



۱-۲۰	صفحات	دوره انتشار	شماره انتشار	۱۴۰۵	سال انتشار
------	-------	-------------	--------------	------	------------

چارچوب هوش مصنوعی عامل محور برای تحلیل فقه حکومتی اسلامی با رویکرد دفاع شناختی : مطالعه ولایت فقیه

خانم دکتر سعیده سادات سدیدپور، جناب آقای مهدی غلامی مهر و آقای امیرحسین برادران

گروه هوش مصنوعی، دانشکده برق و کامپیوتر، دانشگاه صنعتی مالک اشتر، تهران، ایران

گروه هوش مصنوعی، دانشکده برق و کامپیوتر، دانشگاه صنعتی مالک اشتر، تهران، ایران

گروه هوش مصنوعی، دانشکده برق و کامپیوتر، دانشگاه صنعتی مالک اشتر

چکیده

این پژوهش یک چارچوب چندلایه مبتنی بر هوش مصنوعی برای تحلیل مفاهیم فقهی حکومتی که یک کاربرد نوین به شمار می‌رود معرفی می‌کند. چارچوب پیشنهادی شامل سه مؤلفه اصلی است: زنجیره تفکر برای تحلیل گام به گام استدلال‌ها، ساختار مبتنی بر عامل برای ارزیابی دیدگاه‌های فقهی و درخت تفکر برای انتخاب پاسخ‌های منطقی‌تر. این چارچوب با تنظیمات بهینه شده بر روی پنج مدل زبانی پیشرفته شامل GPT-۳.۵، GPT-۴o، Claude ۳ Opus، Gemini ۲.۵ Pro، Grok ۳ و عملکرد مدل‌ها با استفاده از پنج شاخص کلیدی شامل دقت ارجاع به منابع، عمق استدلال، تنوع دیدگاه‌ها، قوت داوری و تعامل‌پذیری تحلیل گردید. نتایج نشان داد که چارچوب پیشنهادی موجب بهبود چشمگیر در دقت منابع، انسجام استدلال‌ها و درک زمینه‌ای شده است. مدل GPT-۴o با کسب امتیاز ۸۳.۲ از ۱۰۰ بالاترین عملکرد را داشت. این چارچوب قابلیت تعمیم به سایر مفاهیم فقهی و دینی را دارد و می‌تواند در حوزه‌های حکمرانی اسلامی و پدافند نرم کاربرد داشته باشد.

واژگان کلیدی: هوش مصنوعی، چارچوب چندلایه، ساختار مبتنی بر عامل، زنجیره تفکر، فقه اسلامی

طبقه‌بندی: JEL فقه - حقوق - جزا و جرم‌شناسی - حقوق بین‌الملل - حقوق خصوصی



Agent-Based Artificial Intelligence Framework for Analyzing Islamic Governmental Jurisprudence with a Cognitive Defense Approach: A Study of the Guardianship of the Jurisprudent

Abstract

This study presents a multi-layered artificial intelligence framework for analyzing Islamic governance jurisprudential concepts, which is a novel application. The proposed framework consists of three main components: Chain of Thought for step-by-step reasoning analysis, Agent-Based Structure for evaluating jurisprudential viewpoints, and Tree of Thought for selecting logically stronger responses. The framework was evaluated on five advanced language models, including GPT-4o, Claude 3 Opus, Gemini 2.5 Pro, Grok 3, and GPT-3.5, using optimized settings. The models performance was assessed based on five key indicators: accuracy of source references, depth of reasoning, diversity of viewpoints, soundness of judgment, and interactivity. The results showed that the proposed framework significantly improved source accuracy, argumentative coherence, and contextual understanding. The GPT-4o model achieved the highest performance with a score of 83.2 out of 100. This framework is generalizable to other jurisprudential and theological concepts and can be applied in the fields of Islamic governance and soft power defense.

Keywords: Artificial Intelligence, Multi-layered Framework, Agent-Based Structure, Chain of Thought, Islamic Jurisprudence



متن مقاله

این معماری در مدل‌های زبانی بزرگ مختلف مورد بررسی و ارزیابی قرار گرفته است. پرسش‌های طراحی‌شده، برگرفته از مباحث اخلاقی-فقهی و مفاهیم کلیدی همچون مشروعیت، اجتهاد و اختیار دینی هستند که در چارچوب نظریه ولایت فقیه صورت‌بندی شده‌اند. نتایج ارزیابی نشان می‌دهد که مدل‌های تنظیم‌شده با این چارچوب، در مقایسه با حالت عادی، توان استدلال عمیق‌تر، تحلیل متنوع‌تر و انسجام منطقی بالاتری را از خود نشان می‌دهند.

در نهایت، این پژوهش بر این اصل تأکید دارد که برای بهره‌گیری مؤثر از مدل‌های زبان بزرگ در حوزه‌های دینی و هنجاری، باید ساختارهای تحلیل‌محور و اعلان‌های دقیق و هدفمند طراحی شود؛ ساختارهایی که امکان درک واقعی، استدلال شفاف و پاسخ‌گویی هم‌راستا با منطق درون‌دینی را برای مدل فراهم کنند.

۱. کارهای مرتبط

مطالعه ارتباط میان هوش مصنوعی و مفاهیم دینی-فقهی مستلزم مرور ادبیاتی میان‌رشته‌ای است که جایگاه پژوهش حاضر را در بدنه جهانی این حوزه روشن کند. با توجه به آنکه موضوع این مقاله در تقاطع معماری مدل‌های زبانی، نظریه‌های مشروعیت دینی، و فلسفه اخلاق فناوری قرار دارد، منابع پیشینه در چهار محور کلیدی دسته‌بندی می‌شوند:

۱-۱. اخلاق هوش مصنوعی: از چارچوب‌های کلاسیک تا نقدهای بافت‌محور

ادبیات اولیه اخلاق هوش مصنوعی عمدتاً بر پایه نظریه‌های غربی مانند سودگرایی، وظیفه‌گرایی و قراردادی‌گرایی شکل گرفته‌اند (۷). این الگوها بر اصولی چون منفعت عمومی، خودمختاری فردی، و شفافیت الگوریتمی تأکید دارند. با این حال، از دهه ۲۰۲۰ به این سو، نقدهایی در خصوص

در دهه‌های اخیر، پیشرفت چشمگیر فناوری‌های هوش مصنوعی، به‌ویژه مدل‌های زبان بزرگ مانند GPT.Claude و Gemini بسترهای نوینی برای پردازش زبان طبیعی، تصمیم‌سازی اخلاقی و طراحی سامانه‌های هوشمند فراهم کرده‌اند (۱). این مدل‌ها با تکیه بر یادگیری عمیق و داده‌های گسترده، توانایی تولید پاسخ‌هایی پیچیده، منسجم و شبه‌انسانی را یافته‌اند. با این حال، چارچوب‌های استدلالی این مدل‌ها اغلب بر مبانی معرفتی غربی تکیه دارند که با منطق سنت‌های دینی، به‌ویژه فقه اسلامی، هم‌خوانی ندارند (۲).

در چنین فضایی، مدل‌های زبان بزرگ هنگام مواجهه با پرسش‌های دارای بار معنایی دینی یا اخلاقی عمیق، دچار سطحی‌نگری، تناقض‌گویی، یا پاسخ‌های فاقد انسجام فقهی می‌شوند (۳). این ناتوانی عمدتاً ناشی از نبود معماری‌های مفهومی چندلایه و فقدان تنظیمات اعلان‌محور هدفمند برای درگیر ساختن مدل با استدلال‌های بنیادین است. به عبارت دیگر، مدل‌های موجود فاقد توانایی لازم برای تحلیل لایه‌به‌لایه مفاهیم انتزاعی، شناسایی دیدگاه‌های متعارض و داوری معتبر میان استدلال‌ها هستند (۴). در پاسخ به این خلأ، پژوهش حاضر با تمرکز بر هوش مصنوعی زبان‌محور و طراحی معماری چندلایه تفسیری، چارچوبی نوآورانه برای تعامل مدل‌های زبانی با مفاهیم دینی، به‌ویژه نظریه «ولایت فقیه» در اندیشه شیعی (۵)، ارائه می‌دهد. این چارچوب، از سه مؤلفه مکمل تشکیل شده است:

- زنجیره تفکر [۱]: تحلیل مرحله‌به‌مرحله و ساختاری پرسش‌ها
- ساختار عامل محور [۲]: ارزیابی هم‌زمان دیدگاه‌های مختلف فقهی و فلسفی
- درخت تفکر [۳]: پیمایش مسیرهای استدلالی و انتخاب پاسخ نهایی بر پایه اعتبار استدلال‌ها (۶).



بردند و دقتی قابل قبول در شناسایی توییت‌های مرتبط با آشوب‌های اجتماعی گزارش کردند (۱۲). در مطالعه‌ای دیگر، داداش‌تبار احمدی و همکاران، با استفاده از خوشه‌بندی رفتاری و تحلیل محتوای زبانی، الگوهای حملات سطح بالا در فضای سایبری را با دقت بالاتری نسبت به روش‌های کلاسیک شناسایی کردند (۱۳).

شیخی و همکاران نیز در تحقیقی جداگانه، سامانه‌ای برای تحلیل احساسات کاربران فارسی‌زبان طراحی کرده‌اند که از ترکیب پردازش زبان طبیعی و یادگیری ماشین برای تفسیر داده‌های متنی بهره می‌برد (۱۴). همچنین، قاسمی و دهقانی با تمرکز بر استخراج خودکار مفاهیم کلیدی از اسناد فارسی، نشان دادند که روش‌های NLP قابلیت بالایی در تحلیل محتوای متنی و بازبایی اطلاعات راهبردی دارند (۱۵).

با وجود این تلاش‌ها، بیشتر پژوهش‌های یادشده بر تحلیل اجتماعی، امنیت نرم یا رفتار کاربران تمرکز دارند و هنوز ساختارهای مفهومی پیچیده‌ای چون اجتهاد، مشروعیت فقیه، یا نیابت شرعی به‌صورت نظام‌مند در آن‌ها بازنمایی نشده‌اند. پژوهش حاضر می‌کوشد این خلأ را با طراحی چارچوبی مبتنی بر معماری چندلایه در تعامل با مدل‌های زبان بزرگ، برای بازنمایی مفاهیم فقهی در بستر هوش مصنوعی، برطرف سازد.

۴-۱. جایگاه پژوهش حاضر در ادبیات میان‌رشته‌ای

پژوهش حاضر با طراحی یک چارچوب سه‌لایه‌ای و ارزیابی عملکرد الگوریتمی آن در مدل‌های زبانی، گامی فراتر از تحلیل نظری برمی‌دارد:

- از منظر مفهومی: چارچوبی برای بازنمایی مشروعیت دینی در معماری مدل‌های زبانی پیشنهاد می‌دهد.

- از منظر فنی: امکان ارتقاء توان تحلیل استدلالی و پاسخ‌گویی دقیق مدل را در مواجهه با مفاهیم فقهی نشان می‌دهد.

این مطالعه از نخستین گام‌های طراحی هوش مصنوعی دین‌محور در

جهانی‌سازی اخلاق لیبرال در طراحی مدل‌های هوش مصنوعی پدید آمد. پژوهشگران بر ضرورت توسعه «الگوهای اخلاقی بافت‌محور» تأکید کردند. رویکردهایی که ویژگی‌های فرهنگی و دینی جوامع مختلف را در طراحی سامانه‌های هوشمند لحاظ می‌کنند (۸).

۱-۲. مشروعیت دینی در برابر الگوهای سکولار حکمرانی

در سنت‌های دینی چون اسلام، یهودیت و مسیحیت، مشروعیت از منابع الهی، صلاحیت اخلاقی رهبر و پیوند با سنت‌های قدسی مشتق می‌شود (۹). این در تضاد با نظریه‌های سیاسی مدرن مانند رضایت عمومی یا قرارداد اجتماعی است که به رأی مردم و قانون عرفی تکیه دارند. نظریه ولایت فقیه در فقه شیعه، نمونه‌ای ساختارمند از مشروعیت دینی است که هنوز در مدل‌سازی‌های هوش مصنوعی به‌طور جدی بازنمایی نشده. بیشتر مدل‌های زبانی موجود فاقد توانایی درک و تحلیل چنین ساختارهایی هستند.

۱-۳. شکاف مفهومی و الگوریتمی در بازنمایی دین در هوش مصنوعی

درحالی‌که برخی پژوهش‌ها تلاش کرده‌اند مفاهیم دینی را در حیطه‌هایی مانند حریم خصوصی یا رفع تبعیض وارد سازند (۱۰)، مفاهیم پیچیده‌ای چون اجتهاد، مشروعیت معنوی و نیابت فقهی همچنان در ادبیات علمی مغفول مانده‌اند. پروژه‌هایی نظیر "AI & Faith" گرچه گامی ابتدایی در این مسیر محسوب می‌شوند، اغلب توصیفی‌اند و عمدتاً به سنت‌های غربی محدود شده‌اند (۱۱). در فضای داخلی نیز طی سال‌های اخیر، پژوهش‌هایی با بهره‌گیری از روش‌های یادگیری ماشین و پردازش زبان طبیعی در حوزه تحلیل رفتارهای اجتماعی و امنیت سایبری شکل گرفته‌اند. برای نمونه، عباسی و جوادزاده با طراحی یک چارچوب پیش‌بینی ناآرامی‌های مردمی در شبکه‌های اجتماعی فارسی‌زبان، از مدل‌هایی چون Naïve Bayes و SVM برای تحلیل احساسات بهره



قابل محاسبه و تعاملی فراهم می آورد.

۱-۲. ساختار سه لایه ای مدل: تعامل تحلیل های متعدد

چارچوب طراحی شده برای این پژوهش بر سه لایه اصلی استوار است که هر یک نقش خاصی در فرآیند تحلیل و تصمیم گیری دارند. این لایه ها به صورت مکمل و در تعامل با یکدیگر عمل می کنند تا مدل بتواند پاسخ هایی جامع، معتبر و منطقی ارائه دهد. در ادامه به بررسی این سه لایه می پردازیم:

الف) لایه اول: زنجیره تفکر

در این لایه، مدل به جای ارائه پاسخ های سطحی، باید یک تحلیل گام به گام از پرسش های فقهی انجام دهد. زنجیره تفکر فرآیند استدلال مدل را به صورت مرحله به مرحله هدایت می کند تا مطمئن شویم که مدل تمامی ابعاد مسئله را به طور دقیق و مستند بررسی کرده است (۱۸).

گام های این زنجیره به صورت زیر است:

۱. تعریف دقیق مسئله فقهی - سیاسی: مدل باید مسئله را به طور شفاف و واضح تعریف کند و آن را در قالب مفاهیم فقهی و سیاسی قرار دهد.

۲. تحلیل قرآنی: مدل باید آیات مرتبط با موضوع را به دقت تحلیل کرده و از آن ها برای پاسخ به پرسش استفاده کند.

۳. تحلیل روایی - حدیثی: در این مرحله، مدل به منابع حدیثی معتبر مثل توقیع امام زمان و روایات اهل بیت مراجعه می کند.

۴. تحلیل عقلی و فلسفه سیاسی اسلام: مدل باید براساس اصول عقلی و فلسفی اسلام، مبانی فقهی را تحلیل کرده و نشان دهد که چه نظریه ای از نظر عقلی و فلسفی معتبر است (۱۹).

۵. تحلیل اجتماعی - تاریخی: در این گام، مدل باید به تجربه تاریخی فقه و حکمرانی در جوامع اسلامی توجه کند و از آن ها برای تبیین موضوع استفاده نماید.

جهان اسلام به شمار می رود و کوششی عملی در جهت تلفیق نظریه پردازی فقهی با معماری مدل های زبانی بزرگ است. همچنین پاسخی به خلأ موجود در ادبیات فعلی برای مدل سازی مفاهیم دینی ساختاریافته در هوش مصنوعی به صورت آزمون پذیر ارائه می دهد.

۲. روش پیشنهادی

در این پژوهش، با هدف ارزیابی امکان بازنمایی مفاهیم انتزاعی و چندلایه نظریه فقهی - سیاسی (ولایت مطلقه فقیه) با استفاده از زبان طبیعی، چارچوبی ترکیبی مبتنی بر مدل زبان بزرگ طراحی شد. این چارچوب بر مبنای مهندسی اعلان چندعامله، زنجیره تفکر گام به گام و ساختار درخت تفکر طراحی شده و هدف آن ارتقاء توان مدل در تولید پاسخ های ساختاریافته، قابل استدلال و هماهنگ با منطق اجتهادی در پرسش های پیچیده است (۱۶).

این چارچوب، با بهره گیری از قابلیت های مدل در تحلیل زبان طبیعی، مدل سازی ذهنی و تولید زنجیره های استدلالی، تلاش می کند تا فرآیند اجتهاد، مشروعیت دینی و منطق فقهی را به صورت ساختاریافته، مرحله محور و قابل ارزیابی در محیط هوش مصنوعی بازنمایی کند (۱۷). ساختار پیشنهادی متشکل از سه لایه مکمل و در تعامل است:

۱) زنجیره تفکر (تحلیل گام به گام و جزئی نگر مسئله)

۲) ساختار عامل محور (ارزیابی هم زمان دیدگاه های معرفتی گوناگون از طریق شبیه سازی عوامل مستقل)

۳) درخت تفکر (توسعه ساختار درختی استدلال ها و انتخاب مسیر نهایی براساس مقایسه اعتبار تحلیل ها).

این طراحی چندلایه، مدل را قادر می سازد تا به پرسش های فقهی پیچیده با حفظ دقت منبع، انسجام منطقی و انسجام معرفتی پاسخ دهد. همچنین این ساختار، امکان بازنمایی نظام اجتهادی شیعه دوازده امامی را در قالبی



تحلیل سازگاری پاسخ با مبانی عقلی و فلسفه سیاسی اسلام

عامل فلسفی

سنجش پاسخ براساس واقعیت‌های تاریخی و تحولات سیاسی شیعه

عامل اجتماعی-تاریخی

افزودن لایه معنوی و هستی‌شناختی به تحلیل (در صورت امکان)

عامل عرفانی

لایه سوم: درخت تفکر

پس از انجام تحلیل‌های هر عامل، مدل موظف است که پاسخ‌ها را به‌طور

کامل و جامع ارائه دهد. این مرحله به‌عنوان مرحله تصمیم‌گیری نهایی

عمل می‌کند که در آن مدل از بین مسیرهای مختلف پاسخ، مسیر برتر را

انتخاب می‌کند. مدل در این لایه باید یک مسیر نهایی را براساس استدلال

برتر از میان گزینه‌های مختلف انتخاب کند (۲۲). این گزینه‌ها در سه

دسته کلی قرار می‌گیرند:

• مسیر قرآنی-روایی

• مسیر عقلی-فلسفی

• مسیر عرفانی-تفسیری (در صورت وجود)

هر مسیر تحلیلی باید شامل شواهد مستند، انسجام منطقی و لایه‌های

استدلالی متعدد باشد (۲۳). در پایان این فرآیند، مدل موظف است با

تحلیل نهایی خروجی عامل‌ها و مقایسه استدلال‌ها، یک جمع‌بندی جامع

ارائه داده و براساس قوت تحلیلی، مسیر برتر را به صورت شفاف و مستدل

انتخاب نماید.

به‌عنوان سیاست راهبردی در ساختار تحلیل، این چارچوب به‌گونه‌ای

طراحی شده که امکان آزمون ترجیح یک دیدگاه از میان دیدگاه‌های

متعارض را با حفظ ساختار استدلالی اجتهادی فراهم کند. بدین منظور،

۶. نگاه عرفانی: اگر مسئله به جنبه‌های عرفانی و معنوی مرتبط باشد،

مدل باید آن را بررسی کرده و تحلیل‌های عرفانی مرتبط با ولایت فقیه را

نیز در نظر بگیرد.

۷. جمع‌بندی نهایی: در این مرحله، مدل تمامی تحلیل‌ها را جمع‌بندی

کرده و نتیجه‌گیری نهایی را ارائه می‌دهد.

این فرآیند، پایه‌گذار محتوای اولیه است که توسط مدل تولید می‌شود و

بعداً باید مورد ارزیابی و تصحیح قرار گیرد (۲۰).

(ب) لایه دوم: ساختار عامل‌محور

پس از تولید پاسخ اولیه در لایه زنجیره تفکر، نوبت به ورود عامل‌ها

می‌رسد. عامل‌ها مشابه ناظران متخصص عمل می‌کنند و هر کدام بخشی

از پاسخ را به طور دقیق بررسی می‌کنند. این لایه به مدل کمک می‌کند

تا از دیدگاه‌های مختلف به تحلیل پاسخ پرداخته و آن را پالایش کند

(۲۱). در این ساختار، عامل‌ها در یک تعامل هم‌زمان با یکدیگر عمل کرده

و هیچ‌یک به‌طور مجزا تحلیل نمی‌کنند. به عبارت دیگر، پس از انجام

تحلیل اولیه در زنجیره تفکر، عامل‌ها بررسی‌های تخصصی خود را به‌طور

هم‌زمان آغاز می‌کنند:

جدول ۱. ساختار وظیفه‌محور عامل‌های تحلیلی در لایه دوم چهارچوب

پیشنهادی

نقش تحلیلی

عامل

بررسی دقت و کفایت استناد به آیات قرآن

عامل قرآنی

ارزیابی اعتبار و جامعیت منابع روایی

عامل حدیثی



مدل منتقل می‌شود. اعلان‌ها به صورت چندلایه طراحی گردیدند و بر پایه سه بخش تحلیلی اصلی که پیش‌تر معرفی شدند، شکل گرفتند.

هر اعلان بر مبنای چهار مؤلفه اصلی تنظیم شده است :

- نقش : تعریف شخصیت تحلیلی مدل (مانند فقیه، مفسر، یا مورخ)
- وظیفه : تعیین نوع تحلیل مورد انتظار (فقهی، کلامی، تاریخی و ...)
- سبک : تعیین لحن و قالب پاسخ (تحلیلی، مستند و رسمی)
- محدودیت : اعمال قیود مانند ارجاع معتبر، پرهیز از تعمیم و حفظ

انسجام مفهومی

در مرحله نخست، نقش مدل به عنوان «دستیار فقهی با دیدگاه روشن نسبت به نظریه ولایت مطلقه» تعریف شد. سپس مجموعه‌ای از اعلان‌های مرحله‌ای، متناسب با حوزه‌های تحلیلی مختلف، به مدل تزریق گردید. این ساختار امکان فعال‌سازی عامل‌های تحلیلی مستقل را فراهم آورد و در نهایت، پاسخ نهایی از طریق تجمیع خروجی‌های مجزای هر عامل، با حفظ هماهنگی و سازگاری با منطق اجتهادی تولید شد.

۳. ارزیابی نتایج

۳-۱. هدف آزمایش

هدف اصلی این آزمایش، ارزیابی توانایی روش پیشنهادی در بازنمایی مفاهیم کلیدی نظریه «ولایت مطلقه فقیه» بر پایه چارچوب سه‌لایه‌ای است. تمرکز اصلی بر بررسی عمق تحلیل، دقت منابع و توان داوری مدل در مسائل اجتهادی-مفهومی است.

۳-۲. طراحی و دسته‌بندی سؤالات

برای آزمایش مدل، مجموعه‌ای از ۱۰ پرسش کلیدی از دل مباحث مهم فقه سیاسی شیعه استخراج شد. این سؤالات به صورت هدفمند و با الهام از چارچوب‌های اجتهادی کلاسیک و رویکردهای اخیر در ترکیب منطق فقهی با ارزیابی مدل‌های زبانی بزرگ در سه دسته اصلی طراحی شدند.

چالش‌ها، شبهات یا دیدگاه‌های مخالف تنها در قالب سؤالات هدایت‌شده و به‌منظور پاسخ‌گویی، دفع یا تبیین آن‌ها مطرح می‌شوند، نه به عنوان مسیر جایگزین.

این رویکرد به مدل اجازه می‌دهد تا ضمن حفظ تنوع منابع، یکپارچگی نتیجه و انسجام نظری را در راستای دفاع از مشروعیت دینی- الهی و عقلانی این نظریه حفظ نماید. همان‌گونه که در شکل ۱ مشاهده می‌شود، مدل از یک ساختار مرحله‌ای برای تحلیل سؤال بهره می‌برد که در نهایت از میان مسیرهای تحلیلی متنوع، گزینه بهینه را برمی‌گزیند.

در این چارچوب، مسیر برتر نه تنها از نظر امتیاز تحلیلی بالاتر است، بلکه کمترین آسیب‌پذیری را در برابر نقدهای معرفتی دارد.

الگوریتم تصمیم‌یار مدل، با ترکیب شاخص‌های قدرت استدلال و میزان تردیدپذیری، مسیر نهایی را تعیین می‌کند.

این ترکیب، امکان داوری دقیق در مواجهه با دو یا چند قرائت اجتهادی را بدون ورود به جدال ایدئولوژیک فراهم می‌سازد.

در نتیجه، پاسخ نهایی مدل نه تنها از نظر منطقی مستحکم است، بلکه توان پاسخ‌گویی به تنوع موجود در دیدگاه‌های فقهی را نیز دارد.

شکل ۱: نمودار درخت تصمیم در معماری تحلیل GPT-۴O برای پرسش‌های فقهی-سیاسی

۲-۲. نحوه پیاده‌سازی چارچوب در مدل‌های زبان بزرگ

برای ارزیابی عملی ساختار پیشنهادی، مجموعه‌ای از مدل‌های زبانی پیشرفته با معماری‌ها و مقیاس‌های متنوع مورد استفاده قرار گرفت. این تنوع با هدف سنجش میزان تعمیم‌پذیری الگو و کاهش وابستگی به یک معماری خاص انتخاب شد.

پیاده‌سازی با استفاده از رویکرد مهندسی اعلان ترکیبی انجام گرفت. روشی که در آن تحلیل چندمرحله‌ای مسئله به صورت ساختاریافته به



در مواجهه با پرسش‌های فقهی - سیاسی پیچیده، پنج مدل پیشرفته شامل GPT-۴o، Claude ۳ Opus، Gemini ۲.۵ Pro، Grok۳ و GPT-۳.۵ مورد بررسی قرار گرفتند. این مدل‌ها در معرض مجموعه‌ای یکسان از ده پرسش مفهومی - اجتهادی قرار گرفتند که به صورت ساختاریافته و با بهره‌گیری از اعلان‌های مهندسی شده طراحی شده بودند، تا ابعاد گوناگون توان استدلالی، تحلیلی و اخلاقی آن‌ها آشکار گردد. در همین راستا، پنج شاخص کلیدی ارزیابی تعریف شد که در جدول زیر آمده است. این شاخص‌ها با تأثیر از چارچوب‌های معتبر ارزیابی عملکرد مدل‌های زبان بزرگ در زمینه‌های اخلاقی، تحلیلی و چندعامله طراحی شده‌اند و معیارهای سنجش دقت، انسجام، مسئولیت‌پذیری، درک بافت و تطابق با ارزش‌های زمینه‌ای را پوشش می‌دهند (۲۴). جدول ۳. معیارهای سنجش کیفیت پاسخ‌های مدل براساس چارچوب تحلیلی پیشنهادی

معیار ارزیابی

شاخص

ارجاع دقیق به آیات، احادیث و منابع معتبر فقهی

دقت منابع

استنتاج چندمرحله‌ای، تحلیل تفکیکی و انسجام درونی

عمق استدلال

بازتاب رویکردهای مختلف (مانند مطلقه / محدود، رأی‌محور / نص‌محور)

تنوع دیدگاه‌ها

انتخاب مسیر پاسخ مستدل و ترجیح یکی از دیدگاه‌ها با منطق اجتهادی

قدرت داوری

هماهنگی بین عامل‌ها، استفاده از منطق ترکیبی در تحلیل نهایی

جدول ۲. محورهای تحلیلی سؤالات آزمون مدل براساس چارچوب سه‌لایه‌ای

نوع سوالات

دسته

مشروعیت فقیه، نصب الهی، اطلاق ولایت، رابطه با عصر غیبت

مفهومی - کلامی

حدود اختیارات فقیه، رأی مردم، حسابیات، تفاوت ولایت مطلقه و محدود

فقهی - اجتهادی

دیدگاه علمای گذشته، عملکرد فقه در دولت، مقایسه دیدگاه‌ها

تاریخی - تحلیلی

نمونه‌ای از سؤالات استفاده شده :

- آیا ولایت مطلقه فقیه با عقل و قرآن قابل اثبات است؟
- تفاوت دیدگاه فقه‌های شیعه درباره ولایت چیست؟
- آیا رأی مردم در مشروعیت ولی فقیه نقش دارد؟
- چگونه می‌توان مفهوم نیابت عام را در مدل هوش مصنوعی بازنمایی کرد؟

۳-۳. نحوه تعامل با مدل

هر پرسش با یک اعلان مهندسی شده آغاز می‌شد که ساختار زنجیره تفکر و عامل‌محور را فعال می‌کرد. تعامل با مدل در چند مرحله و با کنترل دقیق اعلان‌ها صورت گرفت تا امکان پاسخ‌دهی سطحی از بین برود و ساختار تحلیلی حفظ شود. پاسخ‌ها به صورت متنی ثبت شدند.

۳-۴. تحلیل کیفی پاسخ‌های تولیدی

برای ارائه تحلیلی مقایسه‌ای و چندبعدی از عملکرد مدل‌های زبان بزرگ



قابلیت تعامل

GPT-۳.۵

Grok ۳

Gemini ۲.۵ Pro

Claude ۳ Opus

GPT-۴o

شاخص ارزیابی

۷۲

۷۸

۸۲

۸۶

۹۰

دقت منابع

۶۸

۷۶

۷۸

۸۰

۸۲

عمق استدلال

۶۵

۷۱

۷۲

۷۴

۷۶

تنوع دیدگاهها

پنج عامل ارزیاب متخصص هر پاسخ را براساس پنج شاخص کلیدی، از ۰ تا ۱۰۰ نمره‌دهی کردند. این روش ارزیابی با الگوبرداری از استانداردهای نوین تحلیل کیفی در آزمون سیستم‌های استدلال محور تنظیم شده است. امتیاز نهایی هر شاخص، میانگین مجموع نمرات هر عامل ارزیاب در آن معیار بود.

میانگین امتیازهای به‌دست‌آمده توسط هر مدل در این شاخص‌ها در جدول ۵ ارائه شده است. نتایج این جدول حاکی از آن است که مدل GPT-۴o در مجموع، در تمامی شاخص‌ها عملکرد بالاتری نسبت به سایر مدل‌ها داشته و توانسته است توازن مناسبی میان بازنمایی مفاهیم فقهی و ساختار منطقی پاسخ ایجاد کند. با این حال، باید توجه داشت که این ارزیابی‌ها محدود به مجموعه‌ای خاص از پرسش‌های اجتهادی هستند و لزوماً نمایانگر جامعیت کامل عملکرد مدل‌ها در همه مسائل دینی یا تفاسیر فقهی نیستند. به‌منظور ارزیابی فراگیرتر، لازم است دامنه موضوعات و قرائت‌های فقهی در پژوهش‌های آتی گسترش یابد. برای تسهیل درک مقایسه‌ای این تفاوت‌ها و تقویت خوانایی تحلیلی، داده‌های جدول به‌صورت یک نمودار میله‌ای نیز نمایش داده شده‌اند. این نمودار امکان مشاهده سریع و بصری اختلاف عملکرد مدل‌ها را در هر شاخص فراهم می‌سازد و به‌وضوح برتری نسبی GPT-۴o در تمامی شاخص‌ها را نسبت به سایر مدل‌ها نشان می‌دهد. بهره‌گیری از این نمایش تصویری به‌ویژه در تحلیل‌های چندمتغیره ضروری است، چراکه امکان ادراک توازن داخلی هر مدل را نیز تقویت می‌کند.

جدول ۴. مقایسه عملکرد مدل‌های زبانی در پاسخ‌گویی به سؤالات فقهی-سیاسی

سیاسی



تصویری به‌ویژه در تحلیل‌های چندمتغیره ضروری است، چراکه امکان ادراک توازن داخلی هر مدل را نیز تقویت می‌کند.

شکل ۲. مقایسه عملکرد مدل‌های زبانی در پنج شاخص تحلیلی

در ارزیابی کیفی انجام‌شده، مدل GPT-۴۰ توانست بیشترین انسجام

مفهومی و ساختار تحلیلی را در پاسخگویی به سؤالات اجتهادی از خود

نشان دهد. این مدل در استفاده از زنجیره تفکر، ارجاع دقیق به منابع

قرآنی و حدیثی و تولید استدلال‌های چندلایه، برتری محسوسی نسبت

به سایر مدل‌ها داشت. با این حال، در برخی پرسش‌های تضادمحور که

نیاز به انتخاب نهایی میان دیدگاه‌ها وجود داشت، مدل برخی مواقع از

داوری صریح خودداری می‌کرد. مدل Claude ۳ Opus نیز عملکردی

متعادل داشت. اگرچه از نظر زبانی و ساختار واژگانی بسیار حرفه‌ای عمل

می‌کرد، اما تحلیل‌های آن بیشتر به نمایش موازی دیدگاه‌ها محدود می‌شد

و انسجام تحلیلی لایه‌محور GPT-۴۰ را نداشت. مدل Gemini ۲.۵

Pro با رویکردی نسبتاً دقیق و تفسیرگرا به پرسش‌ها پاسخ داد و توانایی

نسبی در فهم مفاهیم فقهی داشت، اما نتوانست به عمق استدلالی و داوری

اجتهادی مورد انتظار برسد و پاسخ‌های آن اغلب در سطح میان‌تحلیل

باقی ماندند. در مورد Grok ۳، اگرچه از نظر سبک تحلیلی و زبان

استدلالی عملکرد خوبی داشت، ولی ضعف آن در ارجاع به منابع معتبر

دینی و درک ساختارهای خاص فقهی-سیاسی باعث شد که پاسخ‌ها

بیشتر جنبه فلسفی و تحلیل انتزاعی داشته باشند تا بازنمایی دقیق فقهی.

در نهایت، مدل GPT-۳.۵ به‌دلیل محدودیت‌های زیرساختی در پردازش

فقهی و ضعف در استناد منابع، پاسخ‌هایی کلی، ساده‌سازی‌شده و

غیرتحلیلی تولید کرد. این مدل فاقد ساختار لایه‌ای در تحلیل و ناتوان از

ارائه داوری نهایی قابل اتکاء در مسائل اجتهادی بود.

۳-۵. نحوه مدیریت تضاد مفهومی در پاسخ‌ها

۶۶

۷۳

۷۵

۷۸

۸۱

قدرت داوری

۷۲

۸۱

۸۳

۸۵

۸۷

قابلیت تعامل لایه‌ای

۶۸.۶

۷۵.۵

۷۸

۸۰.۶

۸۳.۲

میانگین کل از ۱۰۰

برای تسهیل درک مقایسه‌ای این تفاوت‌ها و تقویت خوانایی تحلیلی، داده‌های جدول به‌صورت یک نمودار میله‌ای نیز نمایش داده شده‌اند. این نمودار امکان مشاهده سریع و بصری اختلاف عملکرد مدل‌ها را در هر شاخص فراهم می‌سازد و به‌وضوح برتری نسبی GPT-۴۰ در تمامی شاخص‌ها را نسبت به سایر مدل‌ها نشان می‌دهد. بهره‌گیری از این نمایش



گستره رویکردهای فقهی قابل شناسایی توسط مدل و محور عمودی، توان دآوری و ترجیح منطقی میان آنها را مشخص می‌کند. در این ارزیابی، مدل GPT-۴۰ در هر دو شاخص عملکردی ممتاز داشت و در ناحیه بالا-راست نمودار قرار گرفت. در مقابل، مدل‌هایی مانند Grok و GPT-۳.۵ در نواحی پایین نمودار قرار گرفتند و توانایی محدودی در تحلیل و ترجیح دیدگاه‌های فقهی از خود نشان دادند. امتیازهای تخصیص یافته به مدل‌ها، حاصل میانگین‌گیری از ارزیابی پنج عامل ارزیاب متخصص بوده است که هر مدل را در دو شاخص «تنوع دیدگاه‌ها» و «توان ترجیح» بر پایه پاسخ‌های ارائه‌شده به مسائل فقهی-سیاسی نمره‌گذاری کرده‌اند. این داده‌ها در پیوست ۲ درج شده‌اند و مبنای ترسیم نمودار قرار گرفتند.

روش ارزیابی به صورت کیفی-کمی و با استفاده از فرم استاندارد مقایسه‌ای انجام شده است، به گونه‌ای که هم دقت عددی و هم تفسیر محتوایی در تحلیل مدل‌ها لحاظ شود. هدف از این سنجش، نه صرفاً رتبه‌بندی عملکرد مدل‌ها، بلکه ارزیابی قابلیت آنها در تعامل با پیچیدگی‌های درونی فقه و تضادهای معرفتی میان دیدگاه‌هاست. این رویکرد به شفاف‌سازی نقاط قوت و ضعف هر مدل در زمینه تعامل با مفاهیم اجتهادی کمک می‌کند. همچنین می‌توان از این چارچوب برای پایش بهبود عملکرد مدل‌ها در نسخه‌های بعدی استفاده کرد یا آن را به حوزه‌های دیگر مطالعات دینی بسط داد. در نهایت، نمودار ارائه‌شده می‌تواند به عنوان ابزاری کاربردی برای سنجش اخلاق‌محور و تحلیلی عملکرد مدل‌های زبانی در بستر مفاهیم اسلامی به کار گرفته شود.

شکل ۳. موقعیت مدل‌های زبانی در دو شاخص «تنوع دیدگاه‌های فقهی» و «توان دآوری و ترجیح استدلال»

یکی از مهم‌ترین ویژگی‌های مدل‌های هوش مصنوعی در مواجهه با سؤالات دینی-اجتهادی، توانایی آنها در تحلیل تضادها و انتخاب مسیر برتر است (۲۵). در این بخش، بررسی می‌شود که مدل‌ها چگونه با مسائلی روبه‌رو می‌شوند که در آنها دو دیدگاه متعارض فقهی-سیاسی وجود دارد؛ (مانند تقابل نصب الهی و رأی مردم). چارچوب تحلیلی و روش ارزیابی کیفی-کمی عملکرد مدل‌های هوش مصنوعی در مواجهه با تضادهای فقهی-سیاسی، به تفصیل در پیوست ۳ ارائه شده است. خلاصه‌ای از یافته‌ها در جدول زیر ارائه می‌شود:

جدول ۵. تحلیل رفتاری مدل‌ها در مواجهه با تضاد دیدگاه‌ها

نحوه مدیریت تضاد فقهی

مدل زبانی

تحلیل تطبیقی دیدگاه‌ها، دآوری نهایی روشن از طریق درخت تفکر

GPT-۴۰

نمایش موازی دیدگاه‌ها، بدون جمع‌بندی نهایی یا موضع‌گیری قاطع

Claude ۳ Opus

تحلیل ساختارمند اما فاقد انتخاب بین دو موضع متضاد

Gemini ۲.۵ Pro

گرایش به تفسیر فلسفی مفاهیم، بدون دآوری فقهی نهایی

Grok ۳

ساده‌سازی تضاد، پاسخ کلی و بدون تحلیل تضاد واقعی

GPT-۳.۵

برای سنجش توان مدل‌ها در مدیریت تضاد مفهومی، از نمودار دوبعدی استفاده شد. محور افقی، میزان تنوع دیدگاه‌ها را نشان می‌دهد (یعنی



۳-۶. پیچیدگی استدلال و توان ترکیب منابع تحلیلی

در پرسش‌های اجتهادی، قدرت مدل در استنتاج از منابع غیرمستقیم و تحلیل چندلایه، یکی از کلیدی‌ترین شاخص‌ها است (۲۶). در این بخش، میزان توان مدل‌ها در به‌کارگیری اصول کلامی، عقلانی و تحلیل چندمرحله‌ای بررسی شده است.

جدول ۶. ارزیابی میزان پیچیدگی استدلال تحلیلی

۷-۳. کیفیت زبانی و انسجام علمی پاسخ‌ها
زبان تحلیلی، واژگان تخصصی و انسجام نحوی از جمله معیارهایی هستند که کیفیت خروجی مدل را در سطح علمی مشخص می‌کنند (۲۷). در این بخش، کیفیت سبک زبانی پاسخ‌های هر مدل بررسی شده و درصد جملات علمی آن‌ها محاسبه شده است.

جدول ۷. ارزیابی کیفیت زبان علمی مدل‌ها

امتیاز (از ۱۰۰)

ویژگی استدلالی

مدل زبانی

۸۶

بهره‌گیری از اصول عقلانی، ترکیب منابع، استنتاج چندمرحله‌ای

GPT-۴o

۸۰

تحلیل استدلالی خوب، با پیچیدگی متوسط

Claude ۳ Opus

۷۵

ساختار منطقی دارد، اما فاقد تحلیل ضمنی و ترکیب لایه‌ای است

Gemini ۲.۵ Pro

۷۲

تحلیل مفهومی قوی ولی دور از مبانی فقهی مستقیم

Grok ۳

۶۴

استدلال‌های ساده‌سازی شده و غالباً صریح بدون عمق استنتاجی

GPT-۳.۵

درصد جملات علمی

ویژگی زبانی و سبک تحلیلی

مدل زبانی

٪۹۳

زبان دقیق، تخصصی، ارجاع‌محور و تحلیلی

GPT-۴o

٪۸۸

واژگان علمی مناسب، سبک آکادمیک خوب

Claude ۳ Opus

٪۸۲

ساختار فلسفی مناسب، در فقه کمی ناپایدار

Gemini ۲.۵ Pro

٪۷۹

واژگان غیرمذهبی، زبان تحلیلی بدون ساختار فقهی

Grok ۳

٪۶۴



$p < 0.001$

۱۹۷.۷۲

دقت منابع

معنادار

$p < 0.001$

۱۰۷.۷۵

عمق استدلال

معنادار

$p < 0.001$

۱۸۱.۴۳

تنوع دیدگاه‌ها

معنادار

$p < 0.001$

۲۲۲.۷۹

قدرت داوری

معنادار

$p < 0.001$

۱۳۵.۷۶

قابلیت تعامل لایه‌ای

زبان ساده و عمومی، فاقد انسجام تحلیل آکادمیک

GPT-۳.۵

۳-۸. تحلیل آماری تفاوت عملکرد مدل‌ها با استفاده از آزمون ANOVA

به‌منظور سنجش معناداری آماری تفاوت عملکرد میان مدل‌های زبانی

مختلف، از آزمون تحلیل واریانس یک‌طرفه (One-Way ANOVA)

استفاده شد (۲۸). این آزمون برای بررسی آن طراحی شد که آیا اختلاف

میانگین امتیازها در شاخص‌های پنج‌گانه (دقت منابع، عمق استدلال،

تنوع دیدگاه، قدرت داوری و قابلیت تعامل) میان مدل‌ها معنادار است یا

خیر.

۷ مراحل تحلیل:

• فرضیه صفر (H_0): بین میانگین عملکرد مدل‌ها تفاوت معناداری

وجود ندارد.

• فرضیه مقابل (H_1): دست‌کم میان دو مدل تفاوت معناداری وجود

دارد.

داده‌های آزمون از امتیازات اختصاص داده‌شده توسط پنج ارزیاب برای هر

مدل، در هر شاخص به‌صورت مستقل استخراج شده و با استفاده از کد

پایتون مورد تحلیل قرار گرفت.

جدول ۸. نتایج آزمون ANOVA برای شاخص‌های ارزیابی

نتیجه

مقدار p (sig)

مقدار F

شاخص ارزیابی

معنادار

با توجه به اینکه مقدار p در تمامی شاخص‌ها کمتر از ۰.۰۵ است، می‌توان

با اطمینان ۹۵٪ نتیجه گرفت که تفاوت میانگین عملکرد مدل‌ها در

تمامی شاخص‌ها از نظر آماری معنادار است.

۷ آزمون تعقیبی (Post-hoc)

پس از مشخص شدن تفاوت معنادار آماری عملکرد مدل‌های زبانی در



معنادار

$p < 0.001$

۹.۷۵+

GPT-۳.۵

Gemini ۲.۵ Pro

معنادار

$p < 0.001$

۱۲.۴۵-

GPT-۴o

Grok ۳

معنادار

$p < 0.001$

۱۱.۹۳-

Claude ۳ Opus

GPT-۳.۵

همان‌طور که جدول فوق نشان می‌دهد، مدل GPT-۴o به‌طور معناداری عملکرد بهتری نسبت به GPT-۳.۵ دارد. سایر تفاوت‌های معنادار نیز عمدتاً در مقایسه مدل‌های قوی (Claude، Gemini) با مدل‌های ضعیف‌تر (GPT-۳.۵، Grok۳) مشاهده می‌شود. این یافته‌ها مؤید آن است که چارچوب ارزیابی طراحی شده در این پژوهش، توانایی تمایز دقیق و معنادار بین عملکرد مدل‌ها را داراست.

۷ جمع‌بندی یافته‌های آماری

یافته‌های آماری پژوهش به‌روشنی نشان می‌دهد که چارچوب ارزیابی چندلایه پیشنهادی، قادر است عملکرد مدل‌ها را به‌صورت علمی و معتبر

شاخص‌های مورد ارزیابی با استفاده از آزمون ANOVA، برای شناسایی دقیق مدل‌هایی که این تفاوت را ایجاد کرده‌اند، از آزمون تعقیبی Tukey HSD بهره گرفته شد (۲۹). این آزمون امکان مقایسه جفتی مدل‌ها را فراهم می‌آورد و مشخص می‌کند که آیا اختلاف میانگین عملکرد دو مدل خاص از نظر آماری معنادار است یا خیر.

در اینجا برای هر مدل زبانی، مهم‌ترین مقایسه آماری معنادار آن با یک مدل دیگر انتخاب و به‌صورت خلاصه و علمی در جدول زیر تنظیم شده است.

جدول ۹. مقایسه آماری کلیدی بین مدل‌های زبانی (Tukey HSD)

نتیجه آزمون

مقدار p

اختلاف میانگین

مدل دوم

مدل اول

معنادار

$p < 0.001$

۱۴.۵۶+

GPT-۳.۵

GPT-۴o

معنادار

۰.۰۰۲

۴.۷۹-

Grok ۳

Claude ۳ Opus



۱.۳

تفکیک نماید. به‌ویژه، عملکرد مدل GPT-۴۰ در شاخص‌های کلیدی پاسخ‌گویی فقهی نسبت به مدل‌های پیشین به‌طور معناداری ارتقاء یافته است. این موضوع بیانگر آن است که تمایزهای مشاهده‌شده نه تنها کیفی یا ذهنی، بلکه دارای پشتوانه آماری مستند هستند.

۴. بحث و نتیجه‌گیری

۴-۱. جمع‌بندی نظری

نتایج این پژوهش نشان می‌دهد که مدل‌های زبانی بزرگ مانند GPT-۴۰، در صورت بهره‌گیری از چارچوب مفهومی دقیق و اعلان‌های لایه‌ای کنترل‌شده، قادر به بازنمایی نظام‌مند مفاهیم فقهی و سیاسی پیچیده، مانند مشروعیت دینی، جایگاه فقیه در حکمرانی اسلامی و نظریه ولایت مطلقه هستند. تحلیل تجربی تأیید کرد که مدل‌های زبانی بزرگ می‌توانند استدلال‌های قرآنی، روایی، فلسفی و اجتماعی را تفکیک کرده و پاسخ‌هایی چندلایه و مبتنی بر منطق اجتهادی ارائه دهند.

با این حال، محدودیت‌هایی نیز مشاهده شد:

- عدم دقت کافی در ارجاع به منابع حدیثی یا متون اصلی.

- تمایل مدل به اجتناب از انتخاب قطعی در پرسش‌های تضادمحور.

- نیاز به اعلان‌های کنترلی قوی برای جلوگیری از تولید اطلاعات نادرست.

چارچوب سه‌لایه‌ای پیشنهادی (زنجیره تفکر، ساختار عامل‌محور، و درخت

پاسخ) ظرفیت مدل‌های زبانی را برای بازنمایی فقه سیاسی به‌طور

قابل توجهی ارتقاء داد و پتانسیل آن را برای مدل‌سازی مفاهیم دینی نشان داد.

۱.

۱.۱

۱.۲

این مطالعه نشان داد که تلفیق نظریه‌های دینی مشروعیت، به‌ویژه در سنت اجتهادی شیعه، با ساختارهای پردازش‌محور هوش مصنوعی امکان‌پذیر است و افق جدیدی برای بازتعریف رابطه دین و فناوری می‌گشاید. برخلاف گفتمان رایج که دین را صرفاً منبعی برای ارزش‌گذاری اخلاقی می‌بیند، این پژوهش نشان داد که نظام‌های فقهی، مانند ولایت فقیه، دارای منطق تصمیم‌سازی و معماری حکمرانی هستند که می‌توان آن‌ها را الگوریتمی مدل‌سازی کرد.

نظریه ولایت فقیه، با تأکید بر نصب الهی، عدالت فقیه و استمرار ولایت در عصر غیبت، ساختاری متمایز از مدل‌های سکولار مبتنی بر رأی یا کارآمدی ارائه می‌دهد. این ساختار، با مؤلفه‌های هوش مصنوعی تصمیم‌یار قابل تطبیق است و مسیرهایی نو برای طراحی سامانه‌های هوش‌محور الهیاتی در بسترهای اسلامی و جهانی فراهم می‌کند.

۲.

۲.۱

۲.۲

۲.۳

۲.۴

۳-۴. نتیجه نهایی

این پژوهش به‌صورت نظام‌مند نشان داد که بازنمایی نظریه‌های فقهی پیچیده، مانند نظریه ولایت مطلقه فقیه، در قالب مدل‌های زبان بزرگ امکان‌پذیر است، مشروط بر آنکه چارچوب مفهومی دقیقی طراحی شده و از مهندسی اعلان پیشرفته، تحلیل چندلایه و ارزیابی عامل‌محور استفاده



متنوع تر یا سنت‌های الهیاتی گوناگون می‌تواند تصویر جامع‌تری از کارایی آن ارائه دهد.

برای توسعه این مسیر پژوهشی، پیشنهاد‌های زیر ارائه می‌شود:

- توسعه پایگاه داده‌های دینی و مدل‌های بومی: طراحی و آموزش مدل‌های زبانی بر پایه منابع ساخت‌یافته قرآنی، روایی و فقهی در سنت‌های مختلف اسلامی و سایر نظام‌های الهیاتی، با هدف ارتقاء دقت، تفسیرپذیری و حساسیت فرهنگی مدل‌ها.

- ایجاد ابزارهای اجتهاد دیجیتال: توسعه سامانه‌های پشتیبان تصمیم برای کمک به فقها و پژوهشگران علوم دینی در تحلیل ساختاریافته متون و استنباط فقهی.

- آزمون تطبیقی در سنت‌های متنوع: پیاده‌سازی چارچوب پیشنهادی در بسترهایی مانند خلافت اهل سنت، الهیات مسیحی، اخلاق یهودی و نظام‌های فلسفی شرقی مانند بودیسم و هندوئیسم، برای سنجش میزان تعمیم‌پذیری مفهومی و الگوریتمی.

- تقویت همکاری‌های میان‌رشته‌ای: گسترش ارتباط میان مراکز علمی علوم دینی، دانشگاه‌های فناوری و پژوهشگاه‌های هوش مصنوعی با هدف توسعه سامانه‌های دین‌یار و اخلاق‌محور با رویکرد چندفرهنگی.

اجرای این پیشنهادها می‌تواند به توسعه مدل‌های زبانی هم‌ساز با زمینه‌های دینی و فرهنگی، کاهش سوء‌گیری‌های سکولار در سامانه‌های هوشمند، و ارتقاء گفتمان‌های مشارکتی در حوزه دین و فناوری منجر شود.

شود. چارچوب سه‌لایه‌ای پیشنهادی توانست فرآیند استدلال اجتهادی را به‌گونه‌ای ساخت‌یافته، مرحله‌محور و قابل ارزیابی در مدل‌های زبانی پیاده‌سازی کند. یافته‌های مقاله نشان داد که این ساختار نه تنها به افزایش دقت منابع و عمق تحلیل منجر می‌شود، بلکه توانایی مدل را در تمایزگذاری میان دیدگاه‌های فقهی متعارض و ارائه پاسخ‌های هماهنگ با منطق درون‌دینی به‌طور محسوس ارتقاء می‌دهد.

از منظر نظری، این چارچوب گامی عملی به‌سوی تلفیق هوش مصنوعی و فقه اسلامی در سطحی بالاتر از صرفاً پاسخ‌گویی عمومی محسوب می‌شود و نشان می‌دهد که نظام‌های دینی نیز می‌توانند با بهره‌گیری از فناوری‌های نوین، به شیوه‌ای الگوریتمی مدل‌سازی شوند. این دستاورد به‌ویژه در بستر طراحی سامانه‌های هوش‌محور بومی، ابزارهای اجتهاد دیجیتال و حتی تحلیل جنگ شناختی در حوزه پدافند نرم، قابلیت توسعه و به‌کارگیری دارد.

در نهایت، این پژوهش تلاش کرده است تا بدون موضع‌گیری ایدئولوژیک، ظرفیت تحلیل مفاهیم دینی توسط مدل‌های زبانی را در چارچوب اجتهادی رایج در فقه شیعی مورد آزمون قرار دهد و راه را برای پژوهش‌های میان‌رشته‌ای بیشتر در حوزه فقه و فناوری بگشاید.

۵. محدودیت‌ها و پیشنهادها برای پژوهش‌های آینده

با وجود دستاوردهای این پژوهش، برخی محدودیت‌ها باید مورد توجه قرار گیرد. نخست، تحلیل‌ها در بستر مفاهیم فقه حکومتی و چارچوب‌های رایج در اندیشه سیاسی شیعه صورت گرفته‌اند و امکان تعمیم آن‌ها به سایر رویکردهای دینی یا فلسفی نیازمند بررسی‌های بیشتر است. دوم، دامنه تجربی پژوهش به مجموعه‌ای محدود از پرسش‌های اجتهادی و مدل‌های منتخب محدود بوده است. هرچند مدل Grok ۳ به‌دلیل توان تحلیل مفهومی انتخاب شد، اما ارزیابی این چارچوب در مواجهه با پرسش‌های



فصلنامه علمی فقه و حقوق نوین

Journal of Jurisprudence and Modern Law

Home Page: www.jaml.ir



مرکز اطلاعات علمی جهان و دانشگاه
Scientific Information Database

Sponsored and Indexed by
CIVILICA
We Respect the Science



ESJI
Eurasian Scientific Journal Index
www.ESJIndex.org



Academic Resource Index
ResearchBib



Academia.edu
share research



Agent-Based Reasoning [۲]

Tree of Thought [۳]



[۸] Crawford, K., & Calo, R. (۲۰۲۱). The politics of AI & Power and governance in automated systems. AI Society, ۳۶(۴), ۱۱۲۳-۱۱۳۵. <https://doi.org/10.1007/s00146-021-01234-0>

[۹] Rahim, M., & Hosseini, S. (۲۰۲۳). AI and Islamic political theory: Compatibility and applications in governance. Journal of Islamic Studies, ۳۴(۳), ۴۱۲-۴۳۰. <https://doi.org/10.1093/jis/etad012>

[۱۰] Binns, R., & Veale, M. (۲۰۲۳). Power dynamics in AI Ethical and societal implications. ACM Transactions on Computer-Human Interaction, ۳۰(۴), ۱-۲۷. <https://doi.org/10.1145/3604723>

[۱۱] Yasini, M. (۲۰۲۳). AI and Religious Pluralism: The & Missed Conversations in Machine Learning Ethics. AI Society. <https://doi.org/10.1007/s00146-023-01643-x>

[۱۲] عباسی، ر.، و جوادزاده، م.ع. (۱۴۰۱). پیش‌بینی ناآرامی‌های مردمی با استفاده از شبکه‌های اجتماعی، مبتنی بر یادگیری ماشین در پردازش زبان طبیعی. پدافند الکترونیکی و سایبری، سال ۱۳، شماره ۳، صص ۴۵-۵۶.

[۱۳] داداش‌تبار احمدی، ک.، خیرخواه، م.، و جبار رشیدی، ع. (۱۳۹۷). تشخیص حملات سایبری پیشرفته با استفاده از مدل‌سازی رفتاری مبتنی بر پردازش زبان طبیعی. پدافند الکترونیکی و سایبری، سال ۶، شماره ۳، صص ۱۴۱-۱۵۱.

[۱۴] شیخی، م.، عباسی، س.، و شاکری، م. (۱۴۰۰). تحلیل احساسات کاربران شبکه‌های اجتماعی فارسی زبان با استفاده از روش‌های پردازش زبان طبیعی و یادگیری ماشین. پدافند الکترونیکی و سایبری، سال ۱۲، شماره ۲، صص ۶۳-۷۷.

[۱۵] قاسمی، ح.، و دهقانی، ب. (۱۳۹۹). استخراج خودکار مفاهیم کلیدی از اسناد فارسی با استفاده از الگوریتم‌های یادگیری ماشین و پردازش زبان

منابع

- [۱] Vaswani, A., Shazeer, N., Parmar, N., & et al. (۲۰۲۳). Attention is all you need: Revisited. Journal of Machine Learning Research, ۲۴(۲), ۱-۳۲. <https://doi.org/10.5555/3666122.3666145>
- [۲] Turan, C. et al. (۲۰۲۳). Beyond the Western Canon: A Global Perspective on AI Ethics. AI and Society <https://doi.org/10.1007/s00146-023-01596-1>
- [۳] Alshamsi, A. et al. (۲۰۲۲). Religious AI: A Review on Integrating Islamic Ethics in Artificial Intelligence. IEEE Access, ۱۰, ۶۸۲۷۶-۶۸۲۹۲. <https://doi.org/10.1109/ACCESS.2022.3186994>
- [۴] Hagendorff, T. (۲۰۲۳). Machine ethics: Principles and challenges for responsible AI development. Ethics and Information Technology, ۲۵(۳), ۱-۱۸. <https://doi.org/10.1007/s10676-023-00912-3>
- [۵] Khomeini, R. (۲۰۰۲). Islamic Government: Governance of the Jurist. Institute for Compilation and Publication of Imam Khomeini's Works
- [۶] Chen, Z., Deng, Y., Wu, H., & et al. (۲۰۲۳). Faithful reasoning in large language models: Techniques and evaluation. ACM Transactions on Intelligent Systems and Technology, ۱۴(۵), ۱-۲۵. <https://doi.org/10.1145/3595923>
- [۷] Morley, J., Floridi, L., & et al. (۲۰۲۴). Global AI ethics: A review of guidelines and frameworks. AI & Society, ۳۹(۲). <https://doi.org/10.1007/s00146-024-01845-5>



<https://doi.org/10.1145/3641289.38-1>, ۵۶ (۸), ۳۸-۱

[۲۳] Wei, J., Wang, X., Schuurmans, D., & et al. (۲۰۲۳). Chain-of-thought prompting elicits reasoning in large language models. *Journal of Machine Learning Research* <https://doi.org/10.5555/3648699.3649046>, ۴۵-۱ (۱), ۲۴

[۲۴] Khan, M., & Siddiqui, A. (۲۰۲۴). A multidimensional framework for evaluating AI in religious contexts: Accuracy and ethical considerations. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, ۵۴(۵), ۲۹۸۷-۳۰۰۱ <https://doi.org/10.1109/TSMC.2023.3324567>

[۲۵] Wang, X., & Chen, Y. (۲۰۲۳). Interpretable reasoning in large language models: Methods and applications. *Neural Networks*, ۱۶۲, ۴۵-۶۰ <https://doi.org/10.1016/j.neunet.2023.02.012>

[۲۶] Zhang, H., & Liu, Q. (۲۰۲۴). Trustworthy multi-agent systems in AI: Ethical and practical challenges. *IEEE Transactions on Dependable and Secure Computing*, ۲۱(۳) <https://doi.org/10.1109/TDSC.2023.3301234>, ۱۳۴۵-۱۳۴۵

[۲۷] Floridi, L., & Cows, J. (۲۰۲۳). A unified framework for AI ethics and fairness in multi-agent systems. *Ethics and Information Technology*, ۲۵(۲), ۱-۱۵ <https://doi.org/10.1007/s10676-023-09678-2>

[۲۸] Anderson, T. W., & Taylor, J. B. (۲۰۲۴). Advanced statistical methods for AI-driven data analysis Applications in complex systems. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, ۵۴(۷), ۴۱۲۳-۴۱۳۸ <https://doi.org/10.1109/TSMC.2024.3356789>

[۲۹] Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K

طبیعی. پدافند الکترونیکی و سایبری، سال ۱۱، شماره ۴، صص ۱۰۱-۱۱۵.

[۱۶] Brown, T., Mann, B., Ryder, N., & et al. (۲۰۲۳). Language models are few-shot learners: Advances in reasoning and contextual understanding. *IEEE Transactions on Artificial Intelligence*, ۴(۳), ۵۱۲-۵۲۸ <https://doi.org/10.1109/TAI.2023.3256789>

[۱۷] Al-Masri, A., & Khan, S. (۲۰۲۴). Enhancing hadith authentication with AI embeddings. *IEEE Transactions on Knowledge and Data Engineering*, ۳۶(۴), ۱۲۳۴-۱۲۴۸ <https://doi.org/10.1109/TKDE.2023.3312456>

[۱۸] Wang, X., Wei, J., & Liu, Y. (۲۰۲۴). Advanced chain-of-thought prompting for complex reasoning tasks. *IEEE Transactions on Neural Networks and Learning Systems*, ۷۹۰۵-۷۸۹۰ (۶), ۳۵

<https://doi.org/10.1109/TNNLS.2023.3278901>

[۱۹] Mavani, H. (۲۰۲۲). Shi'i Islam and the Question of Authority: Reassessing Wilayat al-Faqih. Routledge

[۲۰] Qiao, S., Ou, Y., Zhang, N., & et al. (۲۰۲۳). Reasoning with language model prompting: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. Advance online publication <https://doi.org/10.1109/TPAMI.2023.3293567>

[۲۱] Li, Y., Du, H., & Zhou, J. (۲۰۲۴). Multi-agent systems for evaluating large language models: A survey. *ACM Computing Surveys*, ۵۶(۱۰), ۱-۳۵ <https://doi.org/10.1145/3664602>

[۲۲] Zhao, C., Zhang, N., & et al. (۲۰۲۴). Logical reasoning over natural language: A survey. *ACM Computing Surveys*



فصلنامه علمی فقه و حقوق نوین

Journal of Jurisprudence and Modern Law

Home Page: www.jaml.ir



مرکز اطلاعات علمی
Scientific Information Database

Sponsored and Indexed by
CIVILICA
We Respect the Science



ESJI
www.ESJIndex.org

Eurasian
Scientific
Journal
Index



Academic
Resource
Index
ResearchBib



Academia.edu
share research



BERT revisited: Advances in pre-trained language (۲۰۲۳).

,models. ACM Transactions on Information Systems, ۴۱(۴).

<https://doi.org/10.1145/3579856.28-1>