



وزارت علوم، تحقیقات و فناوری

دوره‌های آزاد مؤسسه پژوهشی حکمت و فلسفه ایران

نیم‌سال اول ۱۴۰۵-۱۴۰۶

شعبه قم

فلسفه هوش مصنوعی و معرفت‌شناسی مدل‌های زبانی بزرگ

حجت‌الاسلام دکتر حمید محسنی

استادیار پژوهشگاه فرهنگ و اندیشه اسلامی

رئیس مرکز مطالعات فضای مجازی و هوش مصنوعی

سه‌شنبه‌ها: ساعت ۱۶ الی ۱۸

شروع دوره: ۲۱ مهرماه ۱۴۰۵

مؤسسه پژوهشی حکمت و فلسفه ایران (شعبه قم)

طرح‌نامه تدریس: «فلسفه هوش مصنوعی و معرفت‌شناسی مدل‌های زبانی بزرگ»

تعداد جلسات: ۱۶ جلسه دوساعته

مخاطبان اصلی:

✓ طلاب حوزه علمیه سطح دو به بالا:

یعنی طلابی که مقدمات را گذرانده و با منطق، فلسفه اسلامی، کلام، فقه و اصول در حد متوسط آشنا هستند.

✓ دانشجویان رشته‌های علوم انسانی (فلسفه، کلام و الهیات، فلسفه ذهن،

علوم شناختی، فلسفه علم، فلسفه فناوری):

یعنی دانشجویانی که با مباحث فلسفی، معرفت‌شناختی، انسان‌شناختی و علوم‌شناختی و فلسفه ذهن آشنایی دارند، هرچند ممکن است با فلسفه اسلامی یا مباحث فنی و فلسفی هوش مصنوعی کمتر آشنا باشند.

اهداف دوره:

چهار هدف اصلی دوره را می‌توان چنین خلاصه کرد:

- ✓ فهم نظاممند هوش مصنوعی و مبانی فلسفی آن
- ✓ تحلیل معرفت‌شناختی و معناشناختی مدل‌های زبانی بزرگ
- ✓ مقایسه تطبیقی و انتقادی فلسفه اسلامی و فلسفه غرب در مسائل هوش مصنوعی
- ✓ توانمندسازی تحلیلی و پژوهشی، برای مواجهه مسئولانه با هوش مصنوعی

این چهار هدف، هم با محورهای تدریس هم‌راستا هستند و هم نیازهای دو گروه اصلی مخاطبان، یعنی طلاب حوزه و دانشجویان علوم انسانی و فلسفه، را پوشش می‌دهد.

مقدمه:

هوش مصنوعی امروز دیگر یک فناوری صرف نیست؛ بلکه پدیده‌ای است که بنیادهای معرفتی، انسان‌شناختی، ارزشی و دینی زیست‌جهان انسانی را به چالش می‌کشد. مدل‌های زبانی بزرگ به عنوان پیشرفته‌ترین جلوه این پدیده، پرسش‌هایی بنیادین درباره ماهیت علم، فهم، معنا، حیث التفاتی، آگاهی و مسئولیت اخلاقی پدید آورده‌اند. این پرسش‌ها نه صرفاً با دانش فنی پاسخ می‌گیرند و نه با فلسفه ورزی انتزاعی؛ بلکه نیازمند رویکردی میان‌رشته‌ای و تطبیقی‌اند که میان فلسفه هوش مصنوعی، فلسفه ذهن، معرفت‌شناسی، علوم‌شناختی و سنت فلسفه اسلامی گفت‌وگو برقرار کند. دوره حاضر با همین رویکرد طراحی شده و مخاطبان آن طلاب حوزه علمیه سطح دو به بالا و دانشجویان رشته‌های علوم انسانی، فلسفه، کلام و الهیات، فلسفه ذهن، علوم‌شناختی، فلسفه علم و فلسفه فناوری هستند. انتظار می‌رود شرکت‌کنندگان با مبانی منطقی، فلسفه اسلامی، معرفت‌شناسی

و فلسفه ذهن در حد مقدماتی آشنا باشند. این دوره چهار هدف اصلی را دنبال می‌کند: فهم نظام‌مند چیستی هوش مصنوعی و مبانی فلسفی آن؛ تحلیل معرفت‌شناختی و معناشناختی مدل‌های زبانی بزرگ؛ مقایسه تطبیقی و انتقادی فلسفه اسلامی و فلسفه غرب در مسائل هوش مصنوعی؛ و توانمندسازی پژوهشی، اخلاقی-دینی و کاربردی برای مواجهه مسئولانه با هوش مصنوعی. محورهای تدریس از چیستی هوش مصنوعی و فلسفه آن آغاز می‌شود و با مباحثی چون انواع و نسل‌های هوش مصنوعی، مسائل هستی‌شناختی، ارزش‌شناختی، معرفت‌شناختی، انسان‌شناختی و دین‌شناختی، آشنایی با مدل‌های زبانی بزرگ، چالش‌های آن‌ها، معرفت‌شناسی و معناشناسی مدل‌های زبانی، بررسی امکان فهم معنا، و معرفی منابع و مسائل پژوهشی ادامه می‌یابد.

محورهای تدریس:

(۱) چیستی هوش مصنوعی

- ✓ آشنایی با انواع تعاریف هوش مصنوعی از منظر اندیشمندان مختلف
- ✓ تعاریف علوم رایانه‌ای
- ✓ تعاریف فلسفی
- ✓ تعریفی جامع

(۲) چیستی فلسفه هوش مصنوعی

- ✓ انواع تعاریف فلسفه هوش مصنوعی
- ✓ آشنایی با کتب فلسفه هوش مصنوعی
- ✓ آشنایی با مسائل فلسفی هوش مصنوعی

(۳) آشنایی با اندیشمندان و فیلسوفان مطرح در حوزه هوش مصنوعی (۲ جلسه)

- ✓ آلن تورینگ و آزمون تورینگ
- ✓ جان مک‌کارتی و کارگاه دارتموث
- ✓ استوارت راسل و آثار معروفش
- ✓ هوبرت دریفوس و شهود و خبرگی

- ✓ جان سرل و اتاق چینی
- ✓ چالمرز و مسئله آگاهی

(۴) انواع هوش مصنوعی

- ✓ هوش مصنوعی ضعیف
- ✓ هوش مصنوعی قوی
- ✓ هوش مصنوعی عمومی
- ✓ ابر هوش
- ✓ هوش مصنوعی بدنمند

(۵) نسل‌ها و رویکردهای هوش مصنوعی

- ✓ نسل اول: رویکرد محاسباتی: نمادگرا و پیوندگرا
- ✓ نسل دوم: رویکرد بدنمند
- ✓ نسل سوم: رویکرد نظریات ساخت 4E

(۶) مسائل نظری و فلسفی هوش مصنوعی

- ✓ مسائل هستی‌شناختی
- ✓ مسائل ارزش‌شناختی
- ✓ مسائل معرفت‌شناختی
- ✓ مسائل انسان‌شناختی
- ✓ مسائل دین‌شناختی

(۷) آشنایی با مدل‌های زبانی بزرگ

- ✓ آشنایی با انواع مدل‌های نظیر ChatGPT
- ✓ آموزش مدل‌ها
- ✓ انواع داده‌های آموزشی
- ✓ بررسی عملکرد مدل‌ها

(۸) مسائل و چالش‌های مدل‌های زبانی بزرگ

- ✓ اصالت و اعتبار محتوا

- ✓ توهم
- ✓ سوگیری
- ✓ ...

(۹) معرفت‌شناسی مدل‌های زبانی بزرگ

- ✓ «باور صادق موجه» در مدل‌ها
- ✓ مطابقت داده با خارج
- ✓ انسجام‌گرایی درونی مدل‌ها

(۱۰) معنا، دانش و درک در مدل‌های زبانی بزرگ

- ✓ انتساب حالات شناختی به مدل‌ها
- ✓ دیدگاه مهندسان
- ✓ دیدگاه فلاسفه
- ✓ هویت علم در انسان از منظر فلاسفه اسلامی

(۱۱) بررسی امکان درک معنا توسط مدل‌های زبانی بزرگ (۲ جلسه)

- ✓ بررسی امکان درک معنا توسط مدل‌ها از منظر «استدلال اتاق چینی سرل»
- ✓ بررسی فاعل شناسا در مدل‌ها از منظر «برهان فضای آزاد ابن سینا»

(۱۲) معرفی آثار، منابع و مسائل پژوهی در حوزه هوش مصنوعی

- ✓ مقالات و کتب مطرح
- ✓ نشریات مطرح در دنیا
- ✓ مراکز تحقیقاتی و دانشگاهی مطرح در دنیا
- ✓ پیشنهاد مسائل فلسفی پژوهشی