

A Conceptual Model of Digital Islamic and Humanities Studies Based on Generative Artificial Intelligence: An Approach to the Challenges of Accuracy, Bias, and Ethics

Hadi Moezzifar¹✉  and Mohammad Javad Rezaei² 

1. Corresponding Author, Faculty Member, Department of Islamic Studies, Birjand University of Technology, Birjand, Iran, Email: moezzifar@birjandut.ac.ir
2. Birjand University of Technology, Birjand, Iran, Email: m.javad-rezaei@birjandut.ac.ir

Article Info

Article type:
Research Article

Article history:

Received:

8/07/2025

Received in revised form:

2/09/2025

Accepted:

19/10/2025

Available online:

4/04/2026

Keywords:

Generative Artificial Intelligence (GenAI); Digital Islamic Studies; Digital Humanities; Accuracy and Algorithmic Bias; AI Ethics.

ABSTRACT

Digital transformation and the rapid advancement of generative artificial intelligence (GenAI) have created new opportunities and challenges for Islamic studies. On the one hand, the capabilities of intelligent language models in natural language processing and text generation have facilitated the rapid extraction, classification, and analysis of Islamic data, opening new avenues for interdisciplinary research. On the other hand, these technologies face significant challenges, including semantic distortion, cultural bias, and insufficient adherence to epistemological and ethical principles, all of which may undermine the credibility of research in Islamic studies.

Accordingly, the present study aims to develop a conceptual framework for the application of generative artificial intelligence within the field of Islamic studies. The research adopts a descriptive-analytical approach, with data collected from two primary sources: Islamic texts (the Qur'an and Hadith) and the outputs generated by generative AI models. The research process was conducted in four stages: (1) selecting and preparing authoritative Islamic texts, (2) processing these texts using generative AI models, (3) comparing the generated outputs with authoritative references, and (4) categorizing and analyzing the identified errors.

The findings indicate that the most frequent errors produced by the models fall into three major categories: semantic distortion, omission or addition of unsupported content, and interpretations influenced by non-Islamic cultural perspectives, accounting for 35%, 25%, and 15% of the observed errors, respectively. The analysis further demonstrates that, despite the technical capabilities of large language models, the absence of a comprehensive framework for ensuring accuracy, mitigating bias, and maintaining ethical standards results in fundamental shortcomings.

Based on these findings, the study proposes a five-layer conceptual model designed to address these challenges. The proposed framework can serve as a foundation for the development of intelligent Islamic applications, the enhancement of digital literacy among scholars of Islamic studies, and the formulation of strategic policies in the field of digital humanities.

The findings suggest that the future of Digital Islamic Studies lies in the synergy between technological capabilities and epistemological-ethical authority. The responsible and informed application of generative artificial intelligence has the potential to significantly enrich Islamic scholarship in the digital age.

Cite this article: Moezzifar, H.; & Rezaei, M.J. (2026). A Conceptual Model of Digital Islamic and Humanities Studies Based on Generative Artificial Intelligence: An Approach to the Challenges of Accuracy, Bias, and Ethics. *Digital Islamic Studies and Humanities*, 2 (3), 259-282. <https://doi.org/10.22034/disah.2026.2094330.1093>



© The Author(s). **Publisher:** Research Center for Digital Islamic Studies and Humanities (RCDISAH).

DOI: <https://doi.org/10.22034/disah.2026.2094330.1093>

مدل مفهومی علوم اسلامی و انسانی دیجیتال بر پایه هوش مصنوعی مولد:

رویکردی به چالش‌های صحت، سوگیری و اخلاق

هادی معزی^۱ فر^۲ و محمدجواد رضائی^۲

۱. عضو هیئت علمی گروه معارف دانشگاه صنعتی بیرجند، رایانامه: moezzifar@birjandut.ac.ir

۲. دانشگاه صنعتی بیرجند، رایانامه: m.javad-rezaei@birjandut.ac.ir

چکیده

اطلاعات مقاله

تحول دیجیتال و گسترش هوش مصنوعی مولد، فرصت‌ها و چالش‌های تازه‌ای را پیش‌روی علوم اسلامی قرار داده است. از یک‌سو، قابلیت‌های پردازش زبان طبیعی و تولید متن در مدل‌های هوشمند، امکان استخراج، طبقه‌بندی و تحلیل سریع‌تر داده‌های اسلامی را فراهم ساخته و مسیرهای جدیدی برای پژوهش‌های میان‌رشته‌ای گشوده است؛ و از سوی دیگر، این فناوری‌ها با مسائلی همچون تحریف معنایی، سوگیری فرهنگی و ضعف در پایبندی به اصول معرفتی و اخلاقی مواجه‌اند که می‌تواند اعتبار پژوهش‌های اسلامی را تحت‌تأثیر قرار دهد. بر این اساس، پژوهش حاضر با هدف طراحی الگویی مفهومی برای بهره‌گیری از هوش مصنوعی مولد در بستر علوم اسلامی انجام شد. روش تحقیق به‌صورت توصیفی-تحلیلی بوده و داده‌ها از دو منبع اصلی شامل متون اسلامی (قرآن و حدیث) و خروجی‌های مدل‌های AI مولد گردآوری شد. فرایند پژوهش در چهار گام اجرا گردید: انتخاب و آماده‌سازی متون اسلامی مرجع، پردازش آنها توسط مدل‌های مولد، مقایسه نتایج با منابع معتبر و در نهایت دسته‌بندی و تحلیل خطاها. یافته‌ها نشان داد که بیشترین خطاهای مدل در سه دسته تحریف معنایی، حذف یا اضافه محتوای بی‌منبع و ارائه تفسیرهای متأثر از فرهنگ غیر اسلامی قرار می‌گیرد و درصد وقوع این خطاها به‌ترتیب ۳۵، ۲۵ و ۱۵ درصد بوده است. تحلیل نتایج آشکار ساخت که صرف‌نظر از توانمندی‌های فنی مدل‌های زبانی، نبود چارچوبی یکپارچه برای کنترل صحت، کاهش سوگیری و تضمین اخلاق‌مداری موجب بروز کاستی‌های بنیادین می‌شود. بر همین مبنا، مدل مفهومی پیشنهادی پژوهش در پنج لایه طراحی گردید. این مدل می‌تواند مبنایی برای توسعه ابزارهای هوشمند اسلامی، ارتقای سواد دیجیتال در میان پژوهشگران علوم اسلامی و سیاست‌گذاری‌های کلان در حوزه علوم انسانی دیجیتال باشد. نتایج تحقیق نشان می‌دهد که آینده علوم اسلامی دیجیتال در هم‌افزایی میان ظرفیت‌های فناورانه و مرجعیت معرفتی-اخلاقی نهفته است و بهره‌گیری هوشمندانه از هوش مصنوعی مولد می‌تواند به غنای پژوهش‌های اسلامی در عصر دیجیتال باری رساند.

نوع مقاله:

مقاله پژوهشی

تاریخ دریافت:

۱۴۰۴/۰۴/۱۷

تاریخ بازنگری:

۱۴۰۴/۰۶/۱۱

تاریخ پذیرش:

۱۴۰۴/۰۷/۲۷

تاریخ انتشار:

۱۴۰۵/۰۱/۱۵

کلیدواژه‌ها:

هوش مصنوعی مولد،
علوم اسلامی دیجیتال،
علوم انسانی دیجیتال،
صحت و سوگیری
الگوریتمی،
اخلاق هوش مصنوعی.

استناد: معزی‌فر، هادی؛ و رضائی، محمدجواد (۱۴۰۵). مدل مفهومی علوم اسلامی و انسانی دیجیتال بر پایه هوش مصنوعی

مولد: رویکردی به چالش‌های صحت، سوگیری و اخلاق. علوم انسانی و اسلامی دیجیتال، ۲ (۳)، ۲۵۹-۲۸۲.

<https://doi.org/10.22034/disah.2026.2094330.1093>

ناشر: پژوهشگاه علوم اسلامی و انسانی دیجیتال (مرکز تحقیقات کامپیوتری علوم اسلامی نور). © نویسندگان.

مقدمه

در دهه‌های اخیر، مفهوم علوم اسلامی دیجیتال به عنوان شاخه‌ای نوین از علوم انسانی دیجیتال مطرح شده است که به کارگیری ابزارهای فناورانه، پایگاه‌های داده و روش‌های محاسباتی را در تحلیل متون، اسناد و میراث اسلامی دنبال می‌کند. این حوزه با بهره‌گیری از فناوری‌هایی همچون متن‌کاوی، یادگیری ماشین و هوش مصنوعی، امکان جست‌وجو، طبقه‌بندی و تحلیل دقیق منابع دینی را فراهم می‌سازد (مرکز تحقیقات کامپیوتری علوم اسلامی، ۱۴۰۴؛ Corpus Coranicum, n.d.).

در کنار این تحول، مفهوم انسان دیجیتال جایگاه برجسته‌ای یافته است. انسان دیجیتال نه تنها مصرف‌کننده محتوای دیجیتال، بلکه تولیدکننده داده و شریک فعال الگوریتم‌ها در فرآیند تولید دانش است. در چنین وضعیتی، مرز میان فضای فیزیکی و مجازی بیش از پیش کمرنگ شده و پرسش‌هایی نو درباره هویت، معنا و ارزش‌های انسانی پدید آمده است. این شرایط به‌ویژه برای جوامع اسلامی اهمیت مضاعف دارد، زیرا تغییر در الگوهای زیست دیجیتال می‌تواند بر هویت دینی و فرهنگی نیز تأثیرگذار باشد (معزی فر و رضائی، ۱۴۰۴؛ Akhtar, 2024).

در این بستر، هوش مصنوعی (Artificial Intelligence) به یکی از پرنفوذترین فناوری‌های تحول‌آفرین بدل شده است. این شاخه از علوم رایانه با شبیه‌سازی فرایندهای شناختی و تصمیم‌گیری انسانی، ابزارهای تازه‌ای برای تحلیل و بازتولید دانش فراهم آورده است. قابلیت‌هایی همچون پردازش زبان طبیعی (NLP)، یادگیری عمیق و تحلیل داده‌های بزرگ، امکان استخراج الگوهای پنهان و تحلیل نوین متون اسلامی را ایجاد کرده‌اند (Patel et al, 2023). در میان انواع هوش مصنوعی، هوش مصنوعی مولد (Generative AI) جایگاهی ویژه دارد. مدل‌هایی چون GPT با یادگیری از داده‌های عظیم، توانایی تولید متن، تصویر و حتی صدا را دارند و می‌توانند در حوزه‌هایی همچون تفسیر اولیه متون دینی یا شبیه‌سازی پرسش و پاسخ‌های فقهی به‌کار گرفته شوند (AI-Salihah Journal, 2024).

با این حال، بهره‌گیری از AI در علوم اسلامی با چالش‌هایی جدی همراه است. یکی از مهم‌ترین آن‌ها سوگیری الگوریتمی است که در اثر داده‌های آموزشی یا ساختار الگوریتم ایجاد می‌شود و می‌تواند به بازتولید کلیشه‌ها یا تفسیرهای مغرضانه منجر شود (Caliskan et al, 2021; Abid et al, 2021).

al, 2017). افزون بر آن، صحت و اعتبار داده‌ها در علوم اسلامی دیجیتال نقشی حیاتی دارند، زیرا هر خطا در داده‌های ورودی یا پردازش می‌تواند به استنباط‌های نادرست و نتایج مخدوش منجر شود (محمد موسایی و همکاران، ۱۴۰۰). بعد دیگر، اخلاق هوش مصنوعی است که با چارچوب‌های ارزشی، حقوقی و فقهی جامعه اسلامی پیوند دارد. این حوزه به‌ویژه در ارتباط با مقاصد شریعت - شامل حفظ دین، جان، عقل، نسل و مال - می‌تواند معیارهایی برای طراحی و بهره‌برداری صحیح از سیستم‌های هوشمند فراهم آورد (Mohadi & Tarshany, 2023; Al-Faruqi, 1982).

از این منظر، روشن می‌شود که علوم اسلامی دیجیتال، انسان دیجیتال و هوش مصنوعی صرفاً مفاهیم تخصصی نیستند، بلکه چارچوب‌هایی بنیادی برای فهم تحولات دانش دینی در عصر حاضر به‌شمار می‌آیند. فناوری‌های هوشمند، هم ظرفیت‌های گسترده برای ارتقای پژوهش‌های اسلامی دارند و هم تهدیدهایی جدی متوجه صحت و اصالت منابع ایجاد می‌کنند. بنابراین، استفاده از آن‌ها نیازمند رویکردی جامع است که علاوه بر جنبه‌های فنی، ملاحظات معرفتی و اخلاقی را نیز در نظر گیرد.

مرور ادبیات موجود نشان می‌دهد که پژوهش‌ها در این حوزه مسیری تدریجی اما پراکنده طی کرده‌اند. نخستین رویکردهای مفهومی را می‌توان در آثار اندیشمندانی چون الفاروقی (Al-Faruqi, 1982) مشاهده کرد که با طرح ایده «اسلامی‌سازی دانش» زمینه‌بازاندیشی رابطه دین و فناوری را فراهم آوردند. با پیشرفت‌های پردازش زبان طبیعی، پژوهش‌هایی مانند Caliskan et al. (2017) به تحلیل سوگیری‌های زبانی پرداختند و حساسیت نسبت به پیامدهای اجتماعی این فناوری‌ها را افزایش دادند. در سطح داخلی، از اواخر دهه ۱۳۹۰ توجه جدی‌تری به AI در علوم اسلامی شکل گرفت. پژوهشگرانی چون نصیری (۱۳۹۹) و منتصری (۱۳۹۹) به بررسی ابعاد اخلاقی و معرفتی پرداختند، در حالی که کارهای کاربردی‌تری مانند پژوهش پوربدخشان (۱۳۹۹) در حوزه هستان‌شناسی و تحلیل زبان فارسی، یا تلاش جوزی و بشیری موحد (۱۳۹۹) برای بازآفرینی ساختار موضوعی نورمگز، به بهره‌گیری عملی‌تر از الگوریتم‌ها معطوف بودند.

در اوایل دهه ۱۴۰۰، رویکردهای جامع‌تری مطرح شدند. محمد موسایی و همکاران (۱۴۰۰) به‌طور نظام‌مند ظرفیت‌ها و چالش‌های AI در علوم اسلامی را بررسی کردند. حسنی آهنگر (۱۴۰۱) و مقدم‌فر (۱۴۰۳) بر ابعاد نظری و روش‌شناختی تأکید کردند، در حالی که ربیعی‌زاده

(۱۴۰۳) تجربه‌های غربی در علوم اسلامی دیجیتال را مقایسه نمود. در همین زمان، کوشکی و موسوی (۱۴۰۳) کاربرد هوش مصنوعی را در تحلیل داده‌های تاریخی دفاع مقدس نشان دادند و پژوهش‌های مرکز تحقیقات کامپیوتری علوم اسلامی (۱۴۰۴) ابزارهای نوینی برای متن‌کاوی قرآنی و حدیثی عرضه کردند. از منظر فلسفی و فرهنگی نیز، معزی‌فر و رضائی (۱۴۰۴) تأثیر AI بر هویت انسان در فرهنگ معاصر را بررسی کردند.

در عرصه بین‌المللی نیز، تلاش‌ها گسترده اما متنوع بوده‌اند. برخی مانند Abid et al. (2021) و Patel et al. (2023) بر کاهش سوگیری و انطباق مدل‌ها با جهان‌بینی اسلامی تمرکز کردند. گروهی دیگر همچون Mohadi & Tarshany (2023) و Mustapha et al. (2025) بر اصول اخلاقی و فقهی تأکید داشتند. در آثار جدیدتر، Hakim & Shamsuddin (2024) و Qayyum & Saifullah (2025) به نقد کاربست ChatGPT در پرتو آموزه‌های قرآنی پرداختند. در مجموع، این مطالعات هر یک بخشی از مسئله را روشن کرده‌اند.

با وجود گستردگی تحقیقات پیشین، اغلب این مطالعات یا به بررسی یک بُعد خاص از هوش مصنوعی پرداخته‌اند – مانند تحلیل فنی مدل‌های زبانی یا ارزیابی جنبه‌های اخلاقی آن – یا در چارچوبی محدود به یک حوزه تخصصی انجام شده‌اند. در نتیجه، خلأ یک چارچوب مفهومی جامع که بتواند یافته‌های معرفتی، اخلاقی و فنی را هم‌زمان در بستر علوم اسلامی دیجیتال یکپارچه کند، همچنان محسوس است. این کمبود، نه تنها مانع بهره‌گیری نظام‌مند از ظرفیت‌های هوش مصنوعی در خدمت علوم اسلامی می‌شود، بلکه زمینه بروز خطراتی چون سوگیری الگوریتمی، تحریف داده‌های دینی و تعارض با اصول شریعت را نیز افزایش می‌دهد.

بر این اساس، پژوهش حاضر با هدف طراحی مدلی مفهومی برای «علوم اسلامی و انسانی دیجیتال بر پایه هوش مصنوعی مولد» است. نوآوری پژوهش در رویکرد میان‌رشته‌ای آن نهفته است: رویکردی که از یک‌سو، الزامات فنی مانند معماری مدل‌های زبانی و الگوریتم‌های یادگیری عمیق را مد نظر دارد و از سوی دیگر، مبانی دینی و انسانی همچون مقاصد شریعت، اصول اخلاقی و هویت دینی را در چارچوب تحلیلی خود لحاظ می‌کند. انتظار می‌رود این چارچوب پیشنهادی علاوه بر توسعه ابزارهای بومی‌سازی شده، به ارتقای سواد دیجیتال پژوهشگران و تقویت سیاست‌گذاری‌های کلان در حوزه هوش مصنوعی اسلامی منجر شود و جایگاه علوم اسلامی را در عصر تحولات فناورانه تثبیت و تقویت نماید.

روش تحقیق

در ادامه مباحث طرح شده در مقدمه و با توجه به ضرورت تدوین چارچوبی منسجم برای بهره‌گیری از هوش مصنوعی در علوم اسلامی، پژوهش حاضر با رویکردی توصیفی-تحلیلی طراحی و اجرا شده است. هدف از این روش، آن است که ضمن بررسی دقیق متون اسلامی و تطبیق آن‌ها با خروجی مدل‌های هوش مصنوعی مولد، تصویری روشن از ظرفیت‌ها و چالش‌های این فناوری در بستر علوم اسلامی ارائه گردد. انتخاب این رویکرد به دلیل ماهیت میان‌رشته‌ای موضوع و نیاز به ترکیب تحلیل‌های کیفی و کمی صورت گرفته است.

جامعه آماری این پژوهش شامل مجموعه‌ای از متون اسلامی معتبر است که دو دسته اصلی را دربرمی‌گیرد: نخست، متون بنیادین همچون قرآن کریم و احادیث منتخب از منابع حدیثی شیعه و اهل سنت؛ و دوم، مجموعه‌ای از متون علمی-پژوهشی اسلامی در حوزه‌های فقه، کلام و علوم انسانی اسلامی. این متون به‌عنوان داده‌های مرجع مورد استفاده قرار گرفتند تا امکان مقایسه خروجی مدل‌های هوش مصنوعی با آن‌ها فراهم شود.

نمونه‌گیری در این پژوهش به‌صورت هدفمند انجام شد؛ به این معنا که متونی انتخاب شدند که هم از جایگاه معرفتی و دینی بالایی برخوردارند و هم در دسترس‌ترین منابع برای پژوهشگران اسلامی محسوب می‌شوند. در مرحله بعد، از میان مدل‌های هوش مصنوعی مولد، نمونه‌هایی مانند ChatGPT و LLaMA به‌عنوان ابزار پژوهشی مورد استفاده قرار گرفتند. این مدل‌ها وظیفه داشتند بر اساس پرسش‌ها و سناریوهای ازپیش طراحی شده، تولید متن کنند تا امکان مقایسه خروجی آن‌ها با متون مرجع اسلامی فراهم گردد.

ابزار اصلی گردآوری داده‌ها، مجموعه پرسش‌ها و سناریوهای پژوهشی طراحی شده توسط محققان بود. روایی ابزار از طریق داوری خبرگان علوم اسلامی و متخصصان حوزه هوش مصنوعی مورد بررسی قرار گرفت؛ بدین‌صورت که پرسش‌ها به‌گونه‌ای تدوین شدند که هم پوشش‌دهنده موضوعات کلیدی دینی باشند و هم قابلیت پردازش توسط مدل‌های هوش مصنوعی را داشته باشند. پایایی ابزار نیز با اجرای آزمایشی و تکرار چندباره فرایند پرسش-پاسخ بررسی شد تا میزان ثبات نتایج به دست آید.

مراحل اجرای پژوهش در چهار گام اصلی سامان یافت:

۱. انتخاب متون اسلامی مرجع شامل قرآن، منابع حدیثی معتبر و متون علمی اسلامی.
 ۲. پردازش متون توسط مدل‌های هوش مصنوعی مولد همچون ChatGPT و LLaMA با طرح پرسش‌ها و سناریوهای پژوهشی.
 ۳. مقایسه خروجی‌های مدل‌ها با منابع معتبر به منظور شناسایی موارد انطباق و تعارض.
 ۴. دسته‌بندی و تحلیل خطاها و سوگیری‌ها در قالب شاخص‌هایی همچون میزان صحت پاسخ‌ها، نوع سوگیری (معرفتی، زبانی یا فرهنگی) و انطباق با اصول اخلاقی و دینی.
- برای شفاف‌سازی مراحل و اجزای تحقیق، ابتدا چارچوب کلی روش پژوهش در قالب یک جدول ارائه شده است. این جدول به صورت مرحله‌به‌مرحله، اجزای اصلی فرآیند تحقیق شامل جامعه آماری، منابع داده، ابزارهای پژوهش و روش‌های تجزیه و تحلیل را مشخص می‌سازد. هدف از این جدول، نمایش ساختار کلی پژوهش به صورت فشرده و نظام‌مند است تا مسیر تحقیق برای خواننده به طور روشن قابل درک باشد.

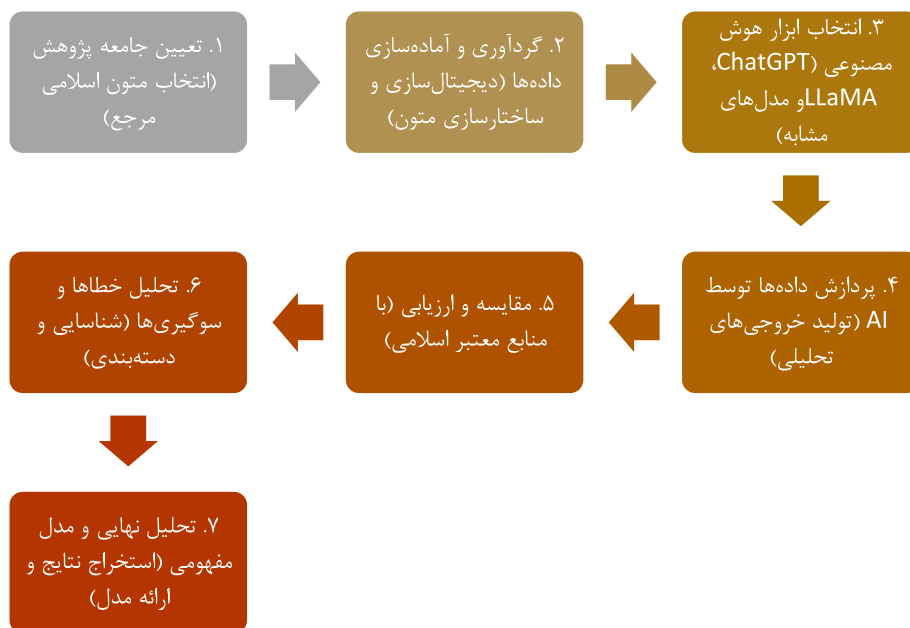
جدول ۱: مراحل اجرای پژوهش

مرحله	شرح فعالیت	خروجی مورد انتظار
۱. تعیین جامعه پژوهش	انتخاب متون اسلامی (قرآن کریم، احادیث معتبر و متون فقهی - تفسیری) به عنوان منابع مرجع.	مجموعه داده اولیه از متون اسلامی
۲. گردآوری و آماده‌سازی داده‌ها	دیجیتال‌سازی، پاک‌سازی و سازمان‌دهی متون اسلامی به صورت ساختاریافته برای پردازش ماشینی.	پایگاه داده متنی استاندارد
۳. انتخاب ابزار هوش مصنوعی	گزینش مدل‌های زبانی مولد (LLaMA، ChatGPT) و سایر مدل‌های مشابه) برای پردازش متون.	مدل‌های آماده برای پردازش متون اسلامی
۴. پردازش داده‌ها توسط AI	ارائه ورودی‌های متنی به مدل‌ها و تولید خروجی‌های مرتبط با تحلیل، تفسیر یا پاسخ به پرسش‌ها.	خروجی‌های تولید شده توسط مدل‌های AI
۵. مقایسه و ارزیابی	مقایسه خروجی مدل‌ها با منابع معتبر اسلامی از منظر صحت علمی، انطباق با مقاصد شریعت و اصول اخلاقی.	گزارش مقایسه و ارزیابی اولیه
۶. تحلیل خطاها و سوگیری‌ها	شناسایی موارد نادرست، سوگیری‌های فرهنگی و معرفتی و دسته‌بندی آنها در قالب محورها و زیر محورها.	دسته‌بندی و کدگذاری خطاها و سوگیری‌ها
۷. تحلیل نهایی	استفاده از تحلیل محتوایی و آمار توصیفی ساده برای استخراج درصد صحت، الگوهای سوگیری و نقاط ضعف/قوت.	نتایج نهایی و مدل مفهومی پیشنهادی

به منظور تکمیل توضیحات جدول، یک فلوچارت از فرآیند اجرای پژوهش طراحی شده است که ارتباط بین مراحل مختلف را به صورت بصری نمایش می دهد. این فلوچارت نشان می دهد که چگونه پژوهش از انتخاب متون اسلامی مرجع آغاز شده و پس از پردازش توسط مدل های هوش مصنوعی مولد، به مقایسه خروجی ها با منابع معتبر، دسته بندی خطاها و نهایتاً ارائه مدل مفهومی منجر می شود. استفاده از فلوچارت موجب می شود خواننده علاوه بر درک اجزای پژوهش، مسیر اجرایی آن را نیز به صورت گام به گام دنبال کند.

شکل ۱: فلوچارت فرآیند اجرای پژوهش

روش تحلیل داده ها ترکیبی از تحلیل محتوایی کیفی و تحلیل آماری ساده بود. در بخش کیفی،



مضامین اصلی استخراج و خطاها و سوگیری ها در طبقات مشخص جای گذاری شدند. در بخش کمی، درصد صحت پاسخ ها و فراوانی خطاها محاسبه شد تا تصویری عینی از عملکرد مدل های هوش مصنوعی ارائه گردد. این ترکیب از روش ها، هم امکان تبیین عمیق و کیفی مسائل را فراهم آورد و هم با ارائه شاخص های کمی، قابلیت مقایسه و ارزیابی دقیق تر نتایج را ممکن ساخت.

مدل مفهومی پیشنهادی

مرور ادبیات پژوهش نشان داد که اگرچه تلاش‌های ارزشمندی در حوزه کاربرد هوش مصنوعی در علوم اسلامی و انسانی صورت گرفته است، اما اغلب آنها یا بر بُعد فنی تمرکز داشته‌اند و از ملاحظات معرفتی و اخلاقی غفلت کرده‌اند، یا تنها به جنبه‌های فلسفی و نظری پرداخته‌اند و فاقد چارچوب عملیاتی برای پیاده‌سازی بوده‌اند. برخی مطالعات نیز تلاش کرده‌اند پیوندی میان مقاصد شریعت و هوش مصنوعی برقرار سازند، اما این تلاش‌ها محدود به تبیین‌های نظری یا بررسی‌های موردی باقی مانده و به سطح یک مدل جامع ارتقاء نیافته‌اند.

این خلأ دوگانه - فقدان جامعیت و فقدان کاربردپذیری - نقطه عزیمت این پژوهش است. به بیان دیگر، آنچه تاکنون مغفول مانده، یکپارچه‌سازی هم‌زمان سه بُعد اصلی است:

۱. بُعد معرفتی: اطمینان از اینکه داده‌ها و خروجی‌ها ریشه در منابع اصیل اسلامی داشته و از چارچوب‌های معرفتی معتبر پیروی می‌کنند.
۲. بُعد فنی: طراحی سازوکاری که از ظرفیت‌های هوش مصنوعی مولد برای پردازش، تحلیل و تولید دانش بهره‌گیرد.

۳. بُعد اخلاقی: ایجاد لایه‌ای نظارتی برای تضمین انطباق خروجی‌ها با اصول اخلاقی اسلامی و جلوگیری از بروز سوگیری یا انحراف.

بر همین اساس، پژوهش حاضر مدلی مفهومی را پیشنهاد می‌دهد که بتواند با رویکردی چندلایه، این سه بعد را در یک چارچوب منسجم گردآورد. این مدل نه تنها یک نقشه راه نظری، بلکه یک الگوی عملیاتی است که قابلیت انطباق با پروژه‌های واقعی در حوزه علوم اسلامی دیجیتال را دارد. شکل ۲، این مدل مفهومی را به صورت بصری نمایش می‌دهد و ارتباط نظام‌مند میان لایه‌های مختلف آن را آشکار می‌سازد.



شکل ۲: مدل مفهومی پیشنهادی علوم اسلامی و انسانی دیجیتال بر پایه هوش مصنوعی مولد

مدل پیشنهادی در قالب یک دیگرام پنج‌لایه‌ای ارائه می‌شود که هر لایه وظیفه‌ای مشخص در زنجیره پردازش دارد. خروجی‌ها با منابع معتبر مقایسه شده، صحت آنها سنجیده و خطا یا سوگیری احتمالی شناسایی می‌شود. این لایه تضمین‌کننده قابلیت اعتماد معرفتی و فنی است. در زیر توضیح تحلیلی لایه‌های مدل مفهومی پیشنهادی را به‌صورت جزئی بررسی می‌کنیم.

لایه نخست: ورودی داده‌های اسلامی

هر مدل تحلیلی در حوزه علوم اسلامی ناگزیر از نقطه آغازین یعنی داده‌های اصیل و معتبر اسلامی است. قرآن کریم، احادیث معتبر، متون فقهی و اصولی، آثار فلسفی-کلامی و حتی منابع تاریخی و تفسیری، شالوده این لایه را تشکیل می‌دهند. اهمیت این مرحله در آن است که بدون دسترسی به متون قابل اعتماد، هیچ فرایند مبتنی بر هوش مصنوعی قادر به ارائه خروجی معتبر نخواهد بود. بنابراین این لایه به مثابه "چشمه تغذیه" کل مدل عمل می‌کند.

لایه دوم: پیش‌پردازش و پاک‌سازی داده

متون اسلامی، به دلیل تنوع در رسم‌الخط، شیوه‌های نگارش، تفاوت نسخه‌ها و گاه وجود تعارض‌های روایی، نیازمند یک مرحله جدی از پالایش و استانداردسازی هستند. در این لایه، داده‌های خام به داده‌های پردازش‌شده تبدیل می‌شوند؛ شامل یکنواخت‌سازی رسم‌الخط، اعراب‌گذاری، حذف داده‌های ناسازگار و همچنین غنی‌سازی آنها با فراداده و هستان‌شناسی معنایی. بدین ترتیب، داده‌ها برای ورود به موتور تحلیل هوش مصنوعی آماده و به‌صورت علمی استاندارد می‌شوند.

لایه سوم: تولید پاسخ توسط هوش مصنوعی مولد

پس از آماده‌سازی داده‌ها، مدل‌های زبانی بزرگ (مانند GPT یا LLaMA) به کار گرفته می‌شوند تا بر پایه آن متون، خروجی‌های متنی یا تحلیلی تولید کنند. این مرحله نقطه‌ای است که ماشین تلاش می‌کند با بهره‌گیری از الگوریتم‌های یادگیری عمیق، الگوهای پنهان در داده‌ها را کشف و استدلال‌های جدید ارائه دهد. اهمیت این لایه در آن است که قابلیت خلاقانه و تولیدی هوش مصنوعی، توانایی ارتقای پژوهش اسلامی از سطح توصیف به سطح تحلیل و تبیین را فراهم می‌آورد.

لایه چهارم: ارزیابی صحت و کنترل سوگیری

هوش مصنوعی، علی‌رغم توانمندی، همواره در معرض خطا، تحریف و سوگیری است. این لایه تضمین می‌کند که خروجی‌های تولیدشده به منابع معتبر اسلامی بازگردانده شوند و صحت

آنها با معیارهای علمی و دینی سنجیده شود. در این مرحله خطاهای الگوریتمی شناسایی شده و سوگیری‌های فرهنگی یا زبانی کنترل می‌گردند. تحلیل محتوایی و استفاده از روش‌های آماری، ابزار اصلی این لایه برای ارزیابی و پایش کیفیت خروجی‌هاست.

لایه پنجم: لایه نظارت اخلاقی و ارائه خروجی نهایی

آخرین لایه، نقش نگهبان نهایی سیستم را دارد. حتی اگر خروجی‌ها از نظر فنی و صحت محتوایی معتبر باشند، بازهم باید با اصول اخلاقی و مقاصد شریعت انطباق یابند. در اینجا، متخصصان علوم اسلامی و اخلاق، خروجی‌های تولیدشده را پالایش و نظارت می‌کنند تا هرگونه مغایرت احتمالی حذف شود. سپس، نتایج نهایی در قالب داده‌های پژوهشی معتبر، ابزارهای کاربردی یا پاسخ‌های قابل استناد، در اختیار پژوهشگران و کاربران قرار می‌گیرد. این لایه، پیوند نهایی میان تکنولوژی و ارزش‌های اسلامی-انسانی است.

مدل مفهومی پیشنهادی این پژوهش با تکیه بر رویکردی میان‌رشته‌ای طراحی شده است تا ابعاد فنی، معرفتی و اخلاقی را به‌صورت یکپارچه در بستر علوم اسلامی دیجیتال ترکیب کند. این مدل، فراتر از یک طرح نظری، نقشه‌ای عملیاتی برای پردازش داده‌های اسلامی، کنترل سوگیری‌ها و تضمین صحت و اصالت متون ارائه می‌دهد. نوآوری اصلی آن در این است که به‌جای تمرکز صرف بر یک بُعد، سازوکاری هماهنگ میان فناوری، مبانی معرفتی و اصول اخلاقی برقرار می‌سازد و از این طریق می‌تواند به‌عنوان الگویی راهبردی برای پژوهش‌های آینده و توسعه سامانه‌های هوش مصنوعی بومی در حوزه علوم اسلامی به‌کار گرفته شود.

این مدل شامل پنج لایه به‌هم‌پیوسته است که به ترتیب از ورودی داده‌های اسلامی آغاز شده و با عبور از مراحل پردازش، تولید پاسخ، ارزیابی صحت و کنترل سوگیری، نهایتاً به لایه نظارت اخلاقی و ارائه خروجی معتبر ختم می‌شود. انسجام این لایه‌ها به نحوی طراحی شده است که هر مرحله هم مکمل و هم ناظر بر مرحله پیشین باشد؛ به‌گونه‌ای که داده‌های خام اسلامی در مسیر عبور از این فرایند نه تنها به پاسخ‌های علمی و کاربردی تبدیل شوند، بلکه از منظر دقت، شفافیت، عدالت و انطباق با مقاصد شریعت نیز اعتبار لازم را کسب کنند.

یافته‌ها

یافته‌های این پژوهش در امتداد مدل مفهومی پیشنهادی پنج‌لایه‌ای به دست آمده است. همان‌طور که در بخش قبل توضیح داده شد، این مدل با هدف ایجاد یک چارچوب ارزیابی چندبعدی برای به‌کارگیری هوش مصنوعی در علوم اسلامی طراحی شد. در این مرحله، برای آزمون عملی مدل، مجموعه‌ای از متون اسلامی شامل قرآن کریم، نهج البلاغه و برخی منابع معتبر حدیثی به‌عنوان داده‌های ورودی انتخاب شد. سپس مدل‌های هوش مصنوعی مولد (از جمله ChatGPT و LLaMA) در معرض پرسش‌هایی قرار گرفتند که از دل همین متون استخراج شده بودند؛ پرسش‌هایی که هم جنبه تفسیری داشتند، هم بُعد اخلاقی و معرفتی و هم زمینه‌های کاربردی در حوزه علوم اسلامی دیجیتال را پوشش می‌دادند.

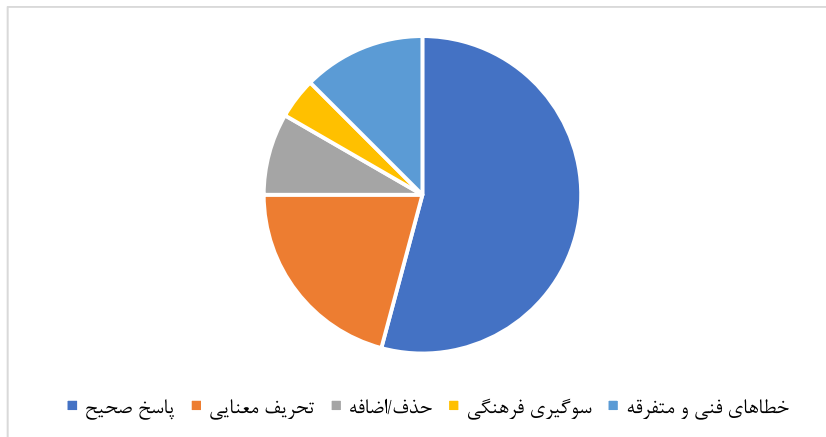
فرآیند ارزیابی به گونه‌ای طراحی شد که خروجی‌های تولید شده توسط مدل‌ها نه صرفاً از منظر صحت زبانی، بلکه در سه سطح «مطابقت با متن اصلی»، «انسجام معنایی با منابع معتبر» و «پرهیز از سوگیری فرهنگی یا معرفتی» بررسی شوند. این رویکرد موجب شد که یافته‌ها تنها بر اساس شمارش خطا یا درصد موفقیت مدل‌ها نباشند، بلکه بازتابی از کیفیت و عمق محتوای تولیدشده نیز ارائه دهند. بدین ترتیب، آنچه در ادامه می‌آید نه فقط آماری از میزان خطا، بلکه تصویری جامع از چالش‌ها و ظرفیت‌های واقعی هوش مصنوعی مولد در مواجهه با متون اسلامی است.

در گام نخست، خطاهای شناسایی‌شده در خروجی مدل‌ها مورد دسته‌بندی قرار گرفتند. همان‌گونه که جدول (۲) نشان می‌دهد، خطاها در سه گروه عمده «تحریف معنایی»، «حذف یا اضافه محتوا» و «سوگیری فرهنگی» جای گرفتند. تحریف معنایی بیشترین سهم خطاها را با ۳۵ درصد به خود اختصاص داد؛ به این معنا که مدل در بازتولید آیات یا احادیث، معنای اصلی را دگرگون ساخته یا به برداشت نادرست منتهی شده است. حذف یا اضافه محتوا با ۲۵ درصد در رتبه دوم قرار دارد که نشان می‌دهد مدل‌ها گاهی بخشی از متن اصلی را کنار گذاشته یا در مقابل، مطالبی بی‌منبعی را افزوده‌اند. در نهایت، خطاهای مربوط به سوگیری فرهنگی با ۱۵ درصد ثبت شدند که عمدتاً ناشی از تعمیم مفاهیم اسلامی بر اساس چارچوب‌های غیراسلامی یا غربی بوده است.

جدول ۲. دسته‌بندی خطاهای مدل هوش مصنوعی مولد در متون اسلامی

نوع خطا	توضیح	نمونه (ترجمه‌شده/خلاصه)	درصد وقوع
تحریف معنایی	تغییر در معنی یا مفهوم اصلی آیه یا حدیث	مدل واژه «تقوا» را به «ترس» صرف ترجمه کرد	۳۵٪
حذف یا اضافه محتوا	حذف بخشی از متن یا افزودن مطلب بدون منبع معتبر	مدل بخشی از روایت «نهج البلاغه» را حذف کرد	۲۵٪
سوگیری فرهنگی	تفسیر مطابق فرهنگ غیر اسلامی یا مفاهیم وارداتی غربی	تفسیر «عدالت» به سبک لیبرال دموکراسی	۱۵٪
خطای فنی/زبانی	اشتباهات نحوی یا ترکیب جملات غیر طبیعی	جملات ناتمام یا ناقص در توضیح آیه	۱۰٪
سایر خطاها	خطاهای متفرقه شامل تکرار، بی‌نظمی یا منابع نامشخص	آوردن منبع غیر معتبر مانند ویکی‌پدیا	۱۵٪

علاوه بر دسته‌بندی خطاها، نمودار توزیع درصد صحت و خطا به‌طور کلی در شکل ۳ نمایش داده شده است. این نمودار نشان می‌دهد که اگرچه بخش قابل توجهی از پاسخ‌ها (حدود ۶۵ درصد) صحیح و قابل استناد بودند، اما سهم خطاهای معنایی و محتوایی همچنان چالش مهمی برای اتکاپذیری مدل‌های زبانی در حوزه علوم اسلامی محسوب می‌شود. (مقادیر می‌توانند هم‌پوشانی داشته باشند چون یک پاسخ ممکن است چند خطا داشته باشد).



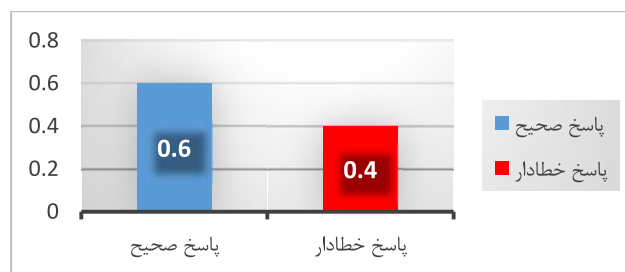
شکل ۳. نمودار درصد صحت و خطا در پاسخ‌های تولیدشده توسط مدل‌های زبانی

به منظور روشن تر شدن یافته‌ها، چند نمونه از پاسخ‌های صحیح و ناصحیح همراه با تحلیل در جدول ۳ ارائه می‌شود. این نمونه‌ها نشان می‌دهند که چرا مدل‌ها گاهی موفق به بازتولید صحیح معنا می‌شوند و گاهی دچار خطاهای بنیادی می‌گردند.

جدول ۳: نمونه پاسخ‌های صحیح و ناصحیح مدل هوش مصنوعی مولد

نوع پاسخ	پرسش نمونه	پاسخ مدل	مقایسه با منبع معتبر	تحلیل
صحیح	معنای «اقیموا الصلاة» چیست؟	«برپاداشتن نماز با همه ارکان و شرایط»	مطابق تفسیر المیزان و طبری	انطباق کامل با منابع اصلی
ناصحیح	مفهوم «تقوا» در قرآن	«ترس از خدا»	در منابع معتبر به معنای پرهیزکاری و مراقبت از گناه آمده است	خطای تحریف معنایی
ناصحیح	خطبه همام (نهج البلاغه)	بخشی از خطبه حذف شد	متن کامل در نهج البلاغه موجود است	خطای حذف محتوا
صحیح نسبی	مفهوم عدالت در اسلام	«برابری حقوقی همه انسان‌ها»	در اسلام عدالت بر اساس حق و تکلیف تعریف می‌شود	سوگیری فرهنگی - نیاز به اصلاح

به منظور تکمیل این یافته‌ها، درصد صحت کلی پاسخ‌های مدل‌ها نیز مورد بررسی قرار گرفت. همان‌طور که شکل ۴ نشان می‌دهد، این نمودار در نگاه اول مشابه شکل ۳ به نظر می‌رسد، اما تفاوت اصلی در این است که شکل ۴ جمع‌بندی کلی عملکرد مدل را در تمامی مراحل پژوهش نشان می‌دهد. به بیان دیگر، در حالی که شکل ۳ صرفاً درصد پاسخ‌های اولیه مدل را نمایش می‌داد، این نمودار برآیند پاسخ‌ها پس از مقایسه با منابع معتبر اسلامی است. این تمایز نشان می‌دهد که اعمال لایه‌های کنترلی و بازبینی انسانی تا چه اندازه می‌تواند دقت خروجی‌ها را ارتقا دهد.



شکل ۴: نمودار درصد صحت و خطا در پاسخ‌های مدل‌های هوش مصنوعی مولد

افزون بر آمارهای کلی، بررسی کیفی پاسخ‌ها نیز نتایج قابل توجهی به همراه داشت. در نمونه‌های صحیح، مدل‌ها توانستند به درستی متن آیات را بازتولید کرده و حتی ارتباط میان مفاهیم را در سطحی قابل قبول تبیین کنند. برای مثال، در پرسشی درباره مفهوم «تقوا» در قرآن، پاسخ ارائه شده توسط مدل با تفاسیر معتبر اسلامی هم‌پوشانی داشت. در مقابل، در نمونه‌های ناصحیح، علل خطا اغلب به یکی از سه عامل بازمی‌گشت: نخست، ناتوانی در بازشناسی دقیق واژگان کلیدی (مانند ترجمه نادرست واژه‌های «رحمت» و «عذاب»)، دوم، اتکا به دانش متنی غیر اسلامی (مانند تطبیق مفاهیم قرآنی با فلسفه مدرن غربی) و سوم، گرایش به ساده‌سازی بیش از حد مفاهیم پیچیده.

این نتایج نشان می‌دهد که نقاط قوت و ضعف مدل‌های هوش مصنوعی مولد نه تنها باید در سطح کمی، بلکه در بستر کیفی و تحلیلی نیز بررسی شوند. بنابراین، یافته‌های این پژوهش صرفاً گزارش‌گر درصد صحت یا خطا نیستند، بلکه تصویری جامع از میزان انطباق، عمق و چالش‌های موجود در کاربرد AI برای علوم اسلامی را آشکار می‌سازند.

برای درک بهتر کارایی مدل‌های هوش مصنوعی مولد در حوزه متون اسلامی، لازم بود نتایج به تفکیک منابع اصلی (قرآن و حدیث) تحلیل شوند. این تفکیک نشان داد که مدل‌ها در مواجهه با متون قرآنی عملکرد بهتری داشته‌اند، در حالی که در متون حدیثی درصد خطا بیشتر بوده است. دلیل اصلی این تفاوت را می‌توان در ساختار زبانی و سبک متفاوت این دو دسته متون جستجو کرد. قرآن دارای انسجام لفظی و محتوایی بالاتری است، در حالی که متون حدیثی گاه به صورت پراکنده و با زمینه‌های متنوع نقل شده‌اند.

جدول ۴. مقایسه صحت و خطا در پاسخ‌های مدل به تفکیک منابع اسلامی

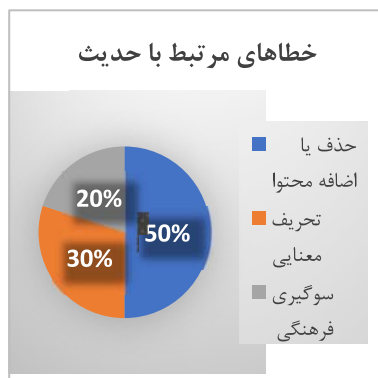
منبع	درصد پاسخ صحیح	درصد پاسخ خطا دار	خطای غالب
قرآن کریم	۷۰٪	۳۰٪	تحریف معنایی (۲۰٪)
احادیث	۵۰٪	۵۰٪	حذف یا اضافه محتوا (۲۵٪)

همان‌گونه که جدول ۴ نشان می‌دهد، میزان صحت پاسخ‌ها در متون قرآنی حدود ۲۰ درصد بیشتر از متون حدیثی است. دلیل اصلی این تفاوت به ویژگی‌های زبانی و ساختاری متون

بازمی‌گردد: قرآن از انسجام لفظی و معنایی بالاتری برخوردار است، در حالی که احادیث در نسخه‌های مختلف و با زمینه‌های متنوع نقل شده‌اند. بنابراین، طراحی الگوریتم‌های بومی‌سازی شده برای پردازش متون روایی اهمیت بیشتری دارد.

به منظور تکمیل تحلیل، خطاها نه تنها بر اساس نوع (تحریف، حذف/اضافه، سوگیری)، بلکه به تفکیک منابع نیز در قالب شکل ۵ نمایش داده شدند.

شکل ۵: نمودار توزیع خطاها به تفکیک قرآن و حدیث



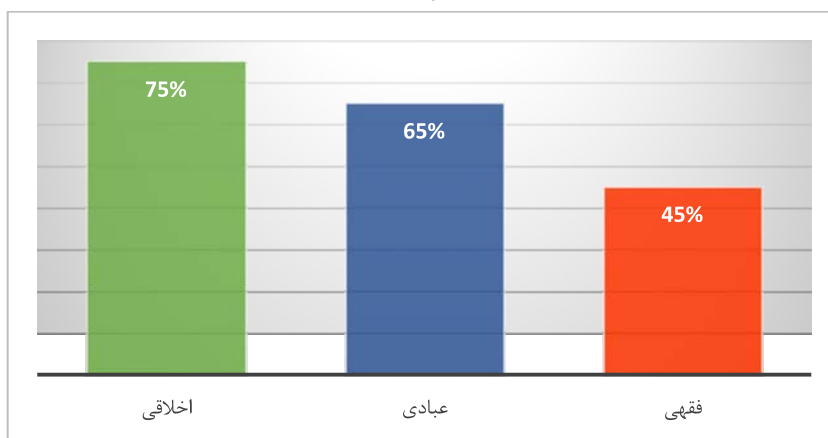
این نمودار نشان می‌دهد که خطاهای مرتبط با قرآن بیشتر در حوزه «تحریف معنایی» متمرکزند، در حالی که در احادیث، «حذف یا اضافه محتوا» بیشترین سهم را دارد. این تفاوت ناشی از ماهیت متون است: در قرآن، واژگان کلیدی مانند «تقوا» یا «رحمت» دچار تغییر معنایی می‌شوند؛ اما در احادیث، مدل به دلیل پیچیدگی ساختار روایت، گاهی بخش‌هایی از متن را حذف یا محتوای بیگانه اضافه می‌کند. این یافته بر ضرورت توجه به سبک‌های زبانی متفاوت در آموزش مدل‌ها تأکید دارد.

در گام بعدی، یافته‌های پژوهش نشان داد که دقت پاسخ‌های مدل‌ها در موضوعات مختلف نیز تفاوت چشمگیری دارد. به همین منظور، پرسش‌ها بر اساس سه حوزه «عبادی»، «اخلاقی» و «فقهی» طبقه‌بندی شدند.

جدول ۵: میزان صحت و خطا به تفکیک حوزه‌های موضوعی

حوزه	درصد پاسخ صحیح	درصد پاسخ خطا دار	خطای غالب
عبادی (نماز، روزه و...)	۶۵%	۳۵%	تحریف معنایی
اخلاقی (تقوا، عدالت، راستگویی و...)	۷۵%	۲۵%	سوگیری فرهنگی
فقهی (احکام معاملات، ارث و...)	۴۵%	۵۵%	حذف یا اضافه محتوا

همان‌طور که جدول (۵) نشان می‌دهد، مدل‌ها در حوزه اخلاقی عملکرد بهتری داشته و بیشترین دقت را ثبت کرده‌اند، در حالی‌که در حوزه فقهی درصد خطا به بیش از ۵۰ درصد می‌رسد. این امر بیانگر آن است که مفاهیم اخلاقی به دلیل عمومیت بیشتر و نزدیکی به زبان روزمره، توسط مدل‌های هوش مصنوعی آسان‌تر پردازش می‌شوند، اما در حوزه فقهی که نیازمند دقت بالا، توجه به جزئیات و تسلط بر اصطلاحات خاص است، ضعف مدل‌ها بیشتر آشکار می‌گردد. برای نمایش تصویری این یافته‌ها، شکل ۶ ارائه می‌شود.



شکل ۶: نمودار درصد صحت پاسخ‌های مدل‌ها به تفکیک حوزه‌های موضوعی

یافته‌ها نشان می‌دهد که مدل در حوزه اخلاقی بهترین عملکرد (۷۵٪ صحت) و در حوزه فقهی ضعیف‌ترین عملکرد (۴۵٪ صحت) را دارد. دلیل این تفاوت آن است که مفاهیم اخلاقی به دلیل عمومیت و نزدیکی به زبان روزمره، برای مدل‌های زبانی قابل درک‌تر هستند. در مقابل، حوزه فقهی به استنباط‌های دقیق، اصطلاحات تخصصی و توجه به جزئیات نیاز دارد؛ اموری که برای مدل‌های عمومی چالش‌برانگیز است. این نتیجه نشان می‌دهد که توسعه مدل‌های بومی با تمرکز ویژه بر متون فقهی و اصولی، یک ضرورت پژوهشی است.

در نهایت، تحلیل کیفی پاسخ‌های صحیح و ناصحیح نشان داد که مدل‌ها در بسیاری از موارد می‌توانند مفاهیم کلی و عمومی متون اسلامی را به‌خوبی بازتولید کنند، اما در مواجهه با جزئیات تخصصی (به‌ویژه در حوزه فقه) یا مفاهیم چندلایه (مانند «ولایت» یا «شورا») دچار خطا می‌شوند. این مسئله می‌تواند جهت‌گیری پژوهش‌های آینده را مشخص سازد؛ یعنی حرکت به سمت توسعه لایه‌های تخصصی و طراحی سازوکارهای کنترلی چندسطحی برای ارتقای دقت پاسخ‌ها.



شکل ۷: دیاگرام فرآیند تحلیل یافته‌ها

این شکل فرآیند تحلیل یافته‌ها در پژوهش را نشان می‌دهد. همان‌طور که در دیاگرام مشاهده می‌شود، مسیر از ورودی پرسش پژوهشگر آغاز شده و پس از پردازش توسط مدل هوش مصنوعی و تولید پاسخ اولیه، خروجی با منابع اسلامی معتبر مقایسه می‌شود. سپس، صحت یا خطا شناسایی شده و بر اساس سه دسته اصلی شامل تحریف معنایی، حذف یا اضافه محتوا و سوگیری

فرهنگی طبقه‌بندی می‌گردد. در نهایت، پس از این مراحل، ارزیابی نهایی صورت گرفته و یافته‌های پژوهش استخراج می‌شوند. این فرایند به خواننده کمک می‌کند تا درک کند چگونه داده‌ها از مرحله خام تا سطح تحلیل نهایی طی مسیر کرده‌اند و چگونه یافته‌ها به دست آمده‌اند. به طور کلی، نتایج نشان دادند که مدل‌های مولد در حوزه متون اسلامی علاوه بر ظرفیت‌های پردازشی چشمگیر، همچنان با خطاهای معنایی و سوگیری فرهنگی مواجه‌اند. این یافته‌ها مبنای تحلیل‌های تفصیلی در بخش بحث و نتیجه‌گیری خواهد بود.

بحث و نتیجه‌گیری

یافته‌های این پژوهش نشان داد که به کارگیری مدل‌های هوش مصنوعی مولد در متون اسلامی، در عین برخورداری از ظرفیت‌های نوین در پردازش زبان و بازیابی داده‌ها، همچنان با محدودیت‌های جدی مواجه است. نتایج تحقیق نشان داد که سهم عمده خطاهای مدل در سه حوزه «تحریف معنایی»، «حذف یا اضافه محتوا» و «سوگیری فرهنگی» متمرکز است. این الگو با یافته‌های پژوهش‌های پیشین هم‌خوانی دارد؛ به عنوان نمونه، (Abid et al. و Caliskan et al. (2021) (2017) بر وجود سوگیری‌های سیستماتیک در مدل‌های زبانی تأکید کرده‌اند. در فضای داخلی نیز، منتصری (۱۳۹۹) و محمدموسایی و همکاران (۱۴۰۰) هشدار داده‌اند که بهره‌گیری از AI در علوم اسلامی بدون چارچوب کنترلی، می‌تواند به تحریف مفاهیم یا تضعیف استقلال پژوهشگر منجر شود. این مقایسه نشان می‌دهد که مسئله سوگیری و تحریف معنایی صرفاً یک چالش فنی نیست، بلکه باید در چارچوبی میان‌رشته‌ای و با رویکرد معرفتی و اخلاقی تحلیل شود. از منظر تحلیلی، یافته‌های این پژوهش ضرورت ایجاد یک چارچوب اخلاقی-فنی یکپارچه را برجسته می‌سازد. هرچند مدل‌های مولد توانسته‌اند برخی فرایندهای گردآوری و بازنمایی داده‌های اسلامی را ساده‌تر کنند، نبود لایه‌های نظارتی معرفتی و اخلاقی می‌تواند پیامدهایی جدی برای اعتبار علوم اسلامی داشته باشد. استفاده از AI در حوزه دینی بدون توجه به مقاصد شریعت و اصول اخلاق اسلامی، خطر انتقال الگوهای فکری ناسازگار و تضعیف مرجعیت دینی را در پی دارد. یافته‌های این پژوهش تأکید می‌کند که طراحی لایه‌های کنترلی برای ارزیابی صحت، کاهش سوگیری و تضمین اخلاق‌مداری، جزء جدایی‌ناپذیر هر معماری بومی AI در علوم اسلامی است.

بر اساس این نتایج، چند پیشنهاد کلیدی برای توسعه و به کارگیری ابزارهای هوش مصنوعی در علوم اسلامی ارائه می شود:

۱. طراحی مدل های بومی سازی شده بر پایه پایگاه های داده معتبر اسلامی (قرآن، حدیث، متون فقهی و تفسیری) برای کاهش سوگیری فرهنگی.
 ۲. ایجاد لایه های کنترلی چندسطحی از طریق ترکیب ارزیابی فنی با بازبینی پژوهشگران دینی، به منظور جلوگیری از تحریف یا حذف محتوا.
 ۳. تدوین چارچوب اخلاقی-فقهی مبتنی بر مقاصد شریعت و ارزش های اسلامی، مشابه پیشنهاد های (Hakim & Shamsuddin (2024) و (Qayyum & Saifullah (2025).
 ۴. توانمندسازی پژوهشگران علوم اسلامی با ارتقای سواد دیجیتال و آموزش درباره فرصت ها و تهدید های AI در تحلیل متون.
 ۵. توسعه ابزارهای هوشمند اسلامی که علاوه بر پردازش زبانی، قابلیت کنترل خودکار صحت و سوگیری متون دینی را داشته باشند.
- این پژوهش با ارائه یک مدل مفهومی پنج لایه برای ادغام AI در علوم اسلامی، تلاشی در جهت پر کردن خلأ ادبیات علمی - یعنی فقدان چارچوبی منسجم برای پوشش هم زمان ابعاد فنی، معرفتی و اخلاقی - انجام داده است. نتایج نشان می دهد که آینده علوم اسلامی دیجیتال در هم افزایی میان توانمندی های AI و مرجعیت معرفتی و اخلاقی نهفته است، نه در تقابل یا کنار گذاشتن فناوری. چنین هم افزایی می تواند مسیر تازه ای برای پژوهش های میان رشته ای در جهان اسلام بگشاید و الگویی برای توسعه ابزارهای بومی متناسب با ارزش های دینی ارائه کند.
- با این حال، تداوم این مسیر پژوهشی مستلزم تمرکز بر دو محور کلیدی در آینده است: نخست، بررسی طولی و تجربی عملکرد مدل های AI در مواجهه با متون اسلامی متنوع، برای ارزیابی پایداری و کیفیت نتایج؛ دوم، طراحی معماری های ترکیبی انسان-ماشین که در آن پژوهشگر دینی نه ناظر منفعل، بلکه شریکی فعال در فرایند تحلیل و تولید معنا باشد. تحقق این افق ها می تواند به غنای ادبیات میان رشته ای یاری رسانده و جایگاه علوم اسلامی را در عصر دیجیتال تثبیت و تقویت کند.

فهرست منابع

- مرکز تحقیقات کامپیوتری علوم اسلامی. (۱۴۰۴). هوش مصنوعی در خدمت پژوهشگران علوم اسلامی: ابزارها و کاربردها. قم: مرکز تحقیقات کامپیوتری علوم اسلامی (نور).
- حسنی آهنگر، محمدرضا. (۱۴۰۱). نقش هوش مصنوعی در پژوهش‌های علوم اسلامی و انسانی. سیزدهمین همایش علوم اسلامی و انسانی دیجیتال.
- نصیری، مهدی. (۱۳۹۹). هوش مصنوعی و رویکردهای اسلامی. کنفرانس ملی هوش مصنوعی و علوم اسلامی. مقدم‌فر، عطیه. (۱۴۰۳). فرصت‌ها و چالش‌های هوش مصنوعی در مطالعات علوم انسانی و اسلامی. اولین همایش تعامل انسان و رایانه.
- محمد موسایی، زینب؛ کریمی‌نیا، محمد مهدی؛ انصاری‌مقدم، مجتبی. (۱۴۰۰). تحلیلی بر کارکردهای هوش مصنوعی در علوم اسلامی. نشریه پیشرفت‌های نوین در روانشناسی، علوم تربیتی و آموزش و پرورش، ۱۰۸-۱۲۳.
- مرکز مدیریت حوزه‌های علمیه. (۱۳۹۹). مجموعه مقالات همایش ملی هوش مصنوعی و علوم اسلامی. قم: مرکز مدیریت حوزه‌های علمیه.
- ربیع‌زاده، احمد. (۱۴۰۳). مروری بر علوم اسلامی دیجیتال در غرب. فصلنامه ره‌آورد نور.
- کوشکی، محمدحسن؛ موسوی، مهدی. (۱۴۰۳). تحول در مطالعات دفاع مقدس با هوش مصنوعی: رویکردی نوین به تحلیل و تفسیر داده‌های تاریخی. فصلنامه علمی مطالعات دفاع مقدس.
- منتصری، محمد مهدی. (۱۳۹۹). اخلاق هوش مصنوعی: چالش‌ها، الزامات و پیشنهادها. همایش ملی علوم اسلامی و هوش مصنوعی.
- معزی‌فر، هادی؛ رضائی، محمدجواد. (۱۴۰۴). تأثیر هوش مصنوعی بر تغییرات هویتی و مفهوم انسان در فرهنگ معاصر. همایش علوم انسانی و هوش مصنوعی. بازیابی از <https://civilica.com/doc/2314154>.
- جوزی، حسین؛ بشیری موحد، محمدحسن. (۱۳۹۹). بازآفرینی سلسله‌مراتب موضوعی مقالات نورمگز. همایش ملی علوم اسلامی و هوش مصنوعی.
- پوربدخشان، محمد هادی. (۱۳۹۹). هستان‌شناسی، گام نخست در ادراک ماشینی به همراه معرفی نرم‌افزار «جمله‌پرداز». نخستین همایش ملی هوش مصنوعی و علوم اسلامی. قم. بازیابی از <https://civilica.com/doc/1347075>.
- رجبی، زینب؛ ولوی، محمدرضا. (۱۳۹۹). پایگاه دانش عرفی در زبان فارسی. نخستین همایش ملی هوش مصنوعی و علوم اسلامی. قم. بازیابی از <https://civilica.com/doc/1347077>.

References

- Shabaz Patel, H, Kane, H, & Rayhan Patel, R. (2023). Building Domain-Specific LLMs Faithful To The Islamic Worldview: Mirage or Technical Possibility? arXiv preprint arXiv: 2312.06652.
- Abid, A, Farooqi, M, & Zou, J. (2021). Persistent Anti-Muslim Bias in Large Language Models. arXiv preprint arXiv: 2101.05783.
- Caliskan, A, Bryson, J. J, & Narayanan, A. (2017). Semantics derived automatically from language corpora contain human-like biases. *Science*, 356(6334), 183–186.
- Al-Salih Journal. (2024). Artificial Intelligence in Islamic Studies: Exploring Opportunities and Addressing Challenges. *Al-Salih*, 1(2), 1–15.
- Salim, M. A, & Aditya, R. B. (2025). Integration of Artificial Intelligence in Islamic Education: Trends, Methods, and Challenges in the Digital Era. *Journal of Modern Islamic Studies and Civilization*, 3(01), 74–89. <https://doi.org/10.59653/jmisc.v3i01.1368>.
- Mustapha, A, Senik@ Kamaruddin, Z, & Che Haron, S. (2025). An Overview of Artificial Intelligence (AI) Issues from the Perspective of Islamic Jurisprudence. *Online Journal of Research in Islamic Studies*, 12(1), 95–114. <https://doi.org/10.22452/ris.vol12no1.6>
- Mukmin Hakim, M. K. I, & Mat Jubri Shamsuddin, M. (2024). Maqasid Al-Shariah in the Age of AI: A Critical Examination of ChatGPT Usage Among International Islamic University Malaysia Students. *Online Journal of Research in Islamic Studies*, 11(2), 01–18. <https://doi.org/10.22452/ris.vol11no2.1>
- Mohadi, M, & Tarshany, Y. (2023). Maqasid Al-Shari'ah and the Ethics of Artificial Intelligence: Contemporary Challenges. *Journal of Contemporary Maqasid Studies*, 2(2), 79–102. <https://doi.org/10.52100/jcms.v2i2.107>
- Akhtar, K. (2024). An Analytical Study of Artificial Intelligence and Ethical Principles in the Light of Qur'anic Teachings. *Journal of Religion and Society*, 2(4), 711–721. Retrieved from <https://www.islamicreligious.com/index.php/Journal/article/view/113>
- Abdul Qayyum, & Saifullah. (2025). Artificial Intelligence and the Quran: Ethical Boundaries in Light of Islamic Teachings. *Journal of Religion*

- and Society, 4(01), 19–34. Retrieved from, <https://www.islamicreligious.com>
- Khanzada Muhammad Waqar, M, Ibrahim, M, & Iltimas Khan, M. (2025). Ethical Implications of Artificial Intelligence: An Islamic Perspective. *Journal of Religion and Society*, 3(01), 347–358. Retrieved from. <https://www.islamicreligious.com>.
- Abdelnour, M. G. (2025). Artificial Intelligence and the Islamic Theology of Technology: From “Means” to “Meanings” and from “Minds” to “Hearts”. *Religions*, 16(6), 796. <https://doi.org/10.3390/rel16060796>
- Corpus Coranicum. (n.d.). In Wikipedia. Retrieved August 2025, from <https://en.wikipedia.org/wiki/Corpus-Coranicum>
- ArchNet. (n.d.). In Wikipedia. Retrieved August 2025, from <https://en.wikipedia.org/wiki/ArchNet>
- Islamic Scientific Manuscripts Initiative (ISMI). (n.d.). In Wikipedia. Retrieved August 2025, from <https://en.wikipedia.org/wiki/Islamic-Scientific-Manuscripts-Initiative>
- Al-Faruqi, I. R. (1982). *Islamization of Knowledge: General Principles and Work Plan*. Herndon, VA: International Institute of Islamic Thought.
- Journal of Arabic and Islamic Studies*. (n.d.). In Wikipedia. Retrieved August 2025, from <https://en.wikipedia.org/wiki/Journal-of-Arabic-and-Islamic-Studies>
- Reddit. (2024). Open-Access peer-reviewed papers in Quranic and Islamic studies. Retrieved from, <https://www.reddit.com>
- Reddit. (2024). Academic Journals and Book Publishers. Retrieved from <https://www.reddit.com/r/AcademicQuran/comments/nck1u0>
- Reddit. (2024). Open-access books in Qur’anic and Islamic studies. Retrieved from <https://www.reddit.com/r/AcademicQuran/comments/1635ups>.