

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ



دانشگاه شهید بهشتی
پژوهشکده مطالعات بنیادین علم و فناوری

پایان نامه کارشناسی ارشد

عنوان تکنیکی هوش مصنوعی و امکان پذیری آن

پژوهشگر
باقر رضانی بیانلوجه

استادان راهنما
دکتر محمود مختاری، دکتر زهرا زرگر

شهریور ۱۴۰۴

کلیه حقوق مادی و معنوی مترتب بر دستاوردهای مطالعات، ابتکارات و نوآوری‌های منتج از این اثر متعلق به دانشگاه شهید بهشتی است و دانشجو موظف به رعایت آیین‌نامه و منشور اخلاق در پژوهش برای ارائه یا چاپ مطالب مستخرج از این اثر است. نقل مطالب با ذکر مأخذ آزاد است.

صور تجلسه دفاع از پایان نامه دانشجویان دوره کارشناسی ارشد

بازگشت به مجوز شماره مورخ جلسه هیئت داوران، ارزیابی پایان نامه
خانم/ آقای به شماره شناسنامه صادره از
دانشجوی دوره کارشناسی ارشد رشته

با عنوان :

--

به راهنمایی :

طبق دعوت قبلی در تاریخ
و درجه تصویب شد.
غیرقابل قبول (0-13,99) 2. قابل قبول (14-15,99) 3. خوب (16-17,99) 4. بسیار خوب (18-18,99) 5. عالی (19-20)

سمت در پایان نامه	نام و نام خانوادگی	مرتبه علمی	نام دانشگاه	امضاء
۱. استاد راهنما:				
۲. استاد مشاور:				
۳. داور ۱:				
۴. داور ۲:				
۵. نماینده تحصیلات تکمیلی:				
۶. مدیر گروه:				

تذکر: این صورتجلسه صرفاً مبین نتیجه جلسه دفاع بوده و فاقد ارزش گواهی فراغت از تحصیل است.
بنابراین ارائه و یا پذیرش آن به عنوان گواهی فراغت از تحصیل هیچگونه وجاهت قانونی ندارد.

تقدیم به:

خانواده عزیزم، به‌ویژه مادر مهربانم، همسر و سلوا.

سیاسگزاری

با تشکر از دکتر زهرا زرگر، دکتر محمود مختاری و دکتر روزبه زارع و سایر اساتید و نویسندگان ،
برای راهنمایی‌ها و حمایت‌هایشان.

تعهد قانون حفظ طبع

نام و نام خانوادگی:

عنوان پایان نامه/رساله:

استاد/استادان راهنما:

این جانب..... تهیه کننده پایان نامه کارشناسی ارشد/رساله دکتری حاضر، خود را ملزم به حفظ امانتداری و قدردانی از زحمات سایر محققان و نویسندگان بنا بر قانون حفظ طبع می دانم. بدین وسیله اعلام می نمایم که مسئولیت کلیه مطالب درج شده با این جانب است. در صورت استفاده از شکل ها، جدول ها و مطالب سایر منابع، بلافاصله مرجع آن ذکر شده و سایر مطالب از کار تحقیقاتی این جانب استخراج شده است و امانت داری را به صورت کامل رعایت نموده ام. در صورتی که خلاف این مطلب ثابت شود، مسئولیت کلیه عواقب قانونی با شخص این جانب است.

نام و نام خانوادگی دانشجو:

امضاء و تاریخ:

چکیده

این پژوهش به ارزیابی فلسفی امکان‌پذیری هم‌ارزی هوش مصنوعی با هوش انسانی، به عنوان پیش‌فرض بنیادین نظریه تکینگی فناورانه، می‌پردازد. چشم‌انداز تکینگی، که به نقطه‌ای فرضی در آینده با ظهور هوش فرابشری اشاره دارد، جامعه فکری را به دو قطب موافقان و مخالفان تقسیم کرده است. پذیرش یا رد امکان هم‌ارزی هوش ماشینی با هوش انسانی از جمله محورهای اصلی در مباحث ارائه شده درباره تکینگی است. مسئله اصلی این پژوهش، کاوش در بنیان‌های فلسفی همین مسئله و پاسخ به این پرسش محوری است که آیا «هم‌ارزی» میان هوش انسانی (با تمام ابعاد بدنمند، موقعیت‌مند و شهودی آن) و هوش مصنوعی (به عنوان پدیده‌ای اساساً محاسباتی) از منظر فلسفی (پدیدارشناسی و حکمت متعالیه) امکان‌پذیر است؟

روش تحقیق در این پایان‌نامه، توصیفی-تحلیلی با رویکردی تطبیقی و میان‌رشته‌ای است. در این راستا، ابتدا دیدگاه‌ها و استدلال‌های کلیدی حامیان تکینگی (از جمله کورزویل، چالمرز و تیپلر) بر پایه مفاهیمی چون «قانون بازده شتابان» و «انفجار هوش» تبیین و صورت‌بندی می‌شود. سپس، این دیدگاه‌ها از منظر دو سنت فلسفی مستقل مورد نقد و ارزیابی قرار می‌گیرند: نخست، از دیدگاه پدیدارشناسی و نظریه «شناخت بدنمند» با تمرکز بر آراء هیوبرت دریفوس؛ و دوم، از منظر «حکمت متعالیه» با تکیه بر مبانی هستی‌شناسی نفس در فلسفه ملاصدرا. این پژوهش دارای محدودیت‌هایی از جمله تمرکز بر مباحث مفهومی و فلسفی به جای ابعاد فنی و الگوریتمی، و چالش‌های تفسیری در استنتاج آراء فلاسفه کلاسیک در باب موضوعی معاصر است.

یافته‌های پژوهش نشان می‌دهد که هر دو سنت فلسفی، علی‌رغم تفاوت‌های بنیادین در مبانی و اصطلاحات، در نقد پیش‌فرض هم‌ارزی به نتایجی همگرا می‌رسند. از منظر پدیدارشناسی، هوش انسانی پدیده‌ای جدایی‌ناپذیر از «بدن‌مندی»، «موقعیت‌مندی» و «مهارت» عملی است که به سادگی قابل فروکاستن به پردازش نمادهای انتزاعی نیست. از منظر حکمت متعالیه نیز، قیود و اصولی چون ضرورت وجود «جسم طبیعی» و «مزاج» به عنوان بستر تکامل، «غایت‌مندی ذاتی» نفس، و ماهیت منحصر به فرد «علم حضوری» و خودآگاهی، مانعی ماهوی بر سر راه هم‌ارزی کامل یک موجود مصنوعی با انسان قرار می‌دهد. این همگرایی انتقادی نشان می‌دهد که چالش‌های پیش روی تحقق تکینگی صرفاً فنی نبوده، بلکه ریشه در مسائل عمیق هستی‌شناختی دارد.

نتیجه‌گیری نهایی این پژوهش آن است که هم‌ارزی کامل میان هوش مصنوعی و هوش انسانی از منظر سنت‌های فلسفی بررسی شده، ممکن به نظر نمی‌رسد و در نتیجه، تکینگی فناورانه، به معنای رایج آن که بر این هم‌ارزی استوار است، با موانع بنیادین فلسفی روبروست. بر این اساس، پیشنهاد می‌شود تمرکز پژوهش‌ها از خلق یک «هوش رقیب» به سمت توسعه «هوش افزوده» که مکمل توانایی‌های انسانی باشد، معطوف گردد و مسئولیت اخلاقی به طور کامل بر دوش طراحان و کاربران انسانی این فناوری باقی بماند.

واژه‌های کلیدی: تکنیکی هوش مصنوعی، هوش انسانی، شناخت بدنمند، حکمت متعالیه، هیوبرت دریفوس، ملاصدرا.

فهرست مطالب

عنوان	شماره صفحه
چکیده	ح
فهرست مطالب	ی
فهرست جدول‌ها	م
فهرست شکل‌ها	ن
فصل اول: مقدمه	۱
۱,۱. طرح مسئله	۲
۱,۲. پرسش‌های پژوهش	۲
۱,۳. اهداف پژوهش	۳
۱,۴. اهمیت و ضرورت پژوهش	۳
۱,۵. روش‌شناسی و رویکرد پژوهش	۴
۱,۶. یادداشتی در باب فرآیند نگارش	۴
۱,۷. ساختار پژوهش	۴
فصل دوم: تاریخچه هوش مصنوعی و تعریف مسئله تکنیکی	۶
۲,۱. هوشمندی	۷
۲,۱,۱. هوش انسانی	۸
۲,۱,۲. هوش مصنوعی	۲۸
۲,۲. مسئله تکنیکی	۵۷
۲,۲,۱. سناریوهای تکنیکی	۵۷
۲,۲,۲. دلایل احتمالی برای نزدیکی تکنیکی	۵۸
۲,۲,۳. استدلال‌هایی در مخالفت با تکنیکی	۵۹

۶۰	 ۲,۲,۴. تحلیل فلسفی تکنیکی
۶۳	 ۲,۳. پیشینه پژوهش
۶۳	 ۲,۳,۱. پیشینه پژوهش در منابع انگلیسی
۶۴	 ۲,۳,۲. پیشینه پژوهش در منابع فارسی
۶۵	 ۲,۴. جمع‌بندی
۶۷	 فصل سوم: تکنیکی هوش مصنوعی؛ دیدگاه‌ها و استدلال‌های موافقان
۶۷	 ۳,۱. ریشه‌های مفهوم تکنیکی فناورانه
۶۷	 ۳,۱,۱. تکنیکی از دیدگاه جان فون نویمان
۶۹	 ۳,۱,۲. انفجار هوش (نخستین نظریه تکنیکی هوش مصنوعی از دیدگاه آی.جی. گود)
۷۰	 ۳,۱,۳. تکنیکی از دیدگاه ورنر وینج
۷۲	 ۳,۲. اجتناب‌ناپذیری تکنیکی
۷۲	 ۳,۲,۱. تکنیکی از دیدگاه ری کورزویل
۸۷	 ۳,۲,۲. تکنیکی از دیدگاه فرانک تیپلر
۹۱	 ۳,۳. برهان تکنیکی
۹۴	 ۳,۳,۱. هم‌ارزی
۹۷	 ۳,۳,۲. گسترش
۹۸	 ۳,۳,۳. تقویت
۱۰۰	 ۳,۳,۴. انفجار هوش بدون هوش
۱۰۳	 ۳,۳,۵. موانع تکنیکی
۱۰۵	 ۳,۴. جمع‌بندی
۱۰۷	 فصل چهارم: هوش مصنوعی و شناخت بدنمند
۱۰۸	 ۴,۱. نقد هیوبرت دریفوس بر هوش مصنوعی
۱۰۹	 ۴,۱,۱. شناسایی الگو
۱۰۹	 ۴,۱,۲. نقد دریفوس بر چهار پیش‌فرض بنیادین هوش مصنوعی
۱۱۶	 ۴,۱,۳. بدن، موقعیت و نیازهای انسانی

۴,۱,۴	پیامدها برای تکینگی و میراث دریفوس در عصر یادگیری عمیق	۱۱۹
۴,۲	دیدگاه فرانسیس هیلن: امتداد نقد دریفوسی در زمینه تکینگی	۱۲۱
۴,۳	جمع‌بندی	۱۲۴
فصل پنجم:	بررسی امکان هم‌ارزی هوش مصنوعی و هوش انسانی	۱۲۶
۵,۱	هستی‌شناسی صدرایی نفس	۱۲۷
۵,۱,۱	نفس: کمال اول برای جسم طبیعی آلی	۱۲۸
۵,۱,۲	جسمانیة الحدوث و روحانیة البقاء بودن نفس	۱۳۰
۵,۱,۳	حرکت جوهری	۱۳۲
۵,۱,۴	ترجمه هوشمندی	۱۳۴
۵,۲	صورت‌بندی استدلالی صدرایی به نفع امکان هم‌ارزی هوش مصنوعی و هوش انسانی	۱۳۶
۵,۲,۱	مرحله اول: امکان تحقق نفس در هوش مصنوعی از منظر صدرایی	۱۳۶
۵,۲,۲	مرحله دوم: کفایت نفس مفروض در هوش مصنوعی برای تحقق تکینگی	۱۳۸
۵,۳	نقد و بررسی استدلال صدرایی به نفع هم‌ارزی	۱۴۰
۵,۴	جمع بندی	۱۴۶
فصل ششم:	جمع‌بندی و نتیجه‌گیری	۱۴۷
۶,۱	خلاصه مسیر پژوهش	۱۴۸
۶,۲	تحلیل و تلفیق نتایج	۱۴۸
۶,۳	پاسخ به پرسش اصلی پژوهش	۱۵۰
۶,۴	پیامدها و چشم‌اندازهای آتی	۱۵۰
۶,۵	محدودیت‌های پژوهش و پیشنهادها برای تحقیقات آینده	۱۵۱
فهرست مراجع		۱۵۲
واژه‌نامه		۱۵۴
چکیده انگلیسی		۱۵۷
صفحه عنوان انگلیسی		۱۵۸

فهرست جدول‌ها

جدول ۱. مفهوم هوش در زبان و بسترهای مختلف.....	۹
جدول ۲. دیدگاه‌های فلسفی تأثیرگذار بر هوش انسانی.....	۱۳
جدول ۳. خلاصه نظرات مطرح روانشناختی درباره هوش.....	۱۴
جدول ۴. نمونه نظریه‌های شامل استعاره جغرافیایی.....	۱۸
جدول ۵. خلاصه‌ای از استعاره‌ها و کاربرد آن در بحث‌های آینده.....	۲۶
جدول ۶. خلاصه رویکردهای مطالعه و توسعه هوش مصنوعی.....	۳۰
جدول ۷. رویکردهای هوش مصنوعی بر-اساس نوع پیاده‌سازی.....	۳۵
جدول ۸. انواع یادگیری ماشین.....	۴۴
جدول ۹. مقایسه امواج هوش مصنوعی.....	۴۶
جدول ۱۰. انواع هوش مصنوعی عمومی.....	۴۷
جدول ۱۱. مقایسه هوش مصنوعی محدود و عمومی.....	۵۰
جدول ۱۲. مقایسه هوش زیستی و هوش مصنوعی.....	۵۱
جدول ۱۳. مقایسه جزئیتر با استفاده از اطلاعات منابع مورد بررسی.....	۵۲
جدول ۱۴. وعده‌ها و خطرات بالقوه تکینگی از دیدگاه کورزویل.....	۸۲
جدول ۱۵. پیش‌فرض‌های بنیادین هوش مصنوعی و نقد دریفوسی.....	۱۱۰
جدول ۱۶. همگرایی انتقادی پدیدارشناسی و حکمت متعالیه در نقد هوش مصنوعی.....	۱۴۷

فهرست شکل‌ها

- شکل ۱. نموداری از اندیشه‌های فلسفی درباره ذهن و روابط آنها..... ۱۱
- شکل ۲. دیاگرام یک شبکه عصبی عمیق..... ۱۲
- شکل ۳. CATTELL-HORN-CARROLL'S THREE STRATUM MODEL..... ۱۹
- شکل ۴. مدلی از فرآیند محاسباتی..... ۲۰
- شکل ۵. مغز و برخی نواحی کلیدی مرتبط با هوش..... ۲۱
- شکل ۶. منطقه توسعه تقریبی..... ۲۴
- شکل ۷. ویژگی‌های سیستم‌های هوشمند..... ۳۰
- شکل ۸. هوش مصنوعی بر اساس قابلیت‌ها..... ۳۴
- شکل ۹. رشته‌های تأثیرگذار در هوش مصنوعی..... ۳۶
- شکل ۱۰. مدلی از فرآیندهای تصمیم‌مارکوف..... ۳۹
- شکل ۱۱. بخشی از خط زمانی هوش مصنوعی..... ۴۲
- شکل ۱۲. استدلال چالمرز برای تکینگی..... ۶۱
- شکل ۱۳. A SIMPLE FORM OF MACHINE SELF-REPLICATION..... ۶۸
- شکل ۱۴. مقایسه رشد خطی و نمایی در فناوری..... ۷۴
- شکل ۱۵. نمونه از روند قیمت-عملکرد محاسبات، ۱۹۳۹-۲۰۲۳..... ۷۵
- شکل ۱۶. منحنی‌های S و شش مرحله تکامل..... ۸۱
- شکل ۱۷. مدل بهبود بازگشتی خود (پایه انفجار هوش)..... ۹۲
- شکل ۱۸. مراحل کلیدی در برهان تکینگی چالمرز..... ۹۴
- شکل ۱۹. نمودار ساده نمودار فرآیندی انفجار ظرفیت‌ها..... ۱۰۲
- شکل ۲۰. مدل پنج مرحله‌ای اکتساب مهارت دریفوس..... ۱۱۴

فصل اول

مقدمه

۱. مقدمه

در آغاز قرن بیست و یکم، بشریت خود را در آستانه تحولی می‌یابد که شاید در تاریخ خود بی‌سابقه باشد. هوش مصنوعی^۱، که روزگاری موضوع داستان‌های علمی-تخیلی و مباحثات نظری در آزمایشگاه‌های علوم کامپیوتر بود، اکنون به سرعتی شگرف در حال نفوذ به تمامی شئون زندگی فردی و اجتماعی است. از الگوریتم‌هایی که محتوای دیجیتال ما را مدیریت می‌کنند و الگوهای مصرف ما را پیش‌بینی می‌کنند و نیز سیستم‌هایی که قادر به تشخیص بیماری‌ها و راندن خودروها هستند تا مدل‌های زبانی بزرگی که قادر به تولید متن، تحلیل داده و حتی کدنویسی هستند. هوش مصنوعی از یک مفهوم علمی-تخیلی به واقعیتی فراگیر و تأثیرگذار و از یک ابزار محاسباتی صرف، به یک نیروی فعال و دگرگون‌ساز بدل شده است.

این پیشرفت شتابان، افقی را پیش روی ما گشوده است که همزمان الهام‌بخش و نگران‌کننده است؛ که از آن با عنوان «تکینگی فناوریانه»^۲ یاد می‌شود. این مفهوم، که توسط آینده‌پژوهان و نظریه‌پردازان برجسته‌ای چون ری کورزویل^۳ و دیوید چالمرز^۴ به تفصیل شرح داده شده، به نقطه‌ای فرضی در آینده اشاره دارد که در آن، هوش ماشینی نه تنها با هوش انسانی هم‌ارز خواهد بود، بلکه از آن پیشی می‌گیرد. در این سناریو، هوش مصنوعی با ورود به یک چرخه بازگشتی و نمایی از خودبهبودبخشی، به یک انفجار هوشی دست می‌یابد که پیامدهای آن برای تمدن، فرهنگ و حتی ماهیت انسان، عمیقاً ناشناخته و غیرقابل پیش‌بینی است.

۱.۱. طرح مسئله

چشم‌انداز تکینگی، جامعه فکری را به دو قطب متمایز تقسیم کرده است. از یک سو، موافقان تکینگی با استناد به روندهای نمایی در قدرت محاسباتی و پیشرفت‌های خیره‌کننده در یادگیری ماشین، وقوع آن را اجتناب‌ناپذیر یا دست‌کم بسیار محتمل می‌دانند و آن را گام بعدی در تکامل هوش در کیهان تلقی می‌کنند. از سوی دیگر، منتقدان، این دیدگاه را بیش از حد خوش‌بینانه و مبتنی بر مفروضاتی ساده‌انگارانه از ماهیت پیچیده هوش، آگاهی و وجود انسانی می‌دانند.

گفتمان پیرامون تکینگی، اغلب میان دو قطب خوش‌بینی آرمانشهری (غلبه بر بیماری، فقر و حتی مرگ) و بدبینی ویرانشهری (خطر کنترل‌ناپذیری و انقراض انسان) در نوسان است. با این حال، هر دو دیدگاه بر یک پیش‌فرض بنیادین و اغلب ناگفته امکان‌پذیری فلسفی و فنی هم‌ارزی هوش مصنوعی با هوش انسانی استوارند. واضح است که، برای سخن گفتن از پیشی گرفتن از انسان و وقوع یک انفجار هوشی، تا زمانی که ماشین نتواند به هوشی در سطح عمومی، خلاق و سازگارانه انسان دست یابد، بی‌معنا خواهد بود.

این پژوهش، با فاصله گرفتن از پیش‌بینی‌های صرفاً فناوریانه، به این پیش‌فرض بنیادین می‌پردازد و این پرسش محوری را مطرح می‌کند: تکینگی چیست و پیش‌بینی امکان‌پذیری آن بر چه مفروضاتی بنا شده است؟ و آیا هم‌ارزی هوش مصنوعی با هوش انسانی (به عنوان مقدمه‌ای اساسی در این رویداد) از منظر فلسفی امکان‌پذیر است؟

به بیان دیگر، مسئله اصلی این پژوهش، نه صرفاً بررسی فنی امکان‌سنجی تکینگی، بلکه کاوش در بنیان‌های فلسفی این ایده است. پرسش محوری این است که آیا «هم‌ارزی» میان هوش انسانی (با تمام ابعاد بدنمند، موقعیت‌مند، شهودی و غایت‌مند آن) و هوش مصنوعی (که ذاتاً پدیده‌ای محاسباتی و برساخته است) اساساً امکان‌پذیر است؟ پاسخ به این پرسش، سنگ بنای هرگونه بحث معنادار درباره تکینگی است. اگر این هم‌ارزی از منظر فلسفی با موانع جدی روبرو باشد، آنگاه کل بنای استدلال تکینگی متزلزل خواهد شد.

^۱ Artificial Intelligence (AI)

^۲ Technological Singularity

^۳ Ray Kurzweil

^۴ David Chalmers

۱.۲. پرسش‌های پژوهش

- این پژوهش در صدد است تا با رویکردی تحلیلی و تطبیقی، به پرسش‌های زیر پاسخ دهد:
- آیا هم‌ارزی هوش مصنوعی با هوش انسانی از منظر فلسفی و در نتیجه وقوع تکینگی فناورانه امکان‌پذیر است؟
- برای رسیدن به پاسخ این پرسش اصلی، پرسش‌های فرعی زیر مورد کاوش قرار خواهند گرفت:
۱. مبانی نظری و استدلال‌های کلیدی حامیان نظریه تکینگی چیست؟
 ۲. از منظر پدیدارشناسی و نظریه «شناخت بدنمند» (به‌ویژه در آراء هیوبرت دریفوس)، چه نقدهای بنیادینی بر امکان هم‌ارزی هوش مصنوعی و انسانی وارد است؟
 ۳. از منظر حکمت متعالیه ملاصدرا و با تکیه بر مفاهیمی چون «نفس»، «حرکت جوهری» و «علم حضوری»، چه موانع و امکان‌هایی فراروی هم‌ارزی هوش مصنوعی و انسانی وجود دارد؟
 ۴. نتایج حاصل از این دو رویکرد انتقادی (پدیدارشناسی و حکمت متعالیه) چه تصویری از آینده هوش مصنوعی و نسبت آن با انسان ارائه می‌دهند؟

۱.۳. اهداف پژوهش

هدف اصلی:

- ارزیابی انتقادی امکان‌پذیری فلسفی هم‌ارزی هوش مصنوعی با هوش انسانی به عنوان پیش‌فرض بنیادین نظریه تکینگی.

اهداف فرعی:

۱. صورت‌بندی و تحلیل مبانی نظری و استدلال‌های کلیدی حامیان تکینگی فناورانه.
۲. تبیین و تحلیل نقد پدیدارشناسانه (با تأکید بر آراء دریفوس) بر امکان هوش مصنوعی هم‌سطح انسان.
۳. کاوش و استنتاج موانع و امکان‌های هم‌ارزی هوش مصنوعی و انسانی از منظر حکمت متعالیه ملاصدرا.

۱.۴. اهمیت و ضرورت پژوهش

اهمیت این پژوهش از چند منظر قابل توجیه و در سه سطح قابل تبیین است:

- **سطح میان‌رشته‌ای و تطبیقی؛ خلأ پژوهشی:** با وجود حجم انبوهی از مطالعات فنی و آینده‌پژوهانه در باب هوش مصنوعی، تحلیل‌های عمیق فلسفی که به صورت تطبیقی به نقد بنیان‌های این حوزه بپردازند، به‌ویژه پژوهش‌هایی که میان سنت‌های فکری غرب و حکمت اسلامی پل بزنند، نادر و پراکنده‌اند. این تحقیق می‌کوشد تا این خلأ را تا حدی پر کند. این تحقیق با ایجاد یک گفتگوی انتقادی میان دو سنت فکری قدرتمند اما ظاهراً بی‌ارتباط—پدیدارشناسی غربی (به‌ویژه در نقد هیوبرت دریفوس) و حکمت متعالیه اسلامی (در آراء ملاصدرا)—برای پاسخ به یک مسئله معاصر، گامی نو در جهت کاربست ظرفیت‌های فلسفه برای فهم جهان امروز برمی‌دارد.
- **سطح عملی و راهبردی؛ ضرورت اخلاقی:** نحوه پاسخ ما به پرسش هم‌ارزی، مستقیماً بر رویکرد اخلاقی ما به هوش مصنوعی تأثیر می‌گذارد. اگر هوش مصنوعی یک ابزار پیشرفته است، مسئولیت اخلاقی کاملاً بر دوش انسان است. اما اگر بتواند به یک فاعل هم‌ارز تبدیل شود، مسائل پیچیده‌ای در باب حقوق، مسئولیت و جایگاه آن مطرح می‌شود. روشن‌سازی این مبانی، پیش‌شرط تدوین هرگونه اخلاق کاربردی معنادار در این حوزه است. درک عمیق از امکان یا امتناع هم‌ارزی، مستقیماً بر نحوه سیاست‌گذاری، تخصیص منابع و تدوین چارچوب‌های اخلاقی برای توسعه هوش مصنوعی تأثیرگذار است. اگر هم‌ارزی ناممکن باشد، تمرکز باید از رقابت با انسان به سمت ساخت ابزارهای مکمل (هوش افزوده) معطوف گردد و دغدغه‌های اخلاقی نیز از «حقوق ماشین» به «مسئولیت انسان» منتقل می‌شود.
- **سطح نظری و فلسفی؛ اهمیت انسان‌شناختی:** مواجهه با «دیگری» هوشمند مصنوعی، ما را بیش از هر زمان

دیگری وادار به تأمل در باب چیستی خودمان می‌کند. این پژوهش، با کاوش در مرزهای هوش ماشینی، در نهایت به دنبال درکی عمیق‌تر از ویژگی‌های منحصر به فرد هوش انسانی، بدن‌مندی و معنای وجود انسان است. این مسئله، ما را وادار به بازاندیشی در بنیادین‌ترین پرسش‌ها درباره ماهیت انسان می‌کند: هوش چیست؟ نقش بدن در شناخت چیست؟ و مرز میان موجود طبیعی و مصنوع کجاست؟ این پژوهش با درگیر شدن با این پرسش‌ها، به غنای مباحث در فلسفه ذهن، فلسفه فناوری و انسان‌شناسی فلسفی کمک می‌کند.

۱.۵. روش‌شناسی و رویکرد پژوهش

روش تحقیق در این پایان‌نامه، «توصیفی-تحلیلی» با رویکردی «تطبیقی و میان‌رشته‌ای» است. در این مسیر، ابتدا با مطالعه منابع اصلی در حوزه هوش مصنوعی و آینده‌پژوهی، دیدگاه‌های موافقان تکنیکی توصیف و صورت‌بندی می‌شود. سپس با رجوع به متون کلیدی در سنت پدیدارشناسی و حکمت متعالیه، مبانی این دو مکتب برای نقد دیدگاه نخست استخراج و تحلیل می‌گردد. در نهایت، با تلفیق و مقایسه نتایج حاصل از این دو نقد، پاسخی مستدل به پرسش اصلی پژوهش ارائه خواهد شد. در بخش‌های توصیفی، به شرح دقیق دیدگاه‌های فنی و فلسفی مرتبط با هوش مصنوعی و تکنیکی پرداخته می‌شود. در بخش‌های تحلیلی، این دیدگاه‌ها مورد واکاوی و نقد قرار می‌گیرند. محور اصلی پژوهش، رویکرد تطبیقی است که در آن، استدلال‌های موافقان تکنیکی در برابر دو سنت انتقادی نیرومند، یعنی پدیدارشناسی و حکمت متعالیه، قرار داده می‌شوند تا نقاط ضعف و قوت هر یک آشکار گردد. منابع اصلی تحقیق شامل آثار کلیدی نظریه‌پردازان تکنیکی (مانند کورزویل و چالمرز) و فیلسوفان منتقد (مانند دریفوس و ملاصدرا) و همچنین مقالات و کتب معتبر در حوزه‌های مرتبط است.

۱.۶. یادداشتی در باب فرآیند نگارش

شایان ذکر است که بخش محدودی از فرآیند نگارش این پژوهش، خود نمونه‌ای از تعامل میان هوش انسانی و هوش مصنوعی بوده است. در برخی مراحل این تحقیق، از جمله ویرایش نگارشی جملات بخش‌هایی از متن، و تنظیم ساختار مباحث، از مدل‌های زبانی پیشرفته کمک گرفته شده است. با این حال، تأکید می‌گردد که مسئولیت نهایی جمع‌آوری اطلاعات، خوانش و خلاصه‌سازی، تحلیل انتقادی، صورت‌بندی استدلال‌ها، تلفیق نتایج و جهت‌گیری کلی پژوهش، کاملاً بر عهده نویسنده انسانی بوده است. این امر، خود شاهدی عملی بر نتیجه‌گیری نهایی این رساله است؛ هوش مصنوعی، در وضعیت کنونی، ابزاری بسیار قدرتمند برای تقویت و تسریع اندیشه انسانی است، نه جایگزینی برای آن.

۱.۷. ساختار پژوهش

این پایان‌نامه در شش فصل به شرح زیر سازماندهی شده است:

- فصل اول: فصل حاضر، شمایی کلی از این پژوهش را ارائه می‌دهد.
- فصل دوم: به تعریف مفاهیم کلیدی همچون هوش، هوش مصنوعی و تکنیکی و بررسی تاریخیچه و ابعاد مختلف آن‌ها می‌پردازد تا زمینه مفهومی لازم برای ورود به بحث اصلی فراهم شود.
- فصل سوم: دیدگاه‌ها و استدلال‌های موافقان تکنیکی، از پیشگامان اولیه تا نظریه‌پردازان برجسته معاصر مانند ری کورزویل و دیوید چالمرز، را به تفصیل شرح و تحلیل می‌کند.
- فصل چهارم: نخستین نقد بنیادین بر امکان هم‌ارزی را از منظر فلسفه غرب و با تمرکز بر دیدگاه «شناخت بدمنند» و نقدهای فلسفی هیوبرت دریفوس بر هوش مصنوعی ارائه می‌دهد.
- فصل پنجم: دومین نقد را از منظری متفاوت و با تکیه بر مبانی «حکمت متعالیه» و هستی‌شناسی نفس در فلسفه ملاصدرا صورت‌بندی کرده و موانع متافیزیکی هم‌ارزی را تبیین می‌کند.

- **فصل ششم:** به جمع‌بندی نهایی، تلفیق نتایج حاصل از دو جبهه نقد، ارائه پاسخی جامع به پرسش اصلی پژوهش، و طرح پیامدها و پیشنهادهایی برای تحقیقات آینده اختصاص دارد.

فصل دوم

تاریخچه هوش مصنوعی و تعریف مسئله تکینگی

۲. تاریخچه هوش مصنوعی و تعریف مسئله تکینگی

فصل حاضر به بررسی مفهوم "هوشمندی"^۱ و تاریخچه و جوانب هوش مصنوعی برای تعریف و شناسایی مسئله تکینگی می‌پردازد. هدف از این فصل، کاوش در تعاریف، نظریه‌ها و رویکردهای گوناگون به هوشمندی، از خاستگاه‌های آن در هوش انسانی تا تجلیات نوین آن در هوش مصنوعی است. این بررسی با هدف ایجاد درکی عمیق‌تر از ماهیت هوش مصنوعی و زمینه‌سازی برای مباحث آتی در مورد تکینگی هوش مصنوعی صورت می‌گیرد.

در این فصل، ابتدا به تعاریف اولیه هوش مصنوعی و چالش‌های ناشی از تعریف آن صرفاً بر اساس شباهت رفتاری با انسان پرداخته می‌شود. سپس، با ورود به مبحث هوش انسانی، به بررسی ریشه‌شناختی و معناشناختی مفهوم "هوش" در بسترهای زبانی و فرهنگی مختلف پرداخته و تصورات فلسفی از ذهن و هوش انسان در طول تاریخ مورد واکاوی قرار می‌گیرد. در ادامه، نظریه‌های برجسته روان‌شناسی در مورد هوش، از جمله نظریه عامل g اسپیرمن و نظریه هوش‌های چندگانه گاردنر، معرفی و تحلیل می‌شوند.

بخش بعدی فصل، به بررسی استعاره‌های غالب در مطالعات هوش اختصاص دارد. استعاره‌هایی چون استعاره جغرافیایی، محاسباتی، زیستی، معرفت‌شناختی، جامعه‌شناختی، انسان‌شناختی و سیستمی، هر یک از منظری خاص به تبیین هوش می‌پردازند و درک ما از این پدیده پیچیده را غنی‌تر می‌سازند.

پس از بررسی جامع هوش انسانی، به مبحث هوش مصنوعی پرداخته می‌شود. در این بخش، ضمن ارائه تعاریف مختلف هوش مصنوعی، رویکردهای گوناگون به آن، از جمله رویکرد رفتارگرایانه، مدل‌سازی شناختی، قوانین فکر و عامل منطقی، مورد بررسی قرار می‌گیرند. همچنین، دسته‌بندی‌های مختلف هوش مصنوعی بر اساس قابلیت‌ها (هوش مصنوعی محدود، عمومی و ابرهوش) و نوع پیاده‌سازی (نمادین، اتصال‌گرا و ...) معرفی می‌شوند.

در ادامه، رشته‌های تأثیرگذار بر هوش مصنوعی، از جمله فلسفه، ریاضیات، اقتصاد، علوم اعصاب، روان‌شناسی، مهندسی کامپیوتر، نظریه کنترل و سایبرنتیک، و زبان‌شناسی، به اختصار معرفی و نقش هر یک در شکل‌گیری و توسعه این حوزه تبیین می‌گردد. سپس، زمینه‌های اصلی تحقیقات در هوش مصنوعی، مانند بازنمایی دانش، استدلال خودکار، یادگیری ماشین، پردازش زبان طبیعی، بینایی ماشین، رباتیک، برنامه‌ریزی و سیستم‌های خبره، مورد بررسی قرار می‌گیرند.

بخش پایانی فصل، به بررسی چالش‌های بنیادین در هوش مصنوعی، از جمله ماهیت و محدودیت‌های محاسبات، نقش بازنمایی و دانش، و چالش ارزیابی سیستم‌های هوشمند، اختصاص دارد. همچنین، به مباحث مهمی چون هوش مصنوعی عمومی، مسئله تکینگی، و مقایسه هوش زیستی و هوش مصنوعی پرداخته می‌شود.

در مجموع، این فصل با ارائه تصویری جامع از مفهوم هوشمندی، و با بررسی چالش‌ها و مباحث کلیدی این حوزه، زمینه‌ساز ورود به بحث، در فصول آتی خواهد بود.

۲.۱. هوشمندی

هوش مصنوعی در ابتدای مسیر خود، اغلب با هدف تقلید^۲ و بازآفرینی رفتارهای هوشمندانه مشاهده‌شده در انسان تعریف می‌شد (Schönecker, 2022, p. 5). این دیدگاه اولیه، هوش مصنوعی را به مثابه تلاش برای ساخت ماشین‌هایی می‌نگریست که قادر به انجام وظایفی باشند که در صورت انجام توسط انسان، «هوشمندانه»^۳ تلقی می‌شدند. این تعریف بر جنبه عملکردی هوش تأکید داشت؛ یعنی هوش مصنوعی زمانی محقق می‌شد که سیستم‌های کامپیوتری بتوانند در موقعیت‌های مشخص، رفتاری مشابه انسان از خود بروز دهند. برای مثال، یک برنامه شطرنج‌باز که در سطح قهرمانان جهانی بازی می‌کند، یا یک سیستم

^۱ Intelligence

^۲ Imitation

^۳ Intelligently