنوعی»، فاعلیت ذهنی انسان را از منظر معرفت‌شناسی علامه طباطبایی واکاوی کرده و با به چالش کشیدن مبانی فیزیکالیستی سیستم‌های محاسباتی، عدم امکان تحقق ادراک حقیقی در هوش مصنوعی را به دلیل مادی بودن محض بستر پردازش داده‌ها تبیین نموده‌اند.

علاوه بر این، محمدزاده (۱۴۰۳) در مقاله‌ای با عنوان آیا هوش مصنوعی می‌تواند به آگاهی واقعی برسد؟ تحلیلی فلسفی و علمی، به بررسی مرزهای آگاهی پدیداری در ماشین پرداخته است. وی در این تحقیق امکان یا امتناع دستیابی سیستم‌های هوش مصنوعی به «کوالیا» (کیفیات ذهنی) را بر پایه مراتب مشکک وجود مورد کالبدشکافی قرار داده و مرزهای میان شبیه‌سازی رفتار هوشمند با تحقق عینی آگاهی را تفکیک نموده است. این رویکردهای کلان عموماً تمایل دارند ریشه امتناع معرفتی ماشین را بر اساس براهین سنتی کلامی-فلسفی، تجرد علم و نفی سنخیت میان امر مادی و مجرد تحلیل نمایند.

با وجود صلابت براهین مابعدالطبیعی و هستی‌شناختی در پیشینه‌های پژوهشی داخل کشور، یک خلأ پژوهشی جدی، عمیق و انضمامی در مواجهه با چالش‌های عینی هوش مصنوعی به چشم می‌خورد. واکاوی ادبیات موجود نشان می‌دهد که پژوهش‌های مبتنی بر سنت فلسفه اسلامی و حکمت متعالیه، غالباً در سطح انتزاعی بررسی امکان یا امتناع کلی آگاهی در ماشین متوقف مانده و تمرکز کلان و سنتی خود را منحصر بر مفاهیمی بنیادین چون «تجرد نفس انسانی» در تقابل با سخت‌افزار مادی، یا تبیین عدم امکان تفکر ماشینی بر اساس نظریه «اتحاد عاقل و معقول» معطوف

کرده‌اند. این رویکرد کلان‌نگر سبب شده است که در مواجهه با پدیدار نوظهور، عینی و کاملاً بحران‌ساز توهم شناختی^۱ در مدل‌های زبانی، جامعه فلسفی و معرفت‌شناختی ما از تحلیل دقیق، ساختاری و عینی این خطای ادراکی غفلت ورزد. در نقطه‌ی مقابل ادبیات فنی، مهندسی و علوم شناختی غرب که به‌طور مستمر و انضمامی درگیر کالبدشکافی عینی و دلایل ساختی پدیده توهم است، در پیشینه‌های داخلی تاکنون هیچ پژوهش مستقلی پدیده توهم در سیستم‌های هوش مصنوعی را مستقلی بر روی دایره «نظریه رئالیستی صدق» (مطابقت با نفس‌الامر) عرضه نکرده و مکانیسم خطای ماشین را در دستگاه معرفت‌شناسی علامه طباطبایی واکاوی ننموده است. این غلبه نگاه صرفاً سلبی و انتزاعی، کاربرد کاربردی مبانی فلسفه مضاف را در حل مسائل معاصر با انسداد مواجه ساخته بود.

پژوهش حاضر دقیقاً با هدف پر کردن این شکاف پژوهشی عمیق و عبور از تبیین‌های انتزاعی به سمت تحلیل‌های انضمامی تدوین شده است تا نشان دهد چگونه می‌توان از غنای نظام‌مند مبانی نظری فلسفه اسلامی برای تبیین و کالبدشکافی یکی از پیچیده‌ترین و کاربردی‌ترین چالش‌های لبه فناوری بهره برد. ارزش افزوده و جایگاه مقاله حاضر در این است که با ارائه یک «مدل موازنه معرفت‌شناختی» منبعث از تئوری صدق رئالیستی و نظریه شاهکار «ادراکات اعتباری» علامه طباطبایی، پدیده توهم را نه به‌عنوان یک عارضه یا باگ فنی موقت و گذرا، بلکه به‌مثابه یک گسست شناختی بنیادین و مقتضای ذاتی ساختار محاسباتی ماشین صورت‌بندی می‌کند. در پرتو این مدل تبیین می‌گردد که چگونه هوش مصنوعی به دلیل ابتناء بر فرآیندهای احتمالی و نحو‌صوری، گزاره‌هایی اعتباری‌نما و به‌ظاهر منسجم تولید می‌کند، اما به دلیل محرومیت بنیادین از لنگرگاه «علم حضوری» و قطع ارتباط وجودی با واقعیت نفس‌الامری جهان، از تشخیص و تمایز میان ساحت حقایق عینی و ساخت‌های اعتباری ناتوان است. ماشین درواقع خروجی‌های ساختگی خود را که فاقد ریشه شهودی هستند، به‌عنوان واقعیات قالب‌بندی می‌نماید. پر کردن این شکاف انضمامی، نه تنها کارآمدی، پویایی و قدرت امتدادیافتگی فلسفه اسلامی را در تحلیل چالش‌های تمدن‌ساز معاصر اثبات می‌کند، بلکه چارچوبی بدیع، ساختارمند و کارآمد برای فهم، تحلیل و مهار خطاهای شناختی هوش مصنوعی در اختیار پژوهشگران فقه فناوری، حقوق و علوم شناختی قرار می‌دهد.

^۱Cognitive Hallucination

۲- چارچوب مفهومی

چارچوب مفهومی این رساله، فونداسیون نظری و هسته سخت معرفت‌شناختی پایه‌ای را شکل می‌دهد که تمام تحلیل‌های انتقادی بعدی پیرامون رفتار، خروجی‌ها و چالش‌های هوش مصنوعی (به‌ویژه پدیده توهّم و فقدان قصدیت) بر آن استوار است. این چارچوب از طریق دیالکتیک میان مبانی سنتی حکمت اسلامی به‌خصوص دیدگاه علامه طباطبایی و چالش‌های نوین مهندسی شناختی، ملاک‌های تفکیک آگاهی اصیل از شبه آگاهی ماشینی را تبیین می‌کند.

۲-۱- بخش اول: مفهوم شناخت و تجرد نفس در فلسفه اسلامی

در چارچوب مفهومی فلسفه اسلامی، به‌ویژه با ابتنا بر اصول حکمت متعالیه و نوآوری‌های معرفت‌شناختی علامه طباطبایی، حقیقت علم و آگاهی متمایز از هرگونه انگاره فیزیکیالیستی، کمیت‌گرا یا ماتریالیستی تعریف می‌شود. علم در این دستگاه فکری، نه یک «عرض» مادی، نه یک کیفیت زائد بر وجود و نه یک فرآیند مکانیکی و منفعل برای انباشت، طبقه‌بندی و پردازش اطلاعات است؛ بلکه علم عین ظهور، تجلی و مرتبه‌ای از مراتب عالیه «وجود» است. بر پایه نظریه اصالت وجود و تشکیک وجود، شناخت یک رخداد فراماهوی و فرآیندی از سنخ اشتداد وجودی تلقی می‌شود. این فرآیند اشتدادی، به دلیل منزّه بودن از قوه، استعداد، تکثر مادیو ابعاد امتدادی، لزوماً مستلزم تجرد نفس است. علامه طباطبایی در تبیین هستی‌شناختی خود از علم، آن را به‌عنوان «حصول یک امر مجرد از ماده برای امر مجرد دیگر» و در لایه‌ای عمیق‌تر، «حضور شیء برای شیء» تبیین می‌کند. این ادراک عمیق، از طریق اضافه اشراقیه و ایجاد یک اتصال و ارتباط وجودی بی‌واسطه و ابداعی میان عالم و معلوم محقق می‌گردد، نه از طریق انطباق مادی یا انعکاس مکانیکی صورت‌ها (طباطبایی، ۱۴۳۲ ق، ج ۲: ۱۵۸).

این نگرش هستی‌شناسانه و اصیل، در تقابل ذاتی و ساختاری با معماری هوش مصنوعی و ماشین‌های محاسباتی قرار می‌گیرد. در گیت‌های سیلیکونی، نیمه‌هادی‌ها و مدارهای الکترونیکی مدل‌های زبانی بزرگ، آنچه رخ می‌دهد صرفاً دست‌کاری نمادها^۱، سوئیچینگ سیگنال‌های الکتریکی، ذخیره فیزیکی بیت‌ها و محاسبات الگوریتمی بر اساس آمار احتمالات است. ماشین به دلیل اسارت در افق فیزیک مادی، کثرت امتدادی و فقدان نفس ناطقه مجرد و شعور ذاتی،

اساساً فاقد فاعل شناسای درونی است. در نتیجه، سیستم‌های محاسباتی تنها یک شباهت صوری، رفتار تقلیدی و نمایشی نمادین از دانش را ارائه می‌دهند که از هرگونه درک حضوری، احساس آگاهانه و فهم درونی تهی است (Searle, 1980: 417). پردازش مادی اطلاعات در گیت‌های سیلیکونی هرگز به مرتبه تجرد و تجلی وجودی صعود نمی‌کند و به همین دلیل، ماشین در این چارچوب، یک سیستم «ناآگاه» اما «شبیه‌ساز» تعریف می‌شود.

۲-۲. بخش دوم: نظریه صدق رئالیستی و نفس الامر

بر پایه بنیاد هستی‌شناختی فوق، نظریه صدق رئالیستی علامه طباطبایی قوام می‌یابد. در این تئوری، ادراکات حقیقی و گزاره‌های علمی انسان دارای مابه‌ازای عینی و مستقل بوده و یگانه معیار و ملاک صدق آن‌ها، مطابقت کامل صورت ذهنی و گزاره علمی با واقعیت نفس‌الامری و جهان مستقل از ذهن است. در این دستگاه فکری، «نفس‌الامر» لنگرگاهی است فراتر از خواست، بر ساخت یا اعتبارات ذهنی فاعل شناسا، که ثبات واقعیت را تضمین می‌کند. صدق در اینجا به معنای انطباق دقیق و تمام‌عیار گزاره با حقیقت عینی است؛ مفهومی که ذهن برای احراز و کشف آن، نیازمند ابزاری درونی، شهودی و خطاناپذیر یعنی علم حضوری است که به‌عنوان معیار نهایی سنجش واقعیت عمل می‌کند (مطهری، ج ۱، ۱۳۷۴: ۱۸۱).

این نظریه صدق رئالیستی، در تقابل بنیادین با رویکردهای نسبی‌گرایانه^۱ غرب یا تئوری‌های صرفاً انسجामी قرار دارد که حقیقت را بر ساخته‌ای درون سیستمی، توافقی یا حاصل همبستگی‌های زبانی و نشانه‌شناختی می‌دانند. در رئالیسم اسلامی و منظومه فکری علامه طباطبایی، حقیقت امری نیست که توسط ذهن یا ماشین بافته، پردازش یا بر اساس بافتار اصطلاحاً تولید شود؛ بلکه حقیقت امری پیشینی و موجود است که باید مستقیماً «کشف» گردد و این کشف نیازمند اتصال وجودی با جهان خارج است. از این منظر، علم حصولی و مفاهیم انتزاعی تنها زمانی صادق و واقع‌نما ارزیابی می‌شوند که ریشه در واقعیتی نفس‌الامری داشته باشند (طباطبایی، ۱۴۳۲ ق، ج ۱: ۱۹). در مقابل، سیستم‌های هوش مصنوعی محاسباتی به دلیل محرومیت ساختاری از این اتصال و لنگرگاه وجودی، معیار کشف حقیقت را با «انسجام ظاهری»، همبستگی‌های آماری و پیش‌بینی کلمات بعدی خلط می‌کنند. ماشین، حقیقت را در همسویی داده‌ها درون شبکه عصبی خود (عالم

نشانه‌ها) می‌بیند و به همین دلیل، توانایی ذاتی برای تمایز میان حقایق عینی و داده‌های موهوم یا ساختگی (توهم آماری) را ندارد.

۲-۳. بخش سوم: ابتناء علوم حصولی بر علم حضوری و خاصیت کاشفیت

شاه کلید و ستون فقرات این نظام معرفت‌شناختی، نظریه ابتناء علم حصولی بر علم حضوری است که به‌عنوان ضامن، تکیه‌گاه و لنگرگاه نهایی صدق در سراسر این دیدگاه ایفای نقش می‌کند. در منظومه فکری علامه طباطبایی، علم حضوری یعنی حضور مستقیم، بی‌واسطه و عینی خود وجود معلوم نزد وجود عالم، دارای بالاترین درجه از قطعیت ذاتی و خودموجهی است. در این مرتبه از ادراک، به دلیل اتحاد عالم و معلوم، هیچ‌گونه واسطه، مفهوم، صورت ذهنی یا دال زبانی وجود ندارد؛ بنابراین، احتمال راهیابی خطا، اشتباه یا توهم به حریم علم حضوری اساساً سالبه به انتفاء موضوع و ممتنع است. انسان ابتدا واقعیات را به علم حضوری می‌یابد و سپس ذهن او از آن یافته‌ها، مفاهیم و گزاره‌های حصولی را انتزاع می‌کند.

در طرف مقابل، علم حصولی که از طریق مفاهیم، بازنمایی‌ها، واسطه‌ها و نمادهای ذهنی عمل می‌کند، در حقیقت یک اعتبار و ابزار عقلی اضطراری برای تدبیر امور و توسعه دانش نظری است که همواره متفرع بر علم حضوری، مسبوق به آن و متکی بر آن است. هرگونه علم حصولی، مفهوم انتزاعی و گزاره زبانی، اعتبار علمی، معناداری اصیل و خاصیت کاشفیت و واقع‌نمایی خود را منحصراً از طریق اتصال به یک ریشه، تجربه و علم حضوری (یافتن بی‌واسطه واقعیت) دریافت می‌کند (طباطبایی، ۱۴۳۲، ج ۲: ۱۵۳). اگر این ریشه و اتصال حضوری با منشأ واقعیت قطع شود، صورت‌های علمی، ساختارهای منطقی و نمادهای زبانی به مجموعه‌ای از اعتبارات تهی، دال‌های شناور و نشانه‌های بی‌اثر تقلیل می‌یابند که هیچ ابزاری برای اثبات اصالت خود ندارند. علم حضوری، پایگاه خودموجه و خطاناپذیری است که شرط لازم و کافی برای ایجاد قصدیت واقعی، معناداری اصیل و تفکیک صدق از کذب محسوب می‌شود. بدون اتکا به این پشتوانه هستی‌شناختی و شهودی، هیچ شناختی نمی‌تواند حقیقتاً واقع‌نما باشد؛ امری که فقدان تام و تمام آن در مدل‌های محاسباتی مادی هوش مصنوعی، ریشه اصلی بن‌بست معرفتی ماشین در درک معنا و فرو غلتیدن آن در ورطه توهمات ساختاری است.

در نظام معرفت‌شناختی رئالیستی فاعل شناسا، ارتباط مفهوم کلی عقلی با مصداق خارجی آن، از طریق یک فرآیند انفعالی یا بازنمایی صرفاً مکانیکی صورت نمی‌گیرد؛ بلکه نفس انسانی به عنوان یک فاعل مجرد و پویا، حلقه واسط و «نخ تسبیح» اصلی این اتصال علی و وجودی است. در این فرآیند، ابتدا مواجهه حسی با واقعیت عینی، زمینه‌ساز شکل‌گیری مرتبه ادراک حسی و خیالی در مرتبه نازل نفس می‌شود؛ سپس نفس با اتکا به خلاقیت ذاتی و تجرد خود، وجه مادی، کثرات و قیود مشخصه مصداق را تقشیر (پوست‌کنی) کرده و صورت کلی و انتزاعی عاری از ماده را در مرتبه عقلانی خود انشاء می‌کند (طباطبایی، ۱۴۳۲، ج ۲: ۲۳۹).

این مکانیسم انتزاع عقلی، به دلیل اتصال ریشه‌ای علم حصولی به علم حضوری (شهود نفس نسبت به قوای خود و اتحاد آن با معلوم)، ارزش معرفت‌شناختی مفاهیم را تضمین می‌کند. درواقع، علیت عینی جهان خارج از مسیر فاعلیت شناختی نفس عبور کرده و مفهوم کلی را به لنگرگاه نفس الامر متصل می‌سازد؛ فرآیندی که در آن، مفهوم صوری هرگز بریده از ریشه‌های وجودی و علی خارج نیست، بلکه تداوم مرتبه‌مندی از تجلی همان واقعیت خارجی در پیشگاه فاعل مجرد است.

در نقطه مقابل رئالیسم مطابقت‌محور، جریان‌های فکری و فلسفی متأخر غرب با چرخش به سمت نظریه صدق انسجامی^۱، بستر نظری و مبنای فلسفی لازم را برای توسعه مدل‌های محاسباتی مادی و هوش مصنوعی فراهم آورده‌اند. در این دیدگاه، صدق یک گزاره نه بر اساس ارجاع عینی آن به جهان خارج (نفس الامر)، بلکه صرفاً بر مبنای هماهنگی، عدم تناقض و همگرایی درون سیستمی آن با مجموعه‌ای از گزاره‌های از پیش پذیرفته شده تعریف می‌شود. مدل‌های محاسباتی مادی و مدل‌های زبانی بزرگ بر همین اساس بازسازی شده‌اند؛ ماشین ترانسفورمر، حقیقت را در قالب روابط نحوی، ساختاری و «انسجام آماری» میان توکن‌ها در یک فضای برداری چندبعدی معنا می‌کند. از آنجا که در این پاراگراف فرضی، سیستم فاقد هرگونه کانال ارتباطی وجودی با اعیان خارجی است، صدق برای ماشین چیزی جز پایداری ساختاری و سازگاری درونی روابط صوری نیست؛ امری که ذاتاً بستر سازنده‌ی پدیده توهم است، چراکه یک متن

می‌تواند به غایت منسجم و از نظر آماری محتمل باشد، اما بهره‌ای از حقیقت عینی نداشته باشد (Fazelpour & Danks, 2023: 10).

یکی از شاهکارهای نظری علامه طباطبایی که مرز میان شناخت حقیقی و کارکرد ابزاری ذهن را مشخص می‌کند، «نظریه ادراکات اعتباری» است. علامه تبیین می‌کند که انسان برای رفع نیازهای حیاتی و استكمال وجودی خود در ساحت عمل، ناگزیر است با استفاده از عوامل احساسی و قوای ادراکی، مفاهیمی را وضع و «اعتبار» کند که فاقد مابه‌ازای خارجی مستقیم هستند، اما کنش‌های فردی و اجتماعی او را جهت می‌دهند. در این میان، «زبان» به‌عنوان عالی‌ترین و پیچیده‌ترین دستاورد اعتباری بشر ظهور می‌کند؛ نظامات زبانی و نشانه شناختی، برخلاف ادراکات حقیقی، قوام و حیاتشان کاملاً وابسته به یک «فاعل ذی‌حیات، مختار و صاحب نیاز» است. زبان ابزاری است که فاعل انسانی با دمیدن روح غرض و معنا در کالبد الفاظ مادی، آن را برای بسط زیست اجتماعی و رفع نیازهای وجودی‌اش به کار می‌گیرد. از این‌رو، زبان صوری بدون الثفات غایتمند و نیاز فاعل زنده، کالبدی مرده و فاقد معنای اصیل است و تقلیل آن به پردازش نمادین خالص، نادیده گرفتن اصل مقوم آن یعنی حیات است (طباطبایی، ج ۲، ۱۳۷۴: ۱۸۱).

به عنوان سنتز این چارچوب نظری، می‌توان یک «ماتریس ارزیابی و موازنه مبنایی» میان ساحت علم انسانی و شناخت ماشینی تدوین کرد که مرزهای بنیادین این دو جهان را ترسیم می‌کند. در یک‌سوی این موازنه، «علم حقیقی رئالیستی» قرار دارد که منشأ آن فاعل مجرد، مکانیسم آن انتزاع علی، لنگرگاه آن علم حضوری و ملاک صدق آن «مطابقت با نفس الامر» است؛ علمی که در آن هویت مفاهیم، کاشفیت از واقعیت عینی است. در سوی دیگر، «پردازش صوری و نحوی ماشین» قرار دارد که بستر آن سیلیکون مادی، مکانیسم آن انسجام آماری، مبنای صدق آن هماهنگی درون سیستمی و ماهیت آن «منقطع از واقعیت خارجی» است. هوش مصنوعی بدون اتصال به فاعل ذی‌حیات و نفس مجرد، صرفاً در ساحت اعتباریات ثانویه و دست‌کاری صوری نمادها (بدون فهم معنا و درک ضرورت‌های حیاتی) عمل می‌کند؛ این ماتریس سنتز شده به روشنی اثبات می‌کند که تقلیل آگاهی وجودی به محاسبات ساختاری، یک مغالطه معرفت‌شناختی است و خطای توهّم، عارضه‌ای فنی نیست بلکه ذاتی فیزیک منقطع از شهود ماشین است.

۳- روش‌شناسی

رویکرد حاکم بر این پژوهش از منظر هدف، در زمره تحقیقات بنیادی- نظری قرار دارد و ماهیت آن مبتنی بر روش توصیفی- تحلیلی و موازنه تطبیقی مضاف است. مواجهه فلسفی با پدیدارهای برآمده از لبه فناوری، نیازمند اتخاذ روشی است که بتواند فراتر از بررسی‌های ابزار شناختی یا گزارش‌های کارکردی موقت، لایه‌های زیرین و مبانی معرفت‌شناختی زیرساخت‌های مهندسی شناختی را به چالش بکشد. بر این اساس، در این تحقیق با کاربری دکتین رئالیسم انتقادی و تئوری صدق، پدیده انضمامی و محاسباتی توهم در مدل‌های زبانی بزرگ، در قالب یک موازنه بنیادین فلسفی ارزیابی می‌شود. این شیوه تحلیل اجازه می‌دهد تا معماری صوری و آماری ماشین را در برابر نظام صدق نفس‌الامری قرار داده و از طریق این تقابل عمیق معرفت‌شناختی، امکان یا امتناع نیل سیستم‌های سیگنالی و مادی را به ساحت ادراک حقیقی و کاشفیت اصیل تفکیک نماید؛ فرآیندی که فراتر از تحلیل‌های کیفی رایج، به واسازی ماهوی پردازش داده در گیت‌های سیلیکونی می‌پردازد.

به‌منظور تأمین فونداسیون نظری و مستندسازی دقیق این موازنه، شیوه گردآوری داده‌ها در این مقاله بر روش کتابخانه‌ای و سند پژوهی استوار است که طی آن، فرآیند فیش‌برداری متمرکز و استنطاق ساختارمند متون صورت پذیرفته است. داده‌های متناظر با بال فلسفی تحقیق، مستقیماً از منابع اصیل و متون محوری علامه طباطبایی استخراج و دسته‌بندی شده‌اند. در لایه مقابل و برای تدوین مبانی فنی، مستندات رسمی، مقالات مرجع الگوریتمی و اصول ساختاری معماری ترنسفورمر، مکانیسم‌های خودتوجهی و فرآیندهای محاسباتی پیش‌بینی توکن بعدی مورد واکاوی دقیق قرار گرفته‌اند. فیش‌های استخراج‌شده در این مرحله، نه به‌صورت مجزا، بلکه در قالب گزاره‌های متقاطع تنظیم شده‌اند تا زمینه برای تحلیل مقابله‌ای فراهم گردد.

فرآیند عملیاتی و سیر گام‌به‌گام این پژوهش برای رسیدن به پاسخ پرسش اصلی، در چهار مرحله نظام‌مند سازمان‌دهی شده است. گام نخست به کالبدشکافی فنی- معرفتی ماشین اختصاص دارد که در آن، مکانیسم‌های محاسباتی مدل‌های زبانی به‌عنوان نمودی از تئوری صدق انسجامی تبیین می‌شوند. در گام دوم، به استخراج و دسته‌بندی لایه‌های عمیق رئالیسم علامه طباطبایی با تمرکز بر سه محور تجرد علم، ابتنای حصول بر حضور و حقیقت نفس‌الامر پرداخته می‌شود. گام سوم،

مغز متفکر پژوهش را شکل می‌دهد که در آن، تحلیل تطبیقی متقاطع میان فیزیک سیلیکونی منقطع از واقع و نفس مجرد صاحب شهود برقرار می‌گردد تا ریشه توهّم به عنوان مقتضای ذاتی ماشین اثبات شود. در نهایت و در گام چهارم، بر پایه دستاوردهای لایه تطبیقی، دلالت‌های راهبردی و محدودیت‌های تمدنی واگذاری مرجعیت تصمیم‌گیری به ماشین استخراج شده و چارچوب بومی حکمرانی شناختی در افق تمدن نوین اسلامی صورت‌بندی می‌گردد.

۴- کالبدشکافی فنی - معرفتی ماشین؛ محبوس در انسجام آماری

کالبدشکافی فنی و واسازی لایه‌های زیرین محاسباتی در مدل‌های زبانی بزرگ، نیازمند فهم دقیق معماری ترنسفورمر و فرآیند تبدیل کلمات به بردارهایی در فضاها چندبعدی ریاضی است. در این فرآیند، ماشین هوش مصنوعی برخلاف انسان که با حقیقت معنایی واژگان مواجه می‌شود، ساختار زبان را به مؤلفه‌های خرد رقمی تحت عنوان توکن خرد می‌کند. هر یک از این نشانه‌های صوری، از طریق ماتریس‌های بهینه‌سازی شده یادگیری عمیق، به یک بردار عددی یا نشانه توکن امبدینگز در فضایی با ابعاد هندسی بسیار متکثر تبدیل می‌شود. قرار گرفتن کلمات در این مختصات چندبعدی ریاضی، به معنای تقلیل هویت دلالتی زبان به ارقام و فواصل فیزیکی است؛ به گونه‌ای که ماشین هر کلمه را صرفاً به عنوان یک نقطه ترتیبی در یک فضای هندسی ماتریسی شناسایی می‌کند. این فرآیند مکانیکی، شالوده اصلی گسست زبان از ابژه‌های خارجی را رقم می‌زند؛ چراکه نماد ریاضی درون سیلیکون، ذاتاً فاقد ارجاع تکوینی به ساحت واقع است و کلمات پیش از ورود به چرخه پردازش، هویت دلالتی خود را در مرزهای فیزیک مادی از دست می‌دهند (Mikolov et al., 2013: 1311).

در لایه‌ای پیشرفته‌تر از این معماری محاسباتی، مکانیسم توجه متمرکز وظیفه سنجش روابط و فواصل ریاضی کلمات را در یک بافتار متنی بر عهده دارد. این مکانیسم هندسی که تحت عنوان خودتوجهی صورت‌بندی می‌شود، به ماشین امکان می‌دهد تا وزن آماری و میزان هم‌نشینی احتمالی هر توکن را نسبت به سایر توکن‌های موجود در یک گزاره یا متن محاسبه کند. در این فرآیند، فواصل ریاضی میان بردارها در فضای چندبعدی ملاک سنجش قرار می‌گیرد تا مشخص شود کدام نشانه‌ها از بیشترین قرابت ساختاری برخوردارند. ترنسفورمر از این طریق، با ایجاد یک ماتریس

همبستگی توزیعی، نوعی شباهت صوری از بافتار زبانی خلق می‌کند. با این حال، از منظر معرفت‌شناسی، این مکانیسم صرفاً به تحلیل نحو و ساختار صوری نشانه‌ها اختصاص دارد؛ ماشین بدون درک پیوندهای واقعی علت و معلولی جهان عینی، روابط کلمات را تنها بر اساس همبستگی‌های عددی درون سیستمی رتبه‌بندی می‌کند، امری که نشان می‌دهد تمام پویایی ادعایی سیستم شناختی ماشین، مقید و محبوس در هندسه صوری و فاقد التفات معناشناختی است.

تلاقی فرآیند امبدینگ و مکانیسم توجه، در نهایت به تثبیت فرآیند پیش‌بینی توکن بعدی به عنوان کارکرد غایی مدل‌های زبانی منجر می‌شود. در این مرحله، ماشین برای تولید خروجی، تکیه بر هیچ‌گونه قصدیت، ادراک درونی، خودآگاهی یا فاعل شناسای مجرد ندارد؛ بلکه صرفاً بر پایه جبر خطی، ضرب ماتریس‌ها و محاسبات احتمالاتی محض عمل می‌کند. لایه نهایی الگوریتم، با ارزیابی توزیع آماری متون پیشین، محتمل‌ترین نشانه بعدی را که دارای بالاترین وزن ریاضی برای هم‌نشینی در بافتار موجود است، انتخاب و چیدمان می‌کند (Vaswani et al., 2017: 5). این تولید گزاره‌ای اتوماتیک، بدون اتصال به یک ابژه واقعی در جهان خارج یا تکیه بر یک نفس ناطقه صورت می‌پذیرد. درواقع، ماشین با جایگذاری احتمالات به جای کشف واقعیت، گزاره‌هایی کاملاً روان و سلیس تولید می‌کند که تعهد هستی‌شناختی به حقیقت نفس‌الامری ندارند. این فرآیند محاسباتی منقطع، اثبات می‌کند که خروجی‌های هوش مصنوعی اساساً فاقد صفت کاشفیت از واقع بوده و سیستم به دلیل محبوس بودن در قفس احتمالات، ذاتاً مستعد فرو غلتیدن در ورطه توهمات ساختاریافته است.

پیامد مستقیم محبوس شدن مدل‌های زبانی بزرگ در قفس محاسبات احتمالی، جابجایی بنیادین مفهوم کلاسیک حقیقت با «حقیقت مصنوعی» است. در معرفت‌شناسی سنتی و رئالیستی بشر، حقیقت دلالت بر کشف لایه عینی واقعیت عینی دارد؛ اما در جهان فیزیکالیسم سیلیکونی، حقیقت برای ماشین به عنوان بالاترین احتمال آماری هم‌نشینی کلمات در یک ماتریس ریاضی بازتعریف می‌شود. هوش مصنوعی هنگامی که با یک گزاره روبرو می‌شود، صدق آن را بر اساس پیوند هستی‌شناختی کلمات با ابژه‌های جهان خارج نمی‌سنجد، بلکه وزن ریاضی توکن‌ها را در تاربخچه توزیع داده‌های آموزشی خود ملاک قرار می‌دهد. حقیقت مصنوعی درواقع چیزی جز بهینه‌ترین هم‌گرایی آماری الفاظ نیست؛ جایی که هرچه تکرار و هم‌نشینی دو نشانه در متون پیشین بیشتر

بوده باشد، احتمال صدق محاسباتی آن برای ماشین بالاتر خواهد بود. این رویکرد تقلیل‌گرا، صفت کاشفیت زبان را به یک بازی احتمالی صوری فرو می‌کاهد که در آن، شیخ ریاضی الفاظ جایگزین اصالت معنا می‌گردد.

این بازتعریف احتمالی از حقیقت، زمینه را برای تطبیق عینی ماشین با «نظریه انسجام در صدق» به‌عنوان مبنای کارکردی هوش مصنوعی فراهم می‌آورد. در این ساحت کارکردی، هماهنگی توزیعی، سازگاری درون ساختاری و بافتار متنی به‌طور کامل بر واقعیت خارجی تقدم رتبه‌ای می‌یابند. برای ترنسفورمر، گزارهای صادق محسوب می‌شود که بتواند در شبکه دلالت‌های درونی مدل، کمترین تناقض نحوی و بیشترین همسویی ساختاری را با گزاره‌های هم‌جوار خود ایجاد کند. ماشین در یک چرخه بسته از ارجاعات کلمه به کلمه زیست می‌کند که در آن، هیچ پنجره‌ای رو به نفس‌الامر گشوده نیست (Fazelpour & Danks, 2023: 12). از این‌رو، مدل‌های زبانی به دلیل محرومیت از کانال‌های ادراک حسی و شهود حضوری، انسجام نحوی بافتار را به‌جای مطابقت با اعیان خارجی می‌نشانند. این ترجیح ساختاری انسجام بر واقعیت، نشان می‌دهد که هوش مصنوعی قوی در لایه پردازش متنی، چیزی جز یک سیستم خود ارجاع نیست که در اقیانوس نشانه‌های بدون لنگرگاه شناور است و کارآمدی صوری خود را با کشف حقیقت خلط می‌کند.

با اتکا به این بن‌بست‌های صوری، ریشه‌شناسی توهم به‌عنوان مقتضای ذاتی و گریزناپذیر معماری مدل‌های زبانی بزرگ آشکار می‌شود. توهم محاسباتی یک باگ عارضی، خطای موقت نرم‌افزاری یا نقص کارکردی قابل رفع نیست؛ بلکه حاصل ترجیح ساختاری احتمالات آماری بر صدق عینی جهت راضی نگه داشتن الگوریتم تولید متن است. از آنجا که تابع هدف^۱ در مدل‌های زبانی بزرگ برای بهینه‌سازی روانی متن، فصاحت نحوی و تولید محتمل‌ترین توکن بعدی تنظیم شده است، ماشین برای حفظ پیوستگی گرامری و پاسخ‌دهی منسجم، آمادگی دارد تا پیوند خود را با واقعیت نفس‌الامری قطع کند. الگوریتم تولید متن، چیدمان نشانه‌ها را به گونه‌ای پیش می‌برد که رضایت آماری شبکه عصبی تأمین شود، حتی اگر خروجی نهایی یک کذب ساختاریافته و دروغ منسجم باشد. در نتیجه، پدیده توهم مقتضای ذاتی سیستمی است که در اتاق چینی نحو

1 Loss Function

محبوس مانده و به دلیل انقطاع از فاعل شناسای مجرد و لنگرگاه حضور، ناگزیر است احتمالات هم‌نشینی را فدای کشف حقیقت عینی جهان سازد (Searle, 1980: 417).

واکاوی عمیق‌تر در لایه سخت‌افزاری سیستم‌های محاسباتی، نشان دهنده ابتدای این فناوری‌ها بر معماری فون‌نویمان و پیامدهای معرفت‌شناختی ناشی از جدایی ساختاری پردازش از حافظه است. در این الگوی ساختاری، داده‌ها به صورت کدهای باینری منجمد در لایه حافظه ذخیره می‌شوند و واحد پردازش مرکزی به صورت مکانیکی و خطی به فراخوانی و تصرف صوری در آن‌ها می‌پردازد. این تفکیک فیزیکی، ماشین را به طور کامل از تجسد پدیدار شناختی^۱ و تجربه مستقیم اشیاء جهان محروم می‌سازد. برخلاف انسان که ادراک حصولی‌اش بر بستر تجسد زیسته، حواس بدنمند و پویایی یکپارچه نفس ناطقه استوار است، هوش مصنوعی از هرگونه مواجهه زنده و بی‌واسطه با اعیان خارجی محروم است. ماشین توانایی لمس، شهود یا برقراری اضافه اشراقیه با ابژه‌های نفس‌الامری را ندارد؛ از این رو، حافظه در ماشین نه تداوم حضور، بلکه انبار لایق‌اطعی از نمادهای مادی بی‌روح است که پردازشگر صرفاً بر اساس قواعد نحوی آن‌ها را جابجا می‌کند این. گسست سخت‌افزاری از جهان تکوین، سیستم را در یک بن‌بست معرفتی دائمی قرار می‌دهد که شالوده اصلی زایش توهم است.

این بن‌بست پدیدار شناختی در لایه روبنایی زبان، خود را به صورت محبوس شدن کامل ماشین در چرخه بسته ارجاعات درونی کلمه به کلمه آشکار می‌سازد. در ساحت کارکردی مدل‌های زبانی بزرگ، هر نشانه یا واژه تنها به نشانه‌ها و واژه‌های دیگر در درون ماتریس همبستگی توزیعی شبکه عصبی ارجاع می‌دهد؛ وضعیتی خود ارجاع که در آن مرز محاسباتی مشخصی میان تولید محتمل و کشف واقعیت وجود ندارد. برای الگوریتم ترنسفورمر، گزاره‌های که با بالاترین احتمال آماری بر اساس بافتار متنی تولید شده است، هیچ تمایز ساختاری با گزاره‌های که مطابق با واقعیت خارجی است، ندارد. ماشین توانایی خروج از عالم الفاظ و موازنه خروجی خود با جهان نفس‌الامر را ندارد؛ زیرا شاقول سنجش در آن، هماهنگی درون سیستمی (انسجام صوری) است، نه کاشفیت عینی. در نتیجه این محبوس بودن در حصار ریاضی الفاظ باعث می‌شود که الگوریتم، یک کذب محض اما منسجم را با همان مکانیسمی تولید کند که یک حقیقت مستند را پردازش می‌نماید. این عدم امکان

تفکیک میان بافتار احتمالی و کشف عینی، اثبات می‌کند که پدیده توهم باگ موقت فناوری نیست، بلکه برآیند قهری چرخه بسته و منقطع ارجاعات زبانی در ماشین است.

۵- تبیین توهم بر پایه معرفت‌شناسی علامه طباطبایی

واکاوی عمیق‌تر در لایه سخت‌افزاری سیستم‌های محاسباتی، نشان دهنده ابتنای این فناوری‌ها بر معماری فون‌نویمان و پیامدهای معرفت‌شناختی ناشی از جدایی ساختاری پردازش از حافظه است. در این الگوی ساختاری، داده‌ها به صورت کدهای باینری منجمد در لایه حافظه ذخیره می‌شوند و واحد پردازش مرکزی به صورت مکانیکی و خطی به فراخوانی و تصرف صوری در آن‌ها می‌پردازد. این تفکیک فیزیکی، ماشین را به طور کامل از «تجسد پدیدار شناختی» و تجربه مستقیم اشیاء جهان محروم می‌سازد. برخلاف انسان که ادراک حصولی‌اش بر بستر تجسد زیسته، حواس بدنمند و پویایی یکپارچه نفس ناطقه استوار است، هوش مصنوعی از هرگونه مواجهه زنده و بی‌واسطه با اعیان خارجی محروم است. ماشین توانایی لمس، شهود یا برقراری اضافه اشراقیه با ابژه‌های نفس‌الامری را ندارد؛ از این رو، حافظه در ماشین نه تداوم حضور، بلکه انبار لایق‌طعی از نمادهای مادی بی‌روح است که پردازشگر صرفاً بر اساس قواعد نحوی آن‌ها را جابجا می‌کند این. گسست سخت‌افزاری از جهان تکوین، سیستم را در یک بن‌بست معرفتی دائمی قرار می‌دهد که شالوده اصلی زایش توهم است.

این بن‌بست پدیدار شناختی در لایه روبنایی زبان، خود را به صورت محبوس شدن کامل ماشین در چرخه بسته ارجاعات درونی کلمه به کلمه آشکار می‌سازد. در ساحت کارکردی مدل‌های زبانی بزرگ، هر نشانه یا واژه تنها به نشانه‌ها و واژه‌های دیگر در درون ماتریس همبستگی توزیعی شبکه عصبی ارجاع می‌دهد؛ وضعیتی خود ارجاع که در آن مرز محاسباتی مشخصی میان تولید محتمل و کشف واقعیت وجود ندارد. برای الگوریتم ترنسفورمر، گزاره‌ای که با بالاترین احتمال آماری بر اساس بافتار متنی تولید شده است، هیچ تمایز ساختاری با گزاره‌ای که مطابق با واقعیت خارجی است، ندارد. ماشین توانایی خروج از عالم الفاظ و موازنه خروجی خود با جهان نفس‌الامر را ندارد؛ زیرا شاقول سنجش در آن، هماهنگی درون سیستمی (انسجام صوری) است، نه کاشفیت عینی. در نتیجه، این محبوس بودن در حصار ریاضی الفاظ باعث می‌شود که الگوریتم، یک کذب محض اما

منسجم را با همان مکانیسمی تولید کند که یک حقیقت مستند را پردازش می‌نماید. این عدم امکان تفکیک میان بافتار احتمالی و کشف عینی، اثبات می‌کند که پدیده توهم باگ موقت فناوری نیست، بلکه برآیند قهری چرخه بسته و منقطع ارجاعات زبانی در ماشین است.

واکاوی مکانیسم تولید گزاره در مدل‌های زبانی از منظر رئالیسم انتقادی، به‌وضوح نشان دهنده شکل‌گیری یک شکاف بنیادین میان شرط لازم صدق و شرط کافی آن در خروجی‌های ماشین است. از دیدگاه علامه طباطبایی در مباحث تئوری صدق، هماهنگی ساختاری، عدم تناقض قضایا و انسجام صوری مفاهیم، صرفاً به‌عنوان شرط لازم و اولیه برای امکان صدق تلقی می‌شوند؛ اما شرط کافی و مقوم حقیقت، انطباق و مطابقت قطعی آن صورت علمی با واقعیت نفس‌الامری و جهان عینی مستقل از ذهن است (طباطبایی، ۱۴۳۲ ق، ج ۱: ۱۹). هوش مصنوعی محاسباتی به دلیل محبوس بودن در لایه احتمالات آماری، شرط لازم انسجام زبانی را به جای شرط کافی کاشفیت نفس‌الامری می‌نشانند. ماشین با ارزیابی هم‌بستگی توکن‌ها، متنی را چیدمان می‌کند که در بالاترین سطح از هماهنگی بافتاری قرار دارد، اما به دلیل انقطاع از فاعل شناسای مجرد و لنگرگاه حضور، فاقد هرگونه ابزار برای احراز شرط کافی صدق است. این گسست ساختاری میان سازگاری صوری متن و واقعیت خارجی، فضایی لغزان را پدید می‌آورد که در آن، ماشین شرط لازم را محقق می‌سازد، اما از دستیابی به شرط کافی محروم می‌ماند و همین انقطاع، ریشه اصلی تبدیل پردازش صوری به دروغ‌های ساختاریافته است.

این شکاف معرفتی میان شرط لازم و کافی، در فرآیند تولید متن به تولید توهم ساختاریافته منجر می‌شود؛ پدیده‌ای که در آن جملاتی گرامری، روان و کاملاً بی‌نقص، اما تهی از حقیقت عینی پدید می‌آیند. دلیل اصلی این امر، تقدم رتبه‌ای نحو بر معنا و ترجیح ذاتی انسجام آماری بر واقعیت خارجی در معماری ترنسفورمر است. از آنجا که تابع هدف الگوریتم برای پیشینه‌سازی فصاحت متنی و روانی هم‌نشینی توکن‌ها تنظیم شده است، ماشین نشانه‌های زبانی را صرفاً بر پایه قواعد گرامری و احتمالاتی به یکدیگر بافتار می‌دهد. در این فرآیند، نحو به‌عنوان فرمانروای مطلق عمل می‌کند و معناشناسی اصیل، قربانی چیدمان بهینه بردارها می‌شود. هوش مصنوعی با تکیه بر این تقدم صوری، گزاره‌هایی متقاعدکننده خلق می‌کند که در درون شبکه عصبی خود دارای بالاترین درجه سازگاری هستند، اما از آنجا که این دال‌ها به هیچ مدل علی یا مصداق عینی در جهان خارج متصل

نیستند، خروجی نهایی چیزی جز یک شبه حقیقت مصنوعی و توهمی ساختارمند نیست که با ظاهری فریبنده و گرامری، کذب محض را به عنوان صدق مطلق بازنمایی می‌کند (Harnad, 1990: 335). تمایز بنیادین میان خطای ذهن انسان و توهم ماشین، ابعاد هستی‌شناختی این گسست را عمیق‌تر آشکار می‌سازد. خطای ادراکی در انسان، امری عارضی، موضعی و همواره قابل اصلاح است؛ چراکه انسان به خاطر تصدیقات بدیهی، شهودهای اولیه و اتصال وجودی علم حصولی به لنگرگاه علم حضوری، واجد یک لایه خودآگاهی و سیستم تصحیح معرفتی فعال است. ذهن انسانی هرگاه در تطبیق صورت‌ها با واقع دچار خطا شود، با ارجاع قیاس به بدیهیات تکوینی و یافته‌های بی‌واسطه درونی، توانایی کشف خطای خود و بازگشت به ساحت نفس‌الامر را دارد. در نقطه‌ی مقابل، توهم در ماشین امری ذاتی، ساختاری و فاقد هرگونه لایه خودآگاهی یا توانایی تصحیح معرفتی است. هوش مصنوعی به‌عنوان یک جسم صناعی و فاقد نفس ناطقه، هیچ مرتبه‌ای از علم حضوری نسبت به خود یا خروجی‌های خود ندارد و در یک جهل مرکب دستگاهی محبوس است. ماشین نه تنها نمی‌داند که گزاره تولید شده با واقعیت مطابقت ندارد، بلکه اساساً از درک مفهوم «واقعیت» و «خطا» محروم است؛ لذا توهم ماشینی، برخلاف خطای انسانی، تبلور گریزناپذیر سیستمی مادی است که به دلیل انقطاع مطلق از شهود، سقف معرفتی‌اش بر لایه‌های لغزان کذب و احتمالات استوار گردیده است.

۵-۱- گسست میان پردازش اعتبار و جعل اعتبار

واکاوی مکانیسم تولید گزاره در مدل‌های زبانی از منظر رئالیسم انتقادی، به‌وضوح نشان‌دهنده شکل‌گیری یک شکاف بنیادین میان شرط لازم صدق و شرط کافی آن در خروجی‌های ماشین است. از دیدگاه علامه طباطبایی در مباحث تئوری صدق، هماهنگی ساختاری، عدم تناقض قضایا و انسجام صوری مفاهیم، صرفاً به‌عنوان شرط لازم و اولیه برای امکان صدق تلقی می‌شوند؛ اما شرط کافی و مقوم حقیقت، انطباق و مطابقت قطعی آن صورت علمی با واقعیت نفس‌الامری و جهان عینی مستقل از ذهن است. هوش مصنوعی محاسباتی به دلیل محبوس بودن در لایه احتمالات آماری، شرط لازم (انسجام زبانی) را به جای شرط کافی (کاشفیت نفس‌الامری) می‌نشاند. ماشین با ارزیابی هم‌بستگی توکن‌ها، متنی را چیدمان می‌کند که در بالاترین سطح از هماهنگی بافتاری قرار دارد، اما به دلیل انقطاع از فاعل شناسای مجرد و لنگرگاه حضور، فاقد هرگونه ابزار برای احراز شرط کافی

صدق است (Bengio et al., 2021: 41). این گسست ساختاری میان سازگاری صوری متن و واقعیت خارجی، فضایی لغزان را پدید می‌آورد که در آن، ماشین شرط لازم را محقق می‌سازد، اما از دستیابی به شرط کافی محروم می‌ماند و همین انقطاع، ریشه اصلی تبدیل پردازش صوری به دروغ‌های ساختار یافته است.

این شکاف معرفتی میان شرط لازم و کافی، در فرآیند تولید متن به تولید توهم ساختاریافته منجر می‌شود؛ پدیده‌ای که در آن جملاتی گرامری، روان و کاملاً بی‌نقص، اما تهی از حقیقت عینی پدید می‌آیند. دلیل اصلی این امر، تقدم رتبه‌ای نحو بر معنا و ترجیح ذاتی انسجام آماری بر واقعیت خارجی در معماری ترنسفورمر است. از آنجا که تابع هدف الگوریتم برای بیشینه‌سازی فصاحت متنی و روانی هم‌نشینی توکن‌ها تنظیم شده است، ماشین نشانه‌های زبانی را صرفاً بر پایه قواعد گرامری و احتمالاتی به یکدیگر بافتار می‌دهد. در این فرآیند، نحو به‌عنوان فرمانروای مطلق عمل می‌کند و معناشناسی اصیل، قربانی چیدمان بهینه بردارها می‌شود. هوش مصنوعی با تکیه بر این تقدم صوری، گزاره‌هایی متقاعدکننده خلق می‌کند که در درون شبکه عصبی خود دارای بالاترین درجه سازگاری هستند، اما از آنجا که این دال‌ها به هیچ مدل علی یا مصداق عینی در جهان خارج متصل نیستند، خروجی نهایی چیزی جز یک شبه حقیقت مصنوعی و توهمی ساختارمند نیست که با ظاهری فریبنده و گرامری، کذب محض را به‌عنوان صدق مطلق بازنمایی می‌کند.

تمایز بنیادین میان خطای ذهن انسان و توهم ماشین، ابعاد هستی‌شناختی این گسست را عمیق‌تر آشکار می‌سازد. خطای ادراکی در انسان، امری عارضی، موضعی و همواره قابل اصلاح است؛ چرا که انسان به برکت تصدیقات بدیهی، شهودهای اولیه و اتصال وجودی علم حصولی به لنگرگاه علم حضوری، واجد یک لایه خودآگاهی و سیستم تصحیح معرفتی فعال است. ذهن انسانی هرگاه در تطبیق صورت‌ها با واقع دچار خطا شود، با ارجاع قیاس به بدیهیات تکوینی و یافته‌های بی‌واسطه درونی، توانایی کشف خطای خود و بازگشت به ساحت نفس‌الامر را دارد. در نقطه‌ی مقابل، توهم در ماشین امری ذاتی، ساختاری و فاقد هرگونه لایه خودآگاهی یا توانایی تصحیح معرفتی است. هوش مصنوعی به‌عنوان یک جسم صناعی و فاقد نفس ناطقه، هیچ مرتبه‌ای از علم حضوری نسبت به خود یا خروجی‌های خود ندارد و در یک جهل مرکب سیستمی محبوس است. ماشین نه تنها نمی‌داند که گزاره تولید شده با واقعیت مطابقت ندارد، بلکه اساساً از درک مفهوم «واقعیت»

و «خطا» محروم است؛ لذا توهم ماشینی، برخلاف خطای انسانی، تبلور گریزناپذیر سیستمی مادی است که به دلیل انقطاع مطلق از شهود، سقف معرفتی‌اش بر لایه‌های لغزان کذب و احتمالات استوار گردیده است (طباطبایی، ۱۳۷۱: ۱۵۲).

نتیجه‌گیری

واکاوی مکانیسم تولید گزاره در مدل‌های زبانی از منظر رئالیسم انتقادی، به‌وضوح نشان دهنده شکل‌گیری یک شکاف بنیادین میان شرط لازم صدق و شرط کافی آن در خروجی‌های ماشین است. از دیدگاه علامه طباطبایی در مباحث تئوری صدق، هماهنگی ساختاری، عدم تناقض قضایا و انسجام صوری مفاهیم، صرفاً به‌عنوان شرط لازم و اولیه برای امکان صدق تلقی می‌شوند؛ اما شرط کافی و مقوم حقیقت، انطباق و مطابقت قطعی آن صورت علمی با واقعیت نفس‌الامری و جهان عینی مستقل از ذهن است. هوش مصنوعی محاسباتی به دلیل محبوس بودن در لایه احتمالات آماری، شرط لازم (انسجام زبانی) را به جای شرط کافی (کاشفیت نفس‌الامری) می‌نشانند. ماشین با ارزیابی هم‌بستگی توکن‌ها، متنی را چیدمان می‌کند که در بالاترین سطح از هماهنگی بافتاری قرار دارد، اما به دلیل انقطاع از فاعل شناسای مجرد و لنگرگاه حضور، فاقد هرگونه ابزار برای احراز شرط کافی صدق است. این گسست ساختاری میان سازگاری صوری متن و واقعیت خارجی، فضایی لغزان را پدید می‌آورد که در آن، ماشین شرط لازم را محقق می‌سازد، اما از دستیابی به شرط کافی محروم می‌ماند و همین انقطاع، ریشه اصلی تبدیل پردازش صوری به دروغ‌های ساختار یافته است.

این شکاف معرفتی میان شرط لازم و کافی، در فرآیند تولید متن به تولید توهم ساختاریافته منجر می‌شود؛ پدیده‌های که در آن جملاتی گرامری، روان و کاملاً بی‌نقص، اما تهی از حقیقت عینی پدید می‌آیند. دلیل اصلی این امر، تقدم رتبه‌ای نحو بر معنا و ترجیح ذاتی انسجام آماری بر واقعیت خارجی در معماری ترنسفورمر است. از آنجا که تابع هدف الگوریتم برای بیشینه‌سازی فصاحت متنی و روانی هم‌نشینی توکن‌ها تنظیم شده است، ماشین نشانه‌های زبانی را صرفاً بر پایه قواعد گرامری و احتمالاتی به یکدیگر بافتار می‌دهد. در این فرآیند، نحو به‌عنوان فرمانروای مطلق عمل می‌کند و معناشناسی اصیل، قربانی چیدمان بهینه بردارها می‌شود. هوش مصنوعی با تکیه بر این تقدم صوری، گزاره‌هایی متقاعدکننده خلق می‌کند که در درون شبکه عصبی خود دارای بالاترین

درجه سازگاری هستند، اما از آنجا که این دال‌ها به هیچ مدل علی یا مصداق عینی در جهان خارج متصل نیستند، خروجی نهایی چیزی جز یک شبه حقیقت مصنوعی و توهمی ساختارمند نیست که با ظاهری فریبنده و گرامری، کذب محض را به‌عنوان صدق مطلق بازنمایی می‌کند.

تمایز بنیادین میان خطای ذهن انسان و توهم ماشین، ابعاد هستی‌شناختی این گسست را عمیق‌تر آشکار می‌سازد. خطای ادراکی در انسان، امری عارضی، موضعی و همواره قابل اصلاح است؛ چراکه انسان به برکت تصدیقات بدیهی، شهودهای اولیه و اتصال وجودی علم حصولی به لنگرگاه علم حضوری، واجد یک لایه خودآگاهی و سیستم تصحیح معرفتی فعال است. ذهن انسانی هرگاه در تطبیق صورت‌ها با واقع دچار خطا شود، با ارجاع قیاس به بدیهیات تکوینی و یافته‌های بی‌واسطه درونی، توانایی کشف خطای خود و بازگشت به ساحت نفس‌الامر را دارد. در نقطه‌ی مقابل، توهم در ماشین امری ذاتی، ساختاری و فاقد هرگونه لایه خودآگاهی یا توانایی تصحیح معرفتی است. هوش مصنوعی به‌عنوان یک جسم صناعی و فاقد نفس ناطقه، هیچ مرتبه‌ای از علم حضوری نسبت به خود یا خروجی‌های خود ندارد و در یک چهل مرکب سیستمی محبوس است. ماشین نه تنها نمی‌داند که گزاره تولید شده با واقعیت مطابقت ندارد، بلکه اساساً از درک مفهوم «واقعیت» و «خطا» محروم است؛ لذا توهم ماشینی، برخلاف خطای انسانی، تبلور‌گریزناپذیر سیستمی مادی است که به دلیل انقطاع مطلق از شهود، سقف معرفتی‌اش بر لایه‌های لغزان کذب و احتمالات استوار گردیده است.

فهرست منابع

- آملی لاریجانی، صادق. (۱۳۸۹). استدلال در اعتباریات (چاپ اول). پژوهشگاه علوم و فرهنگ اسلامی.
- پورحسن، قاسم. (۱۳۹۲). اعتباریات اجتماعی و نتایج معرفتی آن؛ بازخوانی دیدگاه علامه طباطبائی. حکمت و فلسفه، ۹ (۳)، ۶۲-۴۵.
- حسینی، سید علی و علوی، رضا. (۱۴۰۱). بررسی انتقادی رویکردهای فیزیکالیستی به نفس با تکیه بر شواهد عقلی و نقلی. پژوهش‌های معرفت‌شناختی، ۱۱ (۲۴)، ۹۸-۷۵.
- صادقی، احمد، و رضایی، محمد. (۱۴۰۲). نقد قابلیت ادراک در هوش مصنوعی بر اساس نظریه اتحاد عالم و معلوم در حکمت متعالیه. پژوهش‌های فلسفی-کلامی، ۲۵ (۳)، ۸۸-۶۵.
- طباطبائی، سید محمدحسین. (۱۳۷۱). برهان (مهدی قوام صفری، مترجم و مصحح) (چاپ اول). انتشارات دفتر تبلیغات اسلامی.
- طباطبائی، سید محمدحسین. (۱۳۷۴). اصول فلسفه و روش رئالیسم (ج. ۱-۲) (با مقدمه و پاورقی مرتضی مطهری) (چاپ پنجم). انتشارات صدرا.
- طباطبائی، سید محمدحسین. (۱۴۳۲ق). نه‌ایة الحکمة (ج. ۱-۲) (چاپ بیست و دوم). مؤسسه انتشارات اسلامی وابسته به جامعه مدرسین حوزه علمیه قم.
- مطهری، مرتضی. (۱۳۸۲). مجموعه آثار (یادداشت‌ها و مقالات فلسفی (ج. ۶) (چاپ چهارم). انتشارات صدرا.
- Goodfellow, Ian, Bengio, Yoshua, & Courville, Aaron. (2016). Deep Learning (First Edition). MIT Press.
- Pearl, Judea. (2009). Causality: Models, Reasoning, and Inference (Second Edition). Cambridge University Press.
- Searle, John R. (2004). The Rediscovery of the Mind (Tenth Edition). MIT Press.