

ترسیم ساختار مفهومی و سیر تحول موضوعی پژوهش‌های فارسی نمایه‌شده در پایگاه آی‌اس‌سی در حوزه یادگیری ماشین: یک مطالعه علم‌سنجی

حامد برنگی

دکتری مهندسی کامپیوتر،

گروه مهندسی کامپیوتر، دانشگاه پیام نور، تهران، ایران

علی اکبر خاصه *

دانشیار؛ گروه علم اطلاعات و دانش‌شناسی، دانشگاه پیام نور، تهران، ایران

دریافت: ۱۴۰۵/۰۳/۳۰ | پذیرش: ۱۴۰۵/۰۶/۱۰ | مقاله برای اصلاح به مدت ۱۵ روز نزد پدیدآوران بوده است.

نشریه علمی (رتبه بین‌المللی)
پژوهشگاه علوم و فناوری اطلاعات ایران
شاخه (پایه) ۸۲۲۳-۲۲۵۱
شاخه (الکترونیکی) ۸۲۳۱-۲۲۵۱
نمایه در: ISC و SCOPUS, LISTA
<http://jipm.irandoc.ac.ir>
دوره XX | شماره X | صص XX-XX
۱۳XX X

نوع مقاله: پژوهشی

به این مقاله به شکل زیر استناد کنید:

درون متن:

(برنگی و خاصه، زودآیند)

در فهرست منابع:

برنگی، حامد، و خاصه، علی اکبر. زودآیند. ترسیم ساختار مفهومی و سیر تحول موضوعی پژوهش‌های فارسی نمایه‌شده در پایگاه آی‌اس‌سی در حوزه یادگیری ماشین: یک مطالعه علم‌سنجی. پژوهشنامه پردازش و مدیریت اطلاعات.

<http://jipm.irandoc.ac.ir> (دسترسی در

روز/ماه/سال)

چکیده: یادگیری ماشین به عنوان یکی از مهم‌ترین شاخه‌های هوش مصنوعی، طی دهه اخیر رشد چشمگیری در تولیدات علمی ایران داشته است. با وجود افزایش تعداد پژوهش‌ها در حوزه یادگیری ماشین در ایران، مطالعات علم‌سنجی محدودی درباره ساختار دانش و سیر تحول موضوعی پژوهش‌های فارسی در حوزه یادگیری ماشین وجود دارد. هدف پژوهش حاضر، تحلیل و نگاشت علمی مدارک نمایه‌شده به زبان فارسی در حوزه یادگیری ماشین در پایگاه استنادی و پایتخت علم و فناوری جهان اسلام (آی‌اس‌سی) طی سال‌های ۱۳۹۲ تا ۱۴۰۳ است. پژوهش حاضر از نوع مطالعات علم‌سنجی با رویکرد توصیفی-تحلیلی است. داده‌های پژوهش از پایگاه آی‌اس‌سی گردآوری شد. پس از جستجو و پالایش داده‌ها، ۳۹۹ مقاله مرتبط با یادگیری ماشین برای تحلیل شناسایی شد. به منظور افزایش دقت نتایج، کلیدواژه‌ها و وابستگی‌های سازمانی استانداردسازی شدند. تحلیل داده‌ها با استفاده از نرم‌افزار بیبلیوشاپی و ووز-ویور انجام شد و شاخص‌هایی نظیر روند تولید علم، نمودار سه‌بخشی نویسندگان-مجلات-کلیدواژه‌ها، شبکه همکاری دانشگاه‌ها، تحلیل هم‌رخدادی کلیدواژه‌ها، نقشه تکامل موضوعی و نمودار راهبردی مورد بررسی قرار گرفت. نتایج نشان داد تعداد مقالات منتشرشده از یک مقاله در سال ۱۳۹۲ به ۱۰۶ مقاله در سال ۱۴۰۲ افزایش یافته است که بیانگر رشد سریع این حوزه در ایران است. از بین ۱۳۶ مؤسسه/دانشگاه که دارای مقاله در حوزه یادگیری ماشین بودند،

دانشگاه آزاد اسلامی، دانشگاه تهران و دانشگاه تربیت مدرس به ترتیب با ۶۵، ۱۰۱ و ۲۳ مقاله، فعال‌ترین مراکز علمی در این حوزه بودند. علاوه بر این، از میان ۱۸۷۳ کلیدواژه شناسایی شده، کلیدواژه‌های «ماشین بردار پشتیبان»، «جنگل تصادفی» و «شبکه عصبی مصنوعی» به ترتیب با فراوانی ۶۰، ۵۹ و ۴۳، پرتکرارترین موضوعات پژوهشی محسوب می‌شوند. بررسی سیر تحول موضوعات بیانگر حرکت تدریجی پژوهش‌ها از مفاهیم پایه‌ای و الگوریتم‌های کلاسیک به سمت کاربردهای تخصصی‌تر و موضوعاتی نظیر یادگیری عمیق، مدل‌سازی و کاربردهای محیط‌زیستی و مخاطرات طبیعی حرکت کرده است. همچنین تحلیل شبکه هم‌رخدادی کلیدواژه‌ها چهار خوشه اصلی شامل کاربردهای محیط‌زیستی، توسعه الگوریتم‌ها، کاربردهای میان‌رشته‌ای و مطالعات مبتنی بر سنجش از دور را آشکار ساخت. نتایج نمودار راهبردی نیز نشان داد موضوعات «دسته‌بندی»، «پیش‌بینی» و «هوش مصنوعی» در زمره مضامین بنیادی قرار دارند، در حالی که خوشه‌های مرتبط با «ماشین بردار پشتیبان»، «جنگل تصادفی» و «سنجش از دور» به عنوان موضوعات تخصصی شناخته می‌شوند. یافته‌ها نشان می‌دهد پژوهش‌های فارسی حوزه یادگیری ماشین در ایران به مرحله‌ای از رشد نسبی رسیده‌اند و مسیر تحول آن‌ها از توسعه الگوریتم‌های پایه به سمت کاربردهای مسئله‌محور و میان‌رشته‌ای در حال حرکت است.

کلیدواژه‌ها: یادگیری ماشین، هوش مصنوعی، ساختار دانش، علم سنجی، تحلیل هم‌رخدادی کلیدواژه‌ها، تکامل موضوعی

*پدیدآور رابط khasseh@pnu.ac.ir

۱. مقدمه

در دهه‌های اخیر، یادگیری ماشین به عنوان یکی از بنیادی‌ترین و پویاترین شاخه‌های هوش مصنوعی، تحولی بنیادین در علوم محاسباتی، مهندسی، پزشکی، اقتصاد و بسیاری از حوزه‌های دیگر ایجاد کرده است (Jordan & Mitchell, 2015). یادگیری ماشین به مجموعه‌ای از الگوریتم‌ها و روش‌های آماری اطلاق می‌شود که به سیستم‌های رایانه‌ای این توانایی را می‌دهد که بدون برنامه‌نویسی صریح، از داده‌ها بیاموزند و عملکرد خود را به تدریج بهبود بخشند (Bishop, 2006). در این راستا، یادگیری ماشین به عنوان یکی از زیرشاخه‌های اصلی هوش مصنوعی شناخته می‌شود که ریشه‌های نظری آن به شدت بر مفاهیم و روش‌های آماری استوار است (سهرابی و دیگران، ۱۳۹۵). این حوزه، که ریشه‌های آن به تحقیقات آلن تورینگ در دهه ۱۹۵۰ میلادی بازمی‌گردد (Turing, 2009)، اکنون با ظهور یادگیری عمیق، شبکه‌های عصبی پیچیده و پردازش داده‌های

کلان، در قلب انقلاب دیجیتال قرن بیست‌ویکم قرار گرفته است (LeCun et al., 2015). در عصر تحول دیجیتال، تکنیک‌های یادگیری ماشین در هر صنعتی مفید واقع می‌شوند و برای حل مشکلات پیچیده‌ای در صنایع دیگری نظیر بهداشت و درمان، سرمایه‌گذاری و بورس، و بیمه مورد استفاده قرار گرفته‌اند (باغانی و دیگران، ۱۴۰۲).

پیشرفت در توان پردازشی، دسترسی به داده‌های کلان و توسعه روش‌های یادگیری عمیق، موجب گسترش کاربردهای یادگیری ماشین در حوزه‌های مختلف شده است. در پی این تحول، توجه جامعه علمی به این حوزه افزایش یافته و تعداد انتشارات مرتبط نیز به‌طور مستمر رشد کرده است. مطالعات نشان می‌دهد که تعداد مقالات مرتبط با یادگیری ماشین در عرصه بین‌المللی به‌طور مستمر افزایش یافته است (Belmonte et al. 2020; Lundberg et al., 2024).

ایران در عرصه تولید علم هوش مصنوعی جایگاه قابل‌توجهی دارد. بر اساس داده‌های جهانی پایگاه سایمگو^۱ در بازه ۱۹۹۶ تا ۲۰۲۵، ایران با ۲۳,۹۰۶ مدرک علمی و ۳۶۷,۵۷۳ استاندارد، رتبه شانزدهم را در جهان و رتبه نخست را در منطقه خاورمیانه در حوزه هوش مصنوعی به خود اختصاص داده است (SCImago Journal & Country Rank, 2025). سیاست‌گذاری‌های کلان علمی و تدوین اسناد ملی در حوزه هوش مصنوعی، همراه با سرمایه‌گذاری در تربیت نیروی انسانی متخصص، نشان‌دهنده عزم کشور برای ارتقای جایگاه علمی در این حوزه است (اکبری و جلال‌نیا، ۱۴۰۴). در همین راستا، بررسی وضعیت و الگوی تحول پژوهش‌های فارسی نمایه‌شده در حوزه یادگیری ماشین می‌تواند تصویری روشن از روند توسعه این حوزه در کشور ارائه دهد. در این میان، مؤسسه استنادی و پایش علم و فناوری جهان اسلام (آی‌اس‌سی) به‌عنوان یکی از مهم‌ترین نظام‌های استنادی منطقه‌ای، نقش کلیدی در نمایه‌سازی و ارزیابی تولیدات علمی کشورهای اسلامی ایفا می‌کند. این پایگاه با پوشش نشریات معتبر فارسی‌زبان و ارائه شاخص‌های استنادی و علم‌سنجی، بستر نسبتاً مناسبی برای تحلیل ساختار و روندهای تولید علم در ایران فراهم کرده است. به‌ویژه، تمرکز بر مجلات فارسی معتبر در این پایگاه، آن را به منبعی مناسب برای بررسی وضعیت پژوهش‌های یادگیری ماشین در سطح ملی تبدیل می‌کند (Mehrad & Ghane, 2021). البته پایگاه‌های دیگری نیز در ایران وجود دارند که در حوزه نمایه‌سازی پژوهش‌های فارسی اقدامات ارزنده‌ای انجام داده‌اند که از آن جمله می‌توان به پژوهشگاه علوم و فناوری

¹ SCImago

اطلاعات ایران (ایرانداک)، پایگاه مرکز اطلاعات علمی جهاد دانشگاهی (اس‌آی‌دی)، پایگاه مجلات تخصصی نور (نورمگز)، بانک اطلاعات نشریات کشور (مگیران)، پرتال جامع علوم انسانی، سیولیکا (مقالات علمی کنفرانس و ژورنال) و برخی موارد دیگر اشاره نمود؛ لکن ارائه برخی اطلاعات خاص علم‌سنجی از جمله ضریب تاثیر سالیانه نشریات و وضعیت استنادی مقالات توسط آی‌اس‌سی و همچنین اتکای برخی آیین‌نامه‌های دانشگاهی به آمار ارائه شده توسط پایگاه آی‌اس‌سی، آن را تا حدودی از سایر پایگاه‌ها متمایز کرده است.

علم‌سنجی به عنوان رویکردی کمی برای تحلیل ساختار و پویایی علم، امکان بررسی الگوهای تولید، همکاری و تحول موضوعی را فراهم می‌سازد (Mingers & Leydesdorff, 2015). این رویکرد با فوننی از قبیل تحلیل هم‌استنادی، تحلیل هم‌واژگانی^۱ و تحلیل شبکه‌ای همکاری‌های علمی، قادر است ساختار و تکامل مفهومی یک حوزه علمی را آشکار سازد (Leydesdorff, 2001; Mokhtarpour & Khasseh, 2020). در این میان، تحلیل هم‌رخدادی کلیدواژه‌ها، که برای شناسایی ساختار مفهومی و موضوعات کلیدی، نوظهور و بنیادین یک حوزه به کار می‌رود (Cobo et al., 2017; Khasseh et al. 2011) از جایگاه خاصی برخوردار است. روی هم رفته، این فنون ابزارهای مؤثری برای درک عمیق‌تر از وضعیت و مسیر توسعه حوزه‌های علمی محسوب می‌شوند.

علیرغم رشد قابل توجه پژوهش‌ها در زمینه یادگیری ماشین، مطالعات علم‌سنجی درباره ترسیم ساختار دانش و سیر تحول موضوعی پژوهش‌های فارسی‌زبان این حوزه همچنان محدود است. این خلأ پژوهشی ضرورت انجام مطالعه‌ای نظام‌مند را برای ترسیم وضعیت موجود و شناسایی روندهای آینده این حوزه برجسته می‌سازد. از این رو، پژوهش حاضر با بهره‌گیری از روش‌های علم‌سنجی و ابزارهای تحلیل شبکه، به بررسی پژوهش‌های فارسی حوزه یادگیری ماشین نمایه شده در آی‌اس‌سی می‌پردازد. پرسش اساسی در پژوهش حاضر این است که ساختار مفهومی و سیر تکامل موضوعی پژوهش‌های فارسی‌زبان حوزه یادگیری ماشین در بازه زمانی ۱۳۹۲ تا ۱۴۰۳ چگونه است و شکاف‌ها و افق‌های پیش روی این حوزه کدامند؟

به منظور دستیابی به این هدف، پژوهش حاضر در پی پاسخ به پرسش‌های زیر است:

۱. روند انتشار مقالات فارسی در حوزه یادگیری ماشین چگونه است؟

¹ Co-word Analysis

۲. بر اساس نمودار سه‌بخشی^۱، روابط میان نویسندگان، مجلات و کلیدواژه‌ها در پژوهش‌های حوزه یادگیری ماشین چگونه است؟

۳. فعال‌ترین دانشگاه‌ها در تولیدات علمی حوزه یادگیری ماشین کدام‌اند و شبکه همکاری میان آن‌ها چه ساختاری دارد؟

۴. سیر تحول و تکامل موضوعات پژوهشی در حوزه یادگیری ماشین چگونه بوده است؟

۵. خوشه‌های موضوعی اصلی یادگیری ماشین در تحلیل شبکه هم‌رخدادی کلیدواژه‌ها کدامند و بر اساس نمودار راهبردی در چه وضعیتی (محرك، بنیادی، تخصصی و نوظهور) قرار دارند؟

در ادامه، ساختار مقاله به این صورت تنظیم شده است. ابتدا مبانی نظری و سپس پیشینه پژوهش ارائه می‌شود. در بخش روش‌شناسی، داده‌ها و رویکرد علم‌سنجی تشریح می‌گردد. یافته‌ها به تحلیل روند تولید علم، همکاری بین دانشگاه‌ها، و ساختار مفهومی اختصاص دارد و در پایان، بحث، نتیجه‌گیری، محدودیت‌ها و پیشنهادها پژوهشی ارائه می‌شود.

۲. مبانی نظری پژوهش

مبانی نظری پژوهش حاضر بر دو مؤلفه مکمل استوار است. مؤلفه نخست، یادگیری ماشین به‌عنوان موضوع مورد مطالعه است که در این بخش به‌صورت مختصر مفاهیم بنیادین و تحول کلی آن تبیین می‌شود. مؤلفه دوم، علم‌سنجی به‌عنوان چارچوب تحلیلی پژوهش است که مبنای مطالعه ساختار مفهومی، روابط میان موضوعات و سیر تحول پژوهش‌های یادگیری ماشین را فراهم می‌کند.

۱.۲. یادگیری ماشین به‌عنوان موضوع پژوهش

یادگیری ماشین یکی از شاخه‌های اصلی هوش مصنوعی است که بر توسعه الگوریتم‌ها و مدل‌هایی تمرکز دارد که می‌توانند الگوها و روابط موجود در داده‌ها را شناسایی کرده و بر اساس آنها وظایفی مانند پیش‌بینی، دسته‌بندی و تصمیم‌گیری را انجام دهند (Abdullah et al., 2025; Alnuaimi & Albaldawi, 2024). این حوزه در پیوند با آمار، نظریه تصمیم، علوم رایانه و هوش

¹ . Three-Field Plot

مصنوعی شکل گرفته و توسعه آن به طور قابل توجهی تحت تأثیر افزایش حجم داده‌ها و توان پردازشی قرار گرفته است (Surur et al., 2025).

از نظر تاریخی، توسعه یادگیری ماشین را می‌توان حاصل تداوم پژوهش‌های مرتبط با محاسبه، هوش مصنوعی و یادگیری آماری دانست. پژوهش‌های اولیه درباره محاسبه و ماشین‌های هوشمند، از جمله آثار تورینگ، زمینه نظری شکل‌گیری هوش مصنوعی و یادگیری ماشین را فراهم کردند (Turing, 2009). در ادامه، توسعه روش‌های آماری و الگوریتم‌هایی نظیر ماشین بردار پشتیبان و درخت‌های تصمیم، دامنه کاربردهای یادگیری ماشین را گسترش داد (Ezugwu et al., 2024). در قرن بیست و یکم نیز ظهور یادگیری عمیق و شبکه‌های عصبی چندلایه، امکان پردازش داده‌های پیچیده و توسعه کاربردهای گسترده‌تر این فناوری را فراهم کرد (Elshenawy et al., 2025).

از منظر پارادایمی، یادگیری ماشین معمولاً در سه گروه اصلی یادگیری نظارت‌شده، یادگیری بدون نظارت و یادگیری تقویتی طبقه‌بندی می‌شود (Patel & Patel, 2025). یادگیری نظارت‌شده بر استفاده از داده‌های دارای برچسب برای یادگیری رابطه میان ورودی و خروجی، یادگیری بدون نظارت بر کشف الگوها و ساختارهای پنهان در داده‌ها، و یادگیری تقویتی بر یادگیری از طریق تعامل عامل با محیط و دریافت بازخورد استوار است. گستردگی این رویکردها سبب شده است یادگیری ماشین در حوزه‌های متنوعی از جمله پزشکی، مهندسی، اقتصاد، علوم اجتماعی و محیط‌زیست به کار گرفته شود (Singh, 2025).

با این حال، در پژوهش حاضر یادگیری ماشین از منظر توسعه یا مقایسه الگوریتم‌ها بررسی نمی‌شود؛ بلکه به عنوان یک حوزه دانشی و پژوهشی مورد مطالعه قرار می‌گیرد. بنابراین، سؤال اصلی این پژوهش آن است که ساختار مفهومی و سیر تحول موضوعی تولیدات علمی فارسی این حوزه در طول دوره مورد بررسی چگونه بوده است؟ پاسخ به این سؤال مستلزم استفاده از رویکردهای علم‌سنجی است.

۲.۲. علم‌سنجی

علم‌سنجی رویکردی کمی برای مطالعه تولید، ارتباط، ساختار و تحول دانش علمی است. در این رویکرد، عناصر قابل مشاهده نظام ارتباطات علمی، مانند مقالات، استنادها، نویسندگان، سازمان‌ها و کلیدواژه‌ها، به عنوان شواهدی برای مطالعه ساختار اجتماعی، فکری و مفهومی علم مورد استفاده

قرار می‌گیرند (Mingers & Leydesdorff, 2015). از این منظر، علم‌سنجی تنها به شمارش تولیدات علمی محدود نیست، بلکه می‌تواند روابط میان عناصر مختلف یک حوزه و تغییر این روابط در طول زمان را نیز آشکار کند.

یکی از حوزه‌های مهم علم‌سنجی، نگاشت علمی^۱ است. هدف نگاشت علمی، بازنمایی ساختار و روابط میان واحدهای مختلف یک حوزه پژوهشی و شناسایی سازمان مفهومی، فکری و اجتماعی آن است (Cobo et al., 2011). در این رویکرد، عناصر علمی به‌صورت شبکه‌ای از گره‌ها و روابط میان آنها نمایش داده می‌شوند و ساختار حاصل می‌تواند برای شناسایی خوشه‌ها، محورهای اصلی و روابط میان موضوعات مورد استفاده قرار گیرد.

نگاشت علمی را می‌توان بر اساس نوع رابطه مورد مطالعه به رویکردهای مختلفی تقسیم کرد. برای مثال، تحلیل هم‌تألفی برای مطالعه روابط و همکاری‌های علمی میان نویسندگان یا سازمان‌ها، تحلیل هم‌استنادی برای شناسایی پیوندهای فکری و آثار علمی مرتبط، و تحلیل هم‌رخدادی کلیدواژه‌ها برای مطالعه ساختار مفهومی و موضوعی یک حوزه به کار می‌روند (Cobo et al., 2011; Khasseh et al., 2017). بنابراین، انتخاب نوع تحلیل باید متناسب با پرسش پژوهش باشد.

در پژوهش حاضر، از آنجا که تمرکز اصلی بر ساختار مفهومی و سیر تحول موضوعی پژوهش‌های فارسی حوزه یادگیری ماشین است، تحلیل هم‌رخدادی کلیدواژه‌ها و روش‌های مرتبط با تحلیل موضوعی اهمیت ویژه‌ای دارند. در این چارچوب، کلیدواژه‌ها به‌عنوان بازنمایی‌های فشرده‌ای از مفاهیم مورد توجه پژوهشگران در نظر گرفته می‌شوند و روابط هم‌رخدادی آنها می‌تواند تصویری از پیوند میان موضوعات پژوهشی ارائه کند.

۳.۲. تحلیل هم‌رخدادی کلیدواژه‌ها و ساختار مفهومی

تحلیل هم‌رخدادی کلیدواژه‌ها یکی از روش‌های اصلی برای مطالعه ساختار مفهومی حوزه‌های علمی است. در این روش که ارتباط میان واژه‌های به‌کاررفته در قسمت‌های مختلف مدارک (از جمله عنوان، چکیده، کلیدواژه‌ها و مانند آن) مورد بررسی قرار می‌گیرد، فرض بر آن است که پربسامدترین واژه‌ها در مقایسه با واژه‌های کم‌بسامد، تأثیر بیشتری بر یک حوزه داشته‌اند. به طور کلی، تحلیل هم‌واژگانی این امکان را فراروی پژوهشگران قرار می‌دهد تا خوشه‌های موضوعی در

¹ Science Mapping

یک مجله یا حوزه خاص را شناسایی کنند (جواهری و همکاران، ۱۴۰۰؛ خاصه و همکاران، ۱۴۰۰؛ فرهنگی و همکاران، ۱۳۹۷).

در شبکه هم‌رخدادی، کلیدواژه‌ها به‌عنوان گره و روابط هم‌رخدادی میان آنها به‌عنوان یال نمایش داده می‌شوند. میزان تکرار هم‌زمان کلیدواژه‌ها می‌تواند قدرت ارتباط میان آنها را نشان دهد. تحلیل این شبکه امکان شناسایی مجموعه‌هایی از کلیدواژه‌های مرتبط را فراهم می‌کند که معمولاً به‌صورت خوشه‌های موضوعی ظاهر می‌شوند. هر خوشه می‌تواند بازتاب‌دهنده یک زیرحوزه، مسئله، روش یا زمینه کاربردی در ساختار دانش مورد مطالعه باشد (Callon et al., 1991; Cobo et al., 2011).

اهمیت این رویکرد برای پژوهش حاضر در آن است که ساختار مفهومی حوزه یادگیری ماشین را می‌توان از طریق روابط میان کلیدواژه‌ها مطالعه کرد، بدون آنکه صرفاً بر فراوانی منفرد هر واژه تکیه شود. به عبارت دیگر، فراوانی کلیدواژه نشان‌دهنده میزان حضور یک مفهوم است، در حالی که الگوی هم‌رخدادی نشان می‌دهد این مفهوم در چه ارتباطی با سایر مفاهیم قرار گرفته است. از این رو، تحلیل هم‌رخدادی می‌تواند تصویری شبکه‌ای از سازمان مفهومی پژوهش‌های این حوزه فراهم کند.

۴.۲. تکامل موضوعی و پویایی ساختار دانش

ساختار مفهومی یک حوزه علمی ثابت نیست و در طول زمان می‌تواند دستخوش ظهور موضوعات جدید، تداوم موضوعات موجود، تغییر ارتباط میان موضوعات و افول برخی مضامین شود. از این رو، مطالعه ساختار موضوعی در یک مقطع زمانی به‌تنهایی برای شناخت سیر تحول یک حوزه کافی نیست و باید بعد زمانی نیز در تحلیل وارد شود.

تحلیل تکامل موضوعی با تقسیم دوره مورد مطالعه به بازه‌های زمانی و بررسی تغییر، تداوم و ظهور مضامین در این بازه‌ها، امکان مطالعه پویایی ساختار مفهومی یک حوزه را فراهم می‌کند. این تحلیل می‌تواند نشان دهد که چه موضوعاتی در طول زمان تداوم یافته‌اند، چه مضامینی در دوره‌های بعدی ظاهر شده‌اند و چه تغییراتی در ارتباط میان مضامین رخ داده است (Cobo et al., 2011).

بنابراین، در پژوهش حاضر، تحلیل تکامل موضوعی مکمل تحلیل هم‌رخدادی است. تحلیل هم‌رخدادی، ساختار مفهومی و روابط میان موضوعات را نشان می‌دهد، در حالی که تحلیل

تکامل موضوعی، مسیر تغییر این ساختار در طول زمان را آشکار می‌کند. این تمایز برای تفسیر نتایج اهمیت دارد؛ زیرا افزایش فراوانی یک کلیدواژه الزاماً به معنای تحول ساختاری حوزه نیست و تحول موضوعی باید بر اساس تغییر در روابط و تداوم یا ظهور مضامین در بازه‌های زمانی تفسیر شود.

۵.۲. نمودار راهبردی و جایگاه مضامین پژوهشی

برای ارزیابی جایگاه نسبی خوشه‌های موضوعی در ساختار یک حوزه، نمودار راهبردی^۱ یکی از چارچوب‌های شناخته‌شده در تحلیل علم‌سنجی است. این چارچوب که در مطالعات تحلیل هم‌واژه توسعه یافته است، جایگاه هر خوشه را بر اساس دو بعد اصلی، یعنی مرکزیت^۲ و تراکم^۳، نمایش می‌دهد (Callon et al., 1991; Cobo et al., 2011).

مرکزیت بیانگر میزان ارتباط یک خوشه با سایر بخش‌های ساختار موضوعی است و بنابراین می‌تواند نشان دهد که یک مضمون تا چه اندازه برای ساختار کلی حوزه اهمیت دارد. در مقابل، تراکم بیانگر میزان انسجام و توسعه‌یافتگی روابط درونی یک خوشه است و نشان می‌دهد که اعضای آن مضمون تا چه اندازه با یکدیگر ارتباط و انسجام موضوعی دارند (رستمی و همکاران، ۱۳۹۹؛ رحیمی و همکاران، ۱۴۰۳). ترکیب این دو شاخص، امکان تفکیک مضامین به چهار ناحیه اصلی را فراهم می‌کند: مضامین محرک^۴ که از مرکزیت و تراکم بالایی برخوردارند و هم در ساختار کلی حوزه اهمیت دارند و هم از توسعه‌یافتگی درونی مناسبی برخوردارند؛ مضامین بنیادی^۵ یا فراگیر که مرکزیت بالا و تراکم پایین دارند و با بخش‌های مختلف حوزه ارتباط دارند، اما روابط درونی آنها هنوز به‌طور کامل توسعه نیافته است؛ مضامین تخصصی^۶ که دارای تراکم بالا و مرکزیت پایین هستند و از انسجام درونی برخوردارند، اما ارتباط محدودی با سایر بخش‌های حوزه دارند؛ و مضامین نوظهور یا رو به افول^۷ که در ناحیه دارای مرکزیت و تراکم پایین قرار می‌گیرند و می‌توانند بیانگر موضوعات تازه‌شکل گرفته یا موضوعاتی باشند که اهمیت

^۱ Strategic Diagram

^۲ Centrality

^۳ Density

^۴ Motor Themes

^۵ Basic Themes

^۶ Niche Themes

^۷ Emerging or Declining Themes

آنها در حال کاهش است. با این حال، تشخیص قطعی نوظهور بودن یا رو به افول بودن یک مضمون صرفاً بر اساس یک نمودار مقطعی امکان‌پذیر نیست و برای این منظور باید شواهد زمانی نیز مورد توجه قرار گیرد (Cobo et al., 2011).

از این منظر، نمودار راهبردی صرفاً ابزاری برای رتبه‌بندی موضوعات نیست، بلکه چارچوبی برای تفسیر جایگاه ساختاری و راهبردی مضامین در شبکه موضوعی فراهم می‌کند. بنابراین، فراوانی بالای یک کلیدواژه به تنهایی به معنای قرار گرفتن آن در زمره مضامین محرک نیست؛ زیرا فراوانی، اطلاعات مستقیمی درباره میزان مرکزیت یا تراکم یک خوشه ارائه نمی‌دهد و تفسیر جایگاه هر مضمون باید بر اساس موقعیت آن در شبکه و شاخص‌های ساختاری مربوط انجام شود (Callon et al., 1991; Cobo et al., 2011).

۳. پیشینه پژوهش

در ایران پژوهش‌های متعددی با رویکرد علم‌سنجی به بررسی وضعیت هوش مصنوعی و زیرحوزه‌های یادگیری ماشین پرداخته‌اند. فرزین یزدی و رضایی شریف‌آبادی (۱۳۹۶) در یکی از نخستین مطالعات علم‌سنجی هوش مصنوعی در خاورمیانه، ایران را از نظر حجم تولیدات علمی در جایگاه نخست منطقه معرفی کردند، هرچند شکاف قابل‌توجهی در شاخص‌های استنادی و همکاری‌های بین‌المللی نسبت به کشورهای پیشرو مشاهده شد. رضایی‌زاده و عسکری (۱۴۰۰) نیز با تحلیل بیش از ۴۴ هزار مقاله نمایه‌شده در وب‌آف‌ساینس، رشد سریع جهانی هوش مصنوعی و جایگاه محوری یادگیری ماشین را در این حوزه تأیید کردند. پس از آن، ظهوریان نادعلی و همکاران (۱۴۰۱) الگوهای همکاری دانشگاه-صنعت-دولت در پژوهش‌های هوش مصنوعی ایران را با کشورهای چین و ایالات متحده مقایسه کردند و نشان دادند که با وجود جایگاه مناسب ایران از نظر حجم تولیدات علمی، همکاری صنعت و دانشگاه همچنان محدود است. همچنین دانشگاه تهران، دانشگاه صنعتی امیرکبیر و دانشگاه صنعتی شریف به‌عنوان فعال‌ترین نهادهای علمی کشور معرفی شدند.

در ادامه، کاشانی و داستانی (۱۴۰۳) نشان دادند که انتشارات هوش مصنوعی پزشکی ایران از سال ۲۰۱۵ رشد قابل توجهی داشته و «یادگیری ماشین» و «یادگیری عمیق» از مهم‌ترین موضوعات بنیادین این حوزه محسوب می‌شوند، در حالی که مفاهیمی نظیر «پزشکی

شخصی‌سازی‌شده» و «بهینه‌سازی سراسری» به‌عنوان موضوعات نوظهور مطرح شده‌اند. توسن و همکاران (۱۴۰۴) نیز با تحلیل پژوهش‌های سنجش از دور و یادگیری ماشین، موضوعات نوظهوری همچون ادغام داده‌های چندمنبعی، جنگل تصادفی و شبکه‌های عصبی مصنوعی را شناسایی کردند. ابراهیمی ترکمانی و محمودی (۱۴۰۴) نیز نشان دادند که پژوهشگران دانشگاه‌های تهران بیشترین تمرکز خود را بر موضوعاتی مانند یادگیری ماشین، شبکه‌های عصبی مصنوعی و پردازش زبان طبیعی معطوف کرده‌اند. سیاحی و همکاران (۱۴۰۵) نیز با مطالعه علم‌سنجی کاربردهای یادگیری ماشین و یادگیری عمیق در تشخیص بیماری کبد چرب غیرالکلی، روند رو به رشد تولیدات علمی و اهمیت همکاری‌های بین‌المللی را نشان دادند. همچنین، آقایی و کریم‌میان (۲۰۲۵) پیوند مفهومی میان حاکمیت هوش مصنوعی و کاربرد هوش مصنوعی در حکمرانی را بررسی کرده و اخلاق، حکمرانی داده و تنظیم‌گری مسئولانه را به‌عنوان مهم‌ترین مضامین پیونددهنده معرفی کردند.

فراتر از مطالعات داخلی، پژوهش‌های بین‌المللی نیز از روش‌های علم‌سنجی برای ترسیم ساختار دانش، الگوهای همکاری علمی و سیر تحول موضوعی پژوهش‌های یادگیری ماشین در بافت‌های ملی یا حوزه‌های تخصصی استفاده کرده‌اند. هیگاکي و همکاران (۲۰۲۰) با تحلیل شبکه هم‌تألفی ۲۸۵۷ مقاله پژوهشی در حوزه قلب و عروق که از یادگیری ماشین استفاده کرده بودند (۲۰۰۹-۲۰۱۹)، ساختار شبکه‌ای مقیاس‌آزاد را آشکار ساختند و نشان دادند که در میان ۲۲ نویسنده‌ای که در حداقل یکی از شاخص‌های مرکزیت در میان ۱۰ نفر برتر قرار داشتند، ۱۴ نویسنده (۶۳٫۶ درصد) دارای مدرک پزشکی بوده و سایر پژوهشگران دارای پیشینه علوم محاسباتی بودند؛ به‌گونه‌ای که همکاری‌های میان‌رشته‌ای میان متخصصان علوم پزشکی و علوم داده نقش مهمی در توسعه این حوزه ایفا می‌کند (Higaki et al., 2020). چن و همکاران (۲۰۲۱) نیز با هدف ترسیم چشم‌انداز پژوهش‌های یادگیری ماشین مبتنی بر رمزنگاری هم‌ریخت، به تحلیل علم‌سنجی تولیدات علمی این حوزه پرداختند و نشان دادند که مفاهیمی نظیر رایانش ابری، شبکه‌های عصبی و داده‌های کلان به‌عنوان موضوعات غالب ظاهر شده‌اند (Chen et al., 2021). ازوگو و همکاران (۲۰۲۳) با تحلیل ۲۷۶۱ مقاله مرتبط با یادگیری ماشین از ۵۴ کشور آفریقایی در بازه ۳۰ ساله، مروری علم‌سنجی انتقادی ارائه دادند که چشم‌اندازهای فعلی، روندهای آینده و شبکه‌های همکاری علمی در سراسر قاره را ترسیم کرد (Ezugwu et al., 2023). لاندبرگ و

همکاران (۲۰۲۴) نیز با تحلیل ۳۹۸,۷۸۲ سند علمی نمایه شده در اسکوپوس طی سال‌های ۲۰۱۳ تا ۲۰۲۲ نشان دادند که یادگیری ماشین به یکی از سریع‌ترین حوزه‌های در حال رشد نظام علمی جهان تبدیل شده است و بر گسترش همکاری‌های بین‌المللی و افزایش سهم موضوعات مرتبط با یادگیری عمیق و هوش مصنوعی مولد تأکید کردند (Lundberg et al., 2024). همچنین، فانگ و همکاران (۲۰۲۵) با تحلیل علم‌سنجی ۱۰۳۶ مقاله مرتبط با کاربرد یادگیری ماشین در انفارکتوس میوکارد، رشد شتابان تولیدات علمی از سال ۲۰۱۶ و ظهور موضوعاتی مانند یادگیری عمیق، شبکه‌های عصبی و ارزیابی ریسک را به‌عنوان مهم‌ترین محورهای پژوهشی نوظهور گزارش کردند (Fang et al., 2025). این مطالعات به‌طور جمعی نشان می‌دهند که تحلیل‌های علم‌سنجی پژوهش‌های یادگیری ماشین قادرند ساختارهای همکاری علمی و سیر تحول موضوعی این حوزه را به‌طور مؤثر آشکار سازند.

در کنار مطالعاتی که یادگیری ماشین را موضوع تحلیل‌های علم‌سنجی قرار داده‌اند، دسته دیگری از پژوهش‌ها از الگوریتم‌های یادگیری ماشین به‌عنوان ابزاری برای ارتقای تحلیل‌های علم‌سنجی استفاده کرده‌اند. دِلیچه و پولیمینی (۲۰۲۰) از روش‌شناسی علم‌سنجی مبتنی بر یادگیری ماشین برای تحلیل روندهای پژوهش کووید-۱۹ استفاده کردند و ضمن شناسایی رشد نمایی انتشارات، شبکه‌های همکاری علمی میان کشورها را ترسیم نمودند (De Felice & Polimeni, 2020). کانتوروویچ-رزنیچنکو و کانتوروویچ (۲۰۲۵) نیز با بهره‌گیری هم‌زمان از ابزارهای علم‌سنجی و یادگیری ماشین، تحول ۶۰ ساله پژوهش‌های حوزه حقوق و اقتصاد را ترسیم کرده و تغییر جهت پژوهش‌ها از مطالعات نظری به پژوهش‌های تجربی و شواهدمحور را نشان دادند (Kantorowicz-Reznichenko & Kantorowicz, 2025). این مطالعات نشان می‌دهند که یادگیری ماشین علاوه بر آنکه خود موضوع تحلیل‌های علم‌سنجی است، می‌تواند به‌عنوان ابزاری کارآمد برای توسعه روش‌های تحلیل علم‌سنجی نیز به کار گرفته شود.

در مجموع، اگرچه مطالعات پیشین بینش‌های ارزشمندی درباره زیردامنه‌های مختلف هوش مصنوعی ارائه کرده‌اند، ادبیات موجود همچنان از فقدان یک نگاهت جامع از ساختار دانش و تولیدات علمی فارسی‌زبان در حوزه یادگیری ماشین رنج می‌برد. افزون بر این، پژوهش‌های پیشین به‌صورت یکپارچه روندهای انتشار، توزیع مجلات، شبکه‌های همکاری دانشگاهی، هم‌رخدادی کلیدواژه‌ها و سیر تکامل موضوعی پژوهش‌های فارسی نمایه شده در پایگاه

آی‌اس‌سی را بررسی نکرده‌اند. از این رو، انجام مطالعه حاضر می‌تواند با ارائه تصویری جامع از ساختار مفهومی و سیر تحول این حوزه، شواهد لازم را برای سیاست‌گذاری پژوهشی و برنامه‌ریزی علمی در سطح ملی فراهم آورد.

۴. روش‌شناسی پژوهش

پژوهش حاضر از نوع مطالعات علم‌سنجی و با رویکرد توصیفی-تحلیلی است که با هدف ترسیم وضعیت تولیدات علمی فارسی حوزه یادگیری ماشین و تحلیل ساختار مفهومی این حوزه انجام شده است. برای دستیابی به این هدف، از شاخص‌های علم‌سنجی و روش‌های تحلیل شبکه‌های علمی شامل تحلیل هم‌رخدادی در سطح دانشگاه‌ها، تحلیل هم‌رخدادی کلیدواژه‌ها و نقشه‌برداری علمی استفاده شد.

منبع داده‌های پژوهش، نمایه استنادی علوم ایران در مؤسسه استنادی و پایش علم و فناوری جهان اسلام (آی‌اس‌سی) است. بازیابی اطلاعات در اردیبهشت ۱۴۰۴ و در بخش عنوان مدارک پایگاه آی‌اس‌سی انجام شد. عبارت جست‌وجو شامل «یادگیری ماشین» بود و کلیه مقالات پژوهشی فارسی نمایه‌شده در بازه زمانی ۱۳۹۲ تا ۱۴۰۳ بازیابی شدند. سایر انواع مدارک از جمله مقالات کنفرانسی، یادداشت‌ها و سایر مستندات غیرپژوهشی در مجموعه داده وارد نشدند. حاصل جست‌جو، بازیابی ۴۰۳ رکورد علمی بود. پس از بررسی محتوایی مدارک بر اساس معیارهای ورود و خروج پژوهش، چهار رکورد به دلیل عدم ارتباط موضوعی با یادگیری ماشین حذف شدند. معیار ورود شامل مقالات پژوهشی فارسی با تمرکز اصلی بر یادگیری ماشین بود. مدارکی که صرفاً به هوش مصنوعی، داده‌کاوی یا سایر حوزه‌های مرتبط بدون تمرکز بر یادگیری ماشین پرداخته بودند، حذف شدند. در نهایت ۳۹۹ مقاله به عنوان مجموعه نهایی داده‌ها برای تحلیل انتخاب گردید.

به منظور افزایش اعتبار تحلیل‌ها، پیش از انجام پردازش‌های علم‌سنجی، داده‌ها تحت فرایند پالایش و استانداردسازی قرار گرفتند. یکی از مهم‌ترین مراحل این فرایند، یکدست‌سازی کلیدواژه‌ها بود. از آنجا که نویسندگان از معادل‌های مختلف فارسی و انگلیسی، صورت‌های متفاوت نگارشی و گاه اصطلاحات مترادف برای اشاره به یک مفهوم استفاده کرده بودند، تمامی کلیدواژه‌ها به صورت دستی بازبینی شدند. در این مرحله، مترادف‌ها، شکل‌های مختلف نوشتاری، خطاهای تایپی و تفاوت‌های ناشی از فاصله و نیم‌فاصله اصلاح و در قالب واژگان استاندارد ادغام

شدند. برای نمونه، اصطلاحات «ماشین بردار پشتیبان» و «SVM»، «اینترنت اشیا»، «اینترنت اشیاء» و «IoT»، «یادگیری جمعی» و «یادگیری گروهی»، و نیز «یادگیری نظارت شده» و «یادگیری بانظارت» به عنوان مفاهیم یکسان در نظر گرفته شدند.

فرایند مشابهی برای استانداردسازی وابستگی‌های سازمانی نیز انجام شد. به دلیل وجود تفاوت‌های نگارشی در نام دانشگاه‌ها و مؤسسات پژوهشی، اشکال مختلف نام یک سازمان در قالب یک عنوان واحد جمع‌گردد. این اقدام به‌ویژه در تحلیل شبکه همکاری دانشگاه‌ها اهمیت داشت و از ایجاد گره‌های تکراری ناشی از تفاوت‌های نوشتاری جلوگیری کرد. برای مثال، اشکال مختلف نام دانشگاه آزاد اسلامی، دانشگاه صنعتی امیرکبیر، دانشگاه صنعتی خواجه نصیرالدین طوسی و دانشگاه علم و صنعت ایران در قالب نام‌های استاندارد یکپارچه شدند. علاوه بر این، نام نویسندگان نیز از نظر اختلاف‌های نگارشی، فاصله و نیم‌فاصله، ترتیب نام و نام خانوادگی و اشکال مختلف ثبت اسامی بازمینی و یکدست‌سازی شد تا از ایجاد گره‌های تکراری در تحلیل‌های شبکه جلوگیری شود.

پس از آماده‌سازی داده‌ها، اطلاعات در محیط نرم‌افزار بیبلوشاینی^۱ و ووز-ویور^۲ تحلیل شد. بیبلوشاینی رابط گرافیکی بسته بیبلیومتریکس در محیط آر^۳ است که امکان انجام طیف گسترده‌ای از تحلیل‌های علم‌سنجی و مصورسازی شبکه‌های علمی را فراهم می‌سازد (Aria & Cuccurullo, 2017). در این پژوهش از قابلیت‌های این نرم‌افزار برای تحلیل روند تولیدات علمی یادگیری ماشین، ترسیم سیر تحول و تکامل موضوعات پژوهشی، ترسیم شبکه‌های همکاری دانشگاه‌ها، و تحلیل ساختار مفهومی این حوزه استفاده شد. همچنین برای ترسیم و مصورسازی شبکه همکاری دانشگاه‌ها از نرم‌افزار ووز-ویور استفاده شد. این نرم‌افزار امکان نمایش گرافیکی روابط همکاری میان دانشگاه‌ها را بر اساس هم‌نویسندگی و وابستگی سازمانی فراهم می‌سازد (van Eck & Waltman, 2010). بدین منظور، تعداد ۴۱ دانشگاه که دارای حداقل ۳ مقاله بودند وارد تحلیل نهایی شدند و با استفاده از روش نرمال سازی Association شبکه همکاری ترسیم شد. همچنین در بخش تحلیل هم‌واژگانی، پس از بررسی فراوانی کلیدواژه‌ها، تعداد ۹۶ کلیدواژه برتر با حداقل فراوانی ۳ در تحلیل نهایی وارد شدند تا میزان هم‌رخدادی آنان در مقاله‌ها به دست

¹ Biblioshiny

² VOSviewer

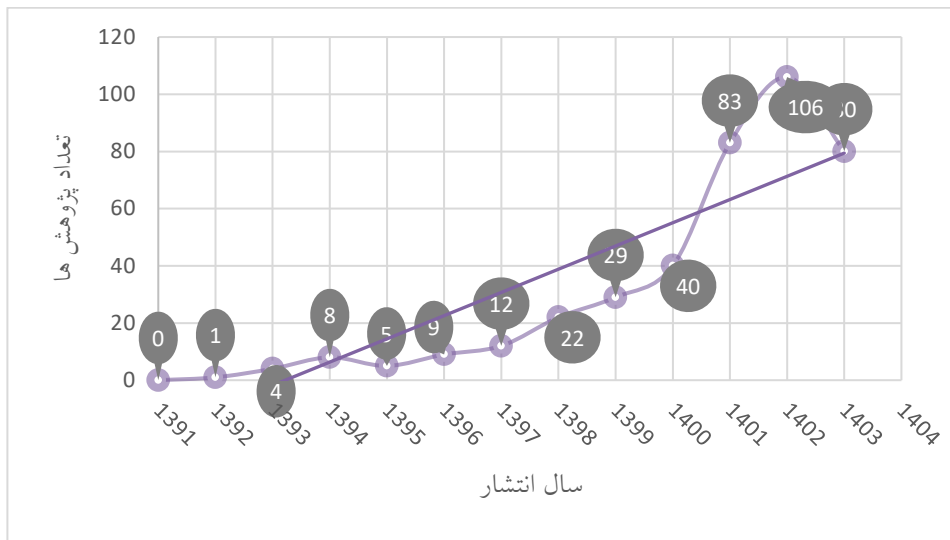
³ R

آید و سپس شبکه هم واژگانی و نمودار راهبردی حوزه داده کاوی با استفاده از بیبلیوشاینی ترسیم شد. در این بخش نیز شبکه هم واژگانی با روش نرمال سازی Association و الگوریتم خوشه بندی Louvain ترسیم شد. با توجه به اینکه تمامی مدارک بازیابی شده به حوزه یادگیری ماشین اختصاص داشتند، کلیدواژه «یادگیری ماشین» از تحلیل‌های متاثر از کلیدواژه ها حذف شد. حضور این واژه در بخش عمده‌ای از مدارک می‌توانست موجب غلبه یک مفهوم عام بر شبکه هم‌رخدادی و کاهش قابلیت شناسایی روابط میان موضوعات تخصصی تر شود. ازاین‌رو، حذف آن امکان آشکارسازی دقیق‌تر ساختار دانشی و مضامین پژوهشی حوزه را فراهم کرد.

۵. یافته‌ها

سوال اول پژوهش: روند انتشار مقالات فارسی در حوزه یادگیری ماشین چگونه است؟

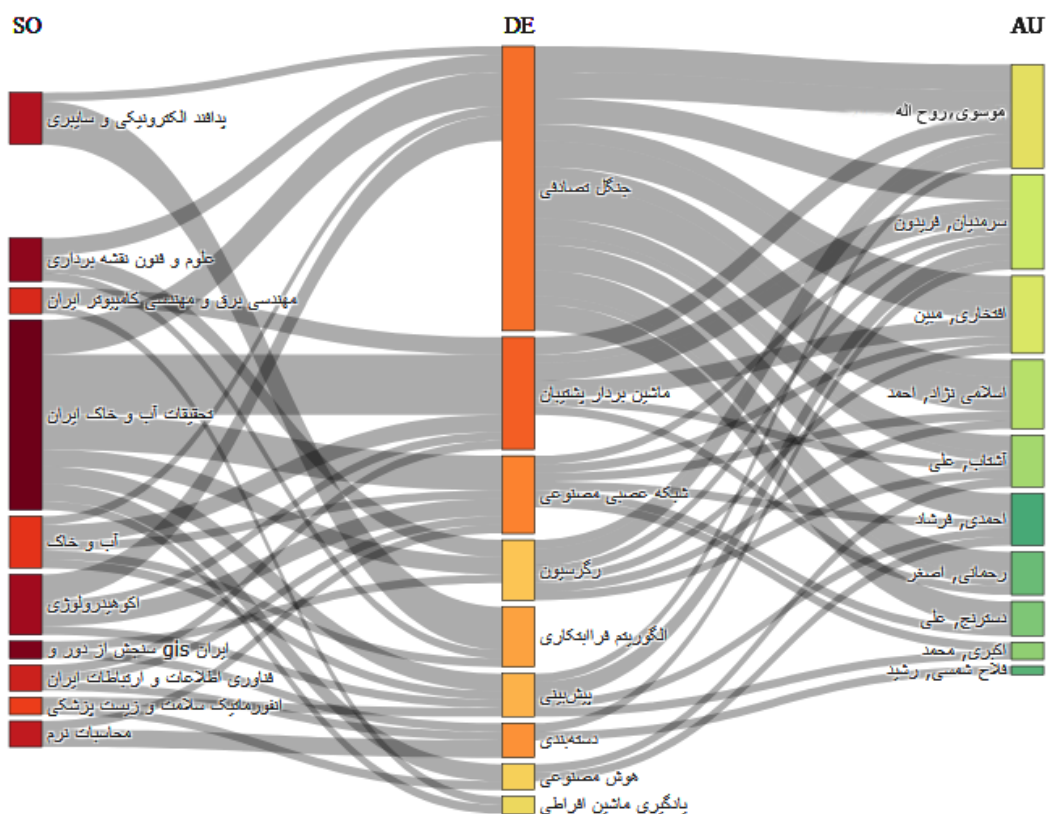
همان‌طور که پیشتر اشاره شد بین سال‌های ۱۳۹۲ تا ۱۴۰۳، در مجموع ۳۹۹ مقاله فارسی در پایگاه آی‌اس‌سی نمایه شده است. تجزیه و تحلیل داده‌ها، افزایش قابل توجهی در پژوهش‌های فارسی یادگیری ماشین را نشان می‌دهد (شکل ۱). نرخ انتشار از یک مقاله در سال ۱۳۹۲ به اوج ۱۰۶ مقاله در سال ۱۴۰۲ رسیده که نشان‌دهنده افزایش چشمگیر در این بازه زمانی است. بر این اساس، طول دوره را می‌توان به سه دوره آغازین، شتاب، و جهش دسته‌بندی کرد که در مرحله آغازین (۱۳۹۲-۱۳۹۶)، با رشد آهسته و خروجی‌های سالانه کمتر از ۱۰ مقاله باقی مانده است. در مرحله شتاب (۱۳۹۷-۱۴۰۰)، روند صعودی قابل توجهی در سال ۱۳۹۷ به خود گرفته است، به طوری که تعداد انتشارات سالانه از ۲۲ مقاله در سال ۱۳۹۸ به ۴۰ در سال ۱۴۰۰ افزایش یافته است. و در مرحله جهش (۱۴۰۱-۱۴۰۳)، سریع‌ترین گسترش در این حوزه را نشان می‌دهد که در سال ۱۴۰۲ به ۱۰۶ مقاله می‌رسد. البته تعداد ۸۰ مقاله در سال ۱۴۰۳ می‌تواند متاثر از تاخیر پایگاه آی‌اس‌سی در نمایه سازی مقالات باشد.



شکل ۱: روند انتشار مقالات فارسی در حوزه یادگیری ماشین (۱۳۹۱-۱۴۰۳)

سوال دوم پژوهش: بر اساس نمودار سه بخشی، روابط میان نویسندگان، مجلات و کلیدواژه‌ها در پژوهش‌های حوزه یادگیری ماشین چگونه است؟

برای پاسخ به این سؤال، از نمودار سه بخشی استفاده شد (شکل ۲). این نمودار ساختار دانش فارسی در حوزه یادگیری ماشین را در سه سطح نویسندگان، مجلات و کلیدواژه‌ها نشان می‌دهد.



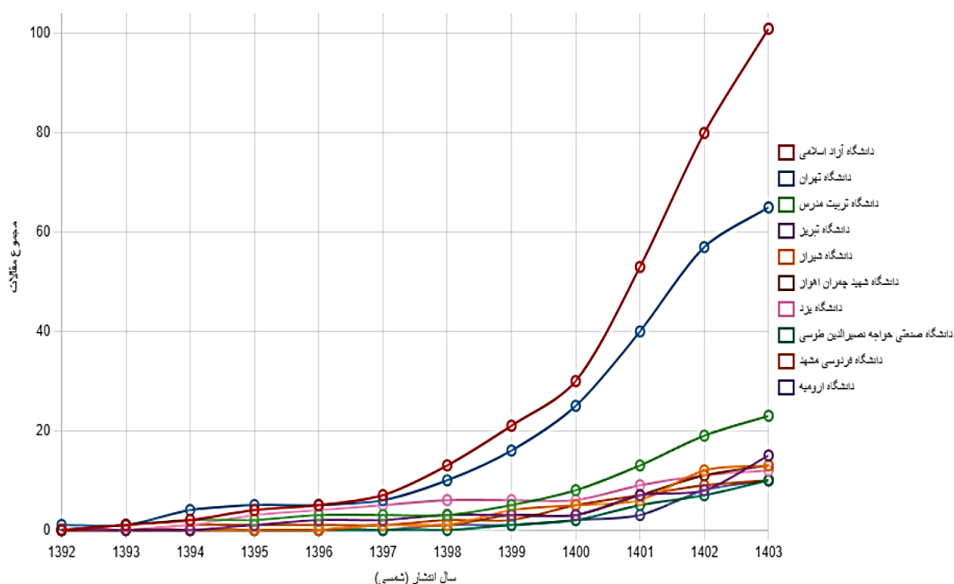
شکل ۲: نمودار سه‌بخشی، ارتباط میان نویسندگان، مجلات و کلیدواژه‌های حوزه یادگیری ماشین

همانطور که در شکل ۲ مشخص است، دو کلیدواژه جنگل تصادفی و ماشین بردار پشتیبان به عنوان گره‌های مرکزی عمل می‌کنند. ضخامت بالای نوارهای جریان منتهی به این دو موضوع، نشان‌دهنده استفاده گسترده و فراگیر از این دو الگوریتم در مقالات فارسی است. همچنین، مجله‌ی تحقیقات آب و خاک ایران بیشترین سهم را در پذیرش مقالات با کلیدواژه‌های اصلی دارد. جریان‌های ورودی این مجله نشان می‌دهد که این نشریه، مقالات کاربردی متعددی را در زمینه‌های جنگل تصادفی و شبکه‌های عصبی مصنوعی جذب کرده است. از جانب دیگر، ساختار جریان‌ها در سمت راست نشان می‌دهد که نویسندگان برتر این حوزه، همگرایی بالایی با روش‌های ماشین بردار پشتیبان و جنگل تصادفی دارند. همچنین، مشاهده می‌شود که برخی نویسندگان جریان‌های بسیار گسترده‌ای به سمت کلیدواژه‌های متنوع دارند که می‌تواند بیانگر

فعالیت چندگانه و تسلط آن‌ها بر طیف وسیعی از الگوریتم‌ها باشد. با توجه به اطلاعات ارائه شده در نمودار سه‌بخشی، می‌توان مشاهده کرد که بخشی از مقالات منتشر شده در مجلات فعال، به کاربرد الگوریتم‌هایی مانند «جنگل تصادفی» و «ماشین بردار پشتیبان» در موضوعاتی از جمله محیط‌زیست و نقشه‌برداری اختصاص یافته‌اند.

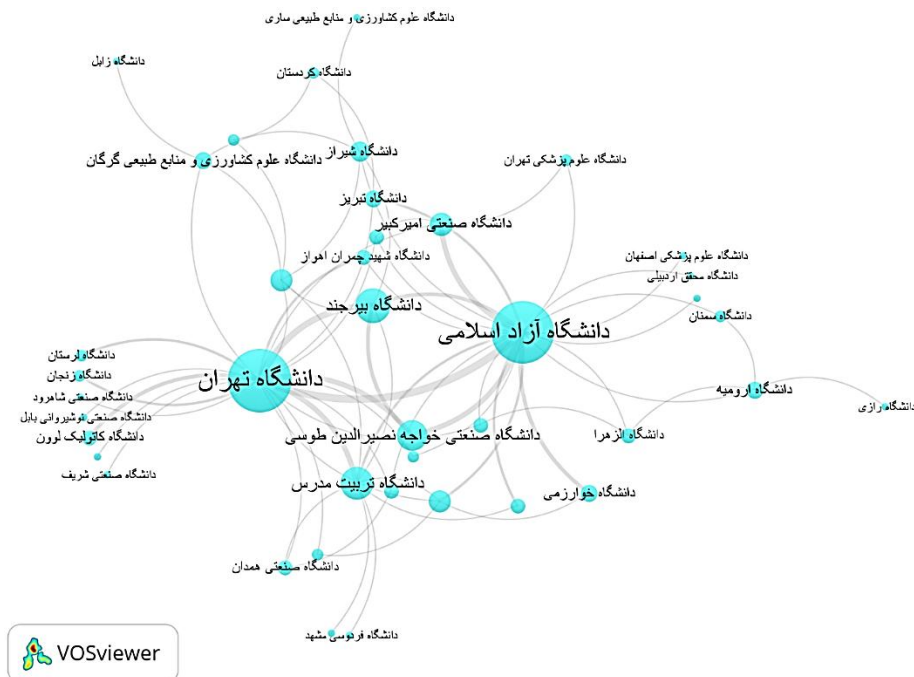
سوال سوم پژوهش: فعال‌ترین دانشگاه‌ها در تولیدات علمی حوزه یادگیری ماشین کدام‌اند و شبکه همکاری میان آن‌ها چه ساختاری دارد؟

بر اساس تحلیل داده‌ها، در مجموع ۱۴۶ دانشگاه و مؤسسه‌ی آموزشی در بازه زمانی ۱۳۹۲ تا ۱۴۰۳ در حوزه یادگیری ماشین فعالیت پژوهشی داشته و اقدام به انتشار مقاله کرده‌اند. این حجم از مشارکت، نشان‌دهنده گسترده‌گی و نفوذ یادگیری ماشین در تمامی سطوح نظام آموزش عالی کشور است. از میان این ۱۴۶ مؤسسه، دانشگاه‌های آزاد اسلامی، تهران و تربیت مدرس به عنوان پرتولیدترین دانشگاه‌ها در این حوزه شناخته می‌شوند. دانشگاه آزاد اسلامی با ۱۰۱ مقاله، در صدر لیست قرار دارد. همان‌طور که در شکل ۳ مشاهده می‌شود، این دانشگاه از سال ۱۳۹۹ روند رشد پرشتابی را آغاز کرده و تا سال ۱۴۰۳ فاصله‌ی معناداری با سایر دانشگاه‌ها ایجاد کرده است. دانشگاه تهران با ۶۵ مقاله، جایگاه دوم را به خود اختصاص داده و همگام با دانشگاه آزاد اسلامی، نقش محوری در پیشبرد پژوهش‌های این حوزه ایفا کرده است و دانشگاه تربیت مدرس نیز با ۲۳ مقاله، در رتبه سوم جای دارد. شکل ۳ وضعیت ده دانشگاه فعال در پژوهش‌های حوزه یادگیری ماشین را در گذر زمان به تصویر می‌کشد.



شکل ۳: وضعیت ده دانشگاه فعال در پژوهش‌های حوزه یادگیری ماشین در گذر زمان

تحلیل شبکه‌ای همکاری علمی میان دانشگاه‌ها در حوزه یادگیری ماشین که در شکل ۴ نمایش داده شده است، بیانگر ساختار همکاری‌های بین‌دانشگاهی در ایران است. این شبکه که با نرم افزار ووز-ویور ترسیم شده است مشتمل بر ۴۱ گره (دانشگاه) است.

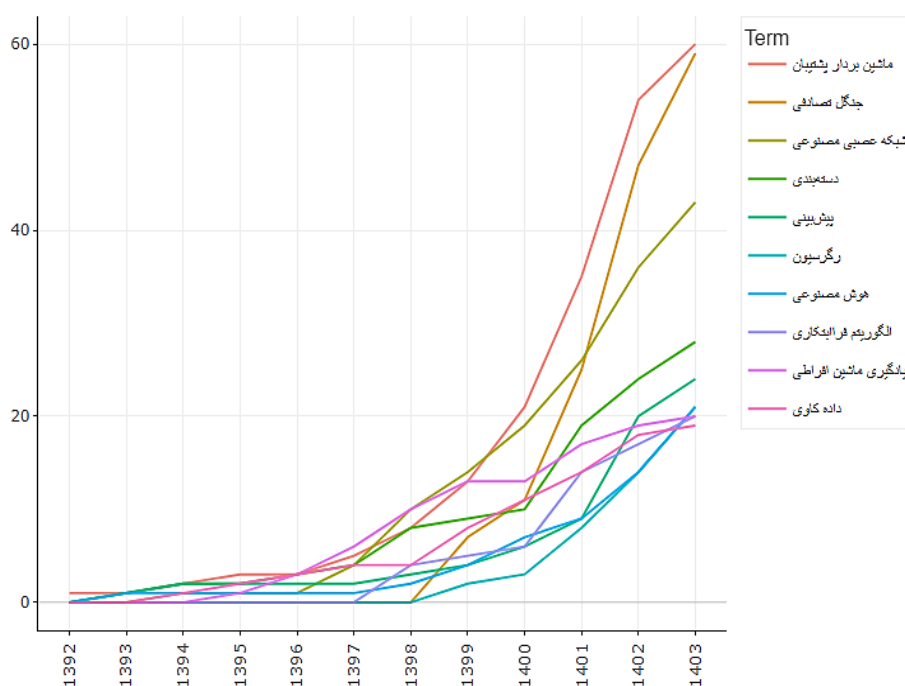


شکل ۴: شبکه همکاری علمی میان دانشگاه‌ها در حوزه یادگیری ماشین

یافته‌ها در این بخش حاکی از آن است که دانشگاه‌های «آزاد اسلامی» و «تهران» به عنوان بزرگ‌ترین گره‌های شبکه، بیشترین نقش را در برقراری پیوندهای علمی و همکاری با سایر دانشگاه‌ها ایفا می‌کنند. ضخامت خطوط ارتباطی متصل به این دو دانشگاه نشان‌دهنده تراکم همکاری‌های پژوهشی آن‌ها با طیف وسیعی از مؤسسات دیگر است. با توجه به مقادیر تراکم شبکه ۰.۰۸۶۶ و میانگین درجه ۳.۴۶، به نظر می‌رسد ساختار شبکه، ساختاری نسبتاً پراکنده با همکاری محدود و متمرکز بر تعداد کمی از دانشگاه‌هاست و هر دانشگاه با تعداد کمی از دانشگاه‌های دیگر ارتباط دارد. البته با توجه به مقدار میانگین درجه، این شبکه دارای چند هسته مرکزی است که همکاری بیشتری دارند و بقیه دانشگاه‌ها پیرامون آنها می‌چرخند.

سوال چهارم پژوهش: سیر تحول و تکامل موضوعات پژوهشی در حوزه یادگیری ماشین چگونه بوده است؟

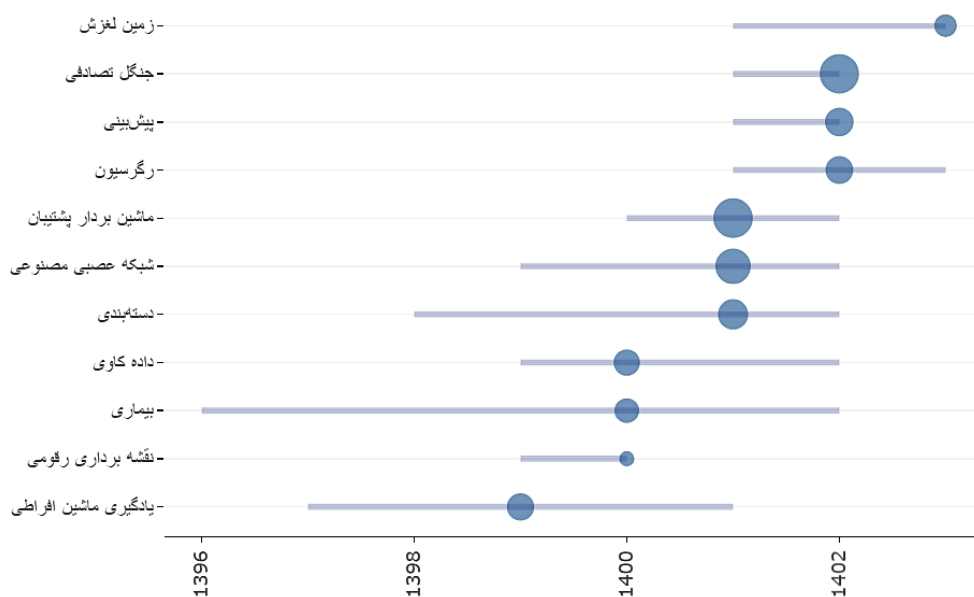
تجزیه و تحلیل داده‌ها نشان داد که با احتساب کلیدواژه «یادگیری ماشین»، مجموعاً ۱۸۷۳ کلیدواژه در پژوهش‌های این حوزه استفاده شده است (میانگین ۴۶۹ کلیدواژه به ازای هر مقاله). فراوانی تجمعی ۱۰ کلیدواژه برتر حوزه یادگیری ماشین در شکل ۵ نشان داده شده است. این شکل که فراز و فرود کلیدواژه‌ها را در بین سالهای ۱۳۹۲ تا ۱۴۰۳ آشکار می‌سازد حاوی اطلاعات ارزنده‌ای است که در ادامه بدان پرداخته شده است.



شکل ۵: کلیدواژه‌های برتر در مقالات یادگیری ماشین در گذر زمان

بررسی دقیق‌تر این نمودار حاکی از آن است که «ماشین بردار پشتیبان» با فراوانی تجمعی ۶۰، پرتکرارترین کلیدواژه در کل دوره است. رشد آن از حدود سال ۱۳۹۸ شتاب می‌گیرد و تا پایان دوره به ۶۰ رخداد می‌رسد. همچنین «جنگل تصادفی» (با فراوانی ۵۹) دومین کلیدواژه پرتکرار است که رشد سریع آن به‌ویژه از سال ۱۳۹۹ به بعد، نشان‌دهنده افزایش توجه پژوهشگران به روش‌های یادگیری ماشین است. کلیدواژه «شبکه عصبی مصنوعی» با فراوانی ۴۳ نیز روندی صعودی را تجربه کرده و در کنار ماشین بردار پشتیبان و جنگل تصادفی، یکی از محورهای اصلی پژوهش‌ها را تشکیل داده است. علاوه بر این، کلیدواژه‌های «دسته‌بندی» (فراوانی ۲۷) و

«پیش‌بینی» (فراوانی ۲۳) از اواخر دهه ۱۳۹۰ رشد قابل‌توجهی را نشان می‌دهند که حاکی از گسترش کاربردهای یادگیری ماشین در تحلیل‌های پیشگویانه است. کلیدواژه‌های «داده‌کاوی»، «هوش مصنوعی»، «الگوریتم فراابتکاری» و «یادگیری ماشین افراطی» (هر کدام با فراوانی‌های ۲۰ تا ۲۱) در سال‌های پایانی دوره وارد مرحله پایداری در ادبیات پژوهشی شده‌اند. در مجموع، الگوی زمانی کلیدواژه‌ها بیانگر تداوم حضور روش‌های یادگیری ماشین در کنار گسترش برخی روش‌ها و مفاهیم تحلیلی و افزایش تنوع موضوعات مورد توجه در سال‌های پایانی دوره است. نمودار سیر تحول موضوعات در حوزه یادگیری ماشین (۱۳۹۶-۱۴۰۳)، سال‌هایی را آشکار می‌سازد که در آن کلیدواژه‌ها بیشتر فعال بوده‌اند (شکل ۶). خط زمانی (محور افقی) نشان می‌دهد که این موضوعات چه زمانی بیشترین رواج را در متون داشته‌اند. اندازه گره‌ها نیز با فراوانی کلیدواژه مطابقت دارد؛ به طوری که دایره‌های بزرگتر، موضوعات پرتکرارتر را نشان می‌دهند.

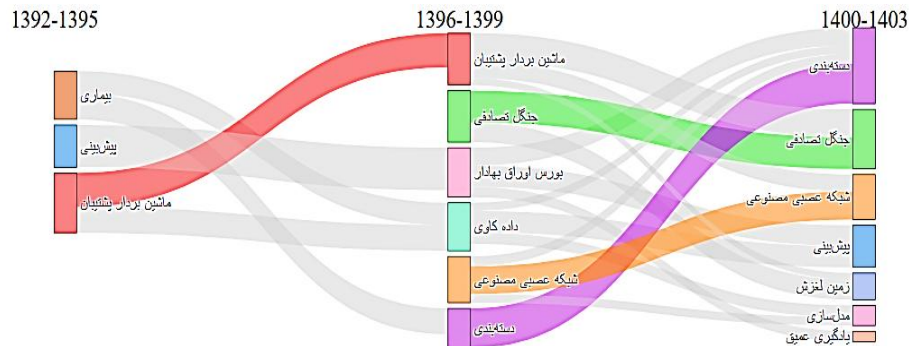


شکل ۶: سیر تحول موضوعات مورد توجه در مقالات حوزه یادگیری ماشین

طبق شکل ۶، کلیدواژه‌های مشاهده‌شده را می‌توان از نظر نقش مفهومی در پژوهش‌های

یادگیری ماشین به سه گروه اصلی تقسیم کرد: کلیدواژه‌های روش‌شناختی که به الگوریتم‌ها و روش‌های مورد استفاده در یادگیری ماشین اشاره دارند، کلیدواژه‌های کاربردی که بیانگر حوزه‌ها یا مسائل مورد استفاده از یادگیری ماشین هستند، و کلیدواژه‌های ناظر بر وظایف تحلیلی که نوع مسئله یا هدف تحلیل را بیان می‌کنند. در گروه نخست، مفاهیمی مانند «ماشین بردار پشتیبان»، «شبکه عصبی مصنوعی»، «جنگل تصادفی» و «یادگیری ماشین افراطی» قرار می‌گیرند. در مقابل، «بیماری»، «زمین لغزش» و «نقشه‌برداری رقومی» بیشتر ماهیت کاربردی دارند و نشان‌دهنده زمینه‌هایی هستند که روش‌های یادگیری ماشین در آنها به کار گرفته شده‌اند. همچنین، مفاهیمی مانند «دسته‌بندی»، «پیش‌بینی» و «رگرسیون» عمدتاً به وظایف یا انواع مسائل تحلیلی اشاره دارند و لزوماً یک روش یا الگوریتم اختصاصی یادگیری ماشین محسوب نمی‌شوند. از منظر زمانی، برخی کلیدواژه‌های کاربردی مانند «بیماری» پیشینه طولانی‌تری داشته و از سال ۱۳۹۶ تا ۱۴۰۲ در پژوهش‌ها حضور داشته‌اند، در حالی که «زمین لغزش» در سال‌های پایانی دوره، به‌ویژه در سال ۱۴۰۳، برجسته شده است. در میان کلیدواژه‌های روش‌شناختی نیز «ماشین بردار پشتیبان» و «شبکه عصبی مصنوعی» در سال ۱۴۰۱ به بیشترین میزان حضور رسیده‌اند و «یادگیری ماشین افراطی» در حوالی سال ۱۳۹۹ مورد توجه قرار گرفته است. بنابراین، مسیر مشاهده‌شده در شکل ۶ را می‌توان نه صرفاً به‌عنوان تغییر موضوعات، بلکه به‌عنوان تغییر هم‌زمان در روش‌های مورد استفاده، وظایف تحلیلی و زمینه‌های کاربردی یادگیری ماشین تفسیر کرد. شایان ذکر است که تفسیر کلیدواژه‌هایی مانند «پیش‌بینی» و «رگرسیون» مستلزم احتیاط کافی است؛ زیرا این مفاهیم در رشته‌های علمی مختلف کاربرد دارند و به‌تنهایی شاخص اختصاصی توسعه یادگیری ماشین محسوب نمی‌شوند. از این‌رو، در تفسیر سیر تحول موضوعی، میان روش‌های اختصاصی یادگیری ماشین، وظایف تحلیلی و زمینه‌های کاربردی تمایز گذاشته شده است.

تحلیل تکامل موضوعی در حوزه یادگیری ماشین که نشان‌دهنده مسیر تغییرات و تداوم موضوعات پژوهشی در سه بازه زمانی مختلف (۱۳۹۲-۱۳۹۵، ۱۳۹۶-۱۳۹۹ و ۱۴۰۰-۱۴۰۳) است در شکل ۷ نشان داده شده است. این تحلیل در واقع نمایی گسترده از مسیر جریان دانش و پویایی مفاهیم را در پژوهش‌های حوزه یادگیری ماشین به تصویر می‌کشد.



شکل ۷: نقشه تکامل موضوعی حوزه یادگیری ماشین در گذر زمان

همان‌طور که در شکل ۷ مشخص است، در بازه ۱۳۹۲-۱۳۹۵، سه مضمون «بیماری»، «پیش‌بینی» و «ماشین بردار پشتیبان» در میان مضامین منتخب این دوره قرار دارند؛ با این حال، این مفاهیم از نظر نقش در پژوهش‌های یادگیری ماشین ماهیت یکسانی ندارند. «بیماری» بیانگر یک زمینه کاربردی، «پیش‌بینی» بیانگر یک وظیفه تحلیلی و «ماشین بردار پشتیبان» یک روش یادگیری ماشین است. بنابراین، قرار گرفتن آنها در یک دوره زمانی به معنای هم‌سطح بودن مفهومی آنها نیست، بلکه بیانگر حضور هم‌زمان این سه نوع مضمون در ساختار موضوعی پژوهش‌های آن دوره است.

در بازه ۱۳۹۶-۱۳۹۹، تنوع مضامین افزایش یافته و علاوه بر تداوم «ماشین بردار پشتیبان»، مضامینی مانند «جنگل تصادفی»، «بورس اوراق بهادار»، «داده کاوی»، «شبکه عصبی مصنوعی» و «دسته‌بندی» در ساختار موضوعی مشاهده می‌شوند. در این میان، تداوم «ماشین بردار پشتیبان» میان دو دوره را می‌توان نشانه‌ای از استمرار توجه به این روش دانست؛ در حالی که ظهور «جنگل تصادفی» و «شبکه عصبی مصنوعی» نشان‌دهنده حضور روش‌های دیگری در کنار آن است. همچنین، مضامینی مانند «بورس اوراق بهادار» و «داده کاوی» نشان‌دهنده گسترش زمینه‌های کاربردی و موضوعی پژوهش‌ها در این دوره هستند.

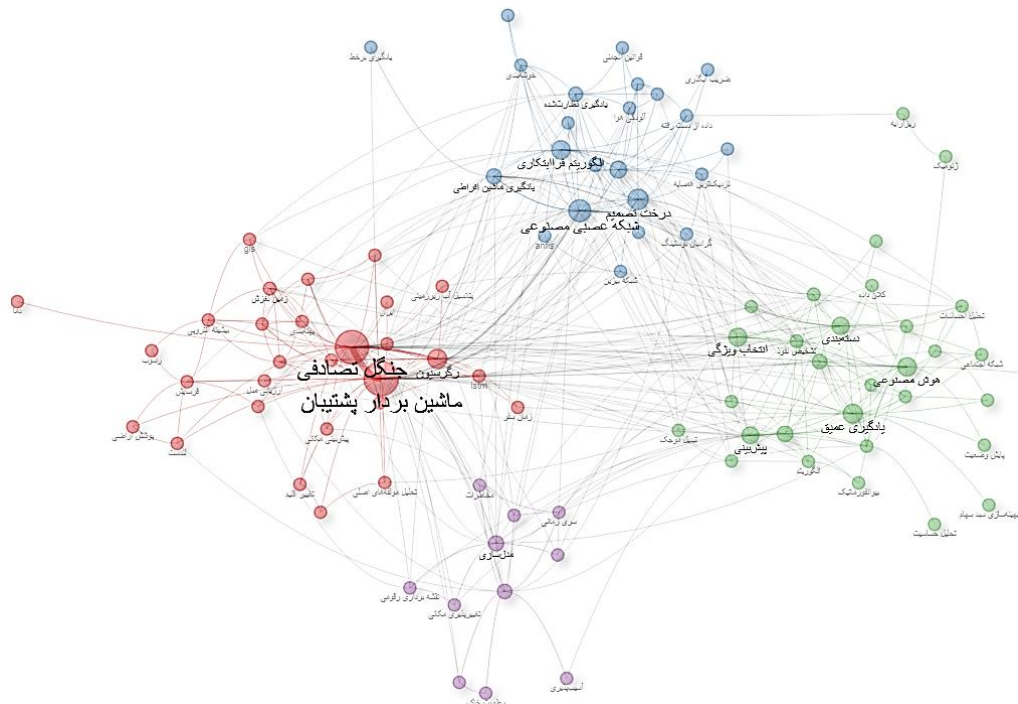
در بازه ۱۴۰۰-۱۴۰۳ نیز «جنگل تصادفی» و «شبکه عصبی مصنوعی» همچنان در ساختار موضوعی حضور دارند و در کنار آنها مضامینی مانند «پیش‌بینی»، «زمین لغزش»، «مدل‌سازی» و «یادگیری عمیق» مشاهده می‌شوند. از این میان، «زمین لغزش» بیشتر بیانگر یک زمینه کاربردی،

«یادگیری عمیق» یک رویکرد روش‌شناختی و «پیش‌بینی» و «مدل‌سازی» مفاهیمی ناظر بر وظایف یا فرایندهای تحلیلی هستند. بنابراین، نقشه تکامل موضوعی نشان می‌دهد که در طول دوره مورد مطالعه، علاوه بر تداوم برخی روش‌های یادگیری ماشین، تنوع زمینه‌های کاربردی و مفاهیم تحلیلی نیز تغییر کرده است.

در مجموع، شکل ۷ بیش از آنکه بیانگر حرکت خطی از «مفاهیم اولیه» به «تکنیک‌های پیچیده‌تر» باشد، تغییر در ترکیب روش‌های یادگیری ماشین، وظایف تحلیلی و زمینه‌های کاربردی در دوره‌های زمانی مختلف را نشان می‌دهد. همچنین، تداوم برخی مضامین روش‌شناختی در دوره‌های متوالی و ظهور برخی مضامین کاربردی و روش‌های جدیدتر، بیانگر تغییر تدریجی ساختار موضوعی پژوهش‌های یادگیری ماشین در ایران است.

سوال پنجم پژوهش: خوشه‌های موضوعی اصلی یادگیری ماشین در تحلیل شبکه هم‌رخدادی کلیدواژه‌ها کدامند و بر اساس نمودار راهبردی در چه وضعیتی (محرك، بنیادی، تخصصی و نوظهور) قرار دارند؟

در راستای کسب بینش ژرف‌تر درباره ساختار دانش پژوهش‌های فارسی حوزه یادگیری ماشین، در این بخش از پژوهش شبکه هم‌رخدادی کلیدواژه‌ها ترسیم شد (شکل ۸). این شبکه مشتمل بر ۹۶ کلیدواژه است. تراکم شبکه ۰.۰۴۶۳، میانگین درجه اش ۴.۳۹۶، و ماژولاریتی آن ۰.۲۹۷۵ می‌باشد. با توجه به این مقادیر، به نظر می‌رسد شبکه دارای تراکم پایین بوده و کلیدواژه‌ها ارتباطات محدودی با یکدیگر دارند. همچنین مقدار ماژولاریتی نشان‌دهنده تمایز قابل قبول خوشه‌های موضوعی شکل گرفته است.



شکل ۸: شبکه هم‌رخدادی کلیدواژه‌ها در حوزه یادگیری ماشین

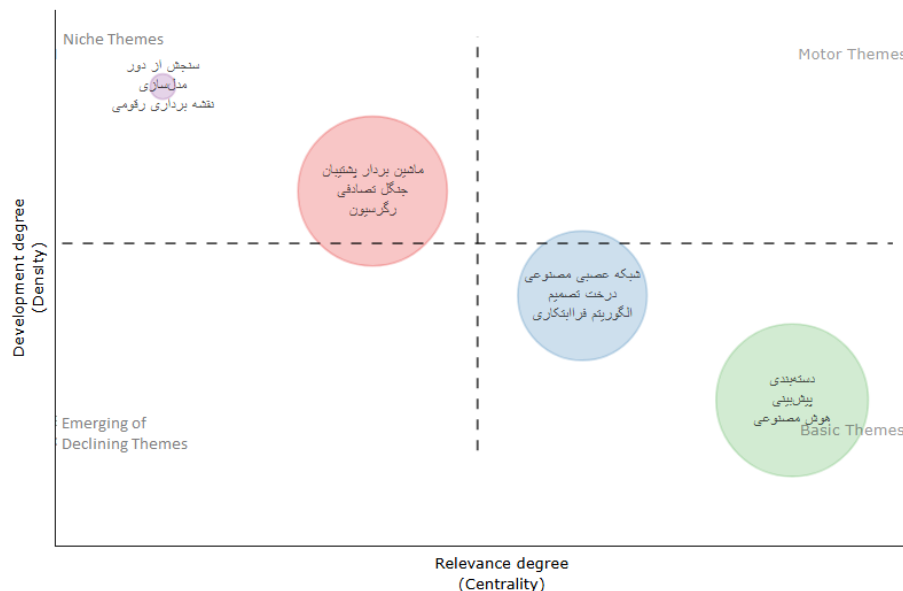
همانطور که در شکل مشخص است، تحلیل هم‌رخدادی کلیدواژه‌های مقالات نمایه‌شده در آی‌اس‌سی طی سال‌های ۱۳۹۲ تا ۱۴۰۳، چهار خوشه اصلی پژوهشی را در حوزه یادگیری ماشین آشکار ساخت که بیانگر ساختار مفهومی و جهت‌گیری‌های موضوعی این حوزه در ایران هستند. خوشه نخست عمدتاً از الگوریتم‌های پرکاربرد نظیر ماشین بردار پشتیبان، جنگل تصادفی و رگرسیون تشکیل شده است که در کنار مفاهیمی همچون زمین‌لغزش، سیل، پهنه‌بندی، حوزه آبخیز، کیفیت آب، فرسایش، GIS و داده‌های سنجش از دور قرار گرفته‌اند. این الگو نشان می‌دهد که بخش مهمی از پژوهش‌های یادگیری ماشین در کشور بر کاربرد این روش‌ها در مسائل مرتبط با علوم زمین، مدیریت منابع طبیعی و ارزیابی مخاطرات محیطی متمرکز بوده است. استفاده گسترده از الگوریتم‌های دسته‌بندی و پیش‌بینی در این مطالعات، بیانگر ظرفیت بالای یادگیری ماشین در تحلیل داده‌های مکانی و ارائه راهکارهای تصمیم‌یار در حوزه محیط‌زیست است.

خوشه دوم بیشتر منعکس‌کننده جنبه‌های روش‌شناختی و توسعه الگوریتم‌های یادگیری ماشین است. حضور شبکه عصبی مصنوعی، یادگیری ماشین افراطی، الگوریتم ژنتیک، الگوریتم‌های فراابتکاری، یادگیری جمعی، درخت تصمیم، خوشه‌بندی و شبکه بیزین نشان می‌دهد که پژوهشگران علاوه بر به کارگیری الگوریتم‌های موجود، به دنبال بهبود عملکرد مدل‌ها از طریق توسعه رویکردهای ترکیبی و بهینه‌سازی فرآیند یادگیری بوده‌اند. این خوشه بیانگر حرکت تدریجی پژوهش‌های داخلی از استفاده از مدل‌های منفرد به سمت طراحی چارچوب‌های هوشمندتر و کارآمدتر است. همچنین، حضور برخی حوزه‌های کاربردی نظیر آلودگی هوا و پیش‌بینی ورشکستگی در این خوشه، نشان‌دهنده تلاش برای انتقال این پیشرفت‌های روش‌شناختی به مسائل واقعی و بین‌رشته‌ای است.

خوشه سوم و چهارم به ترتیب نمایانگر گسترش دامنه کاربردهای یادگیری ماشین و توسعه مطالعات مبتنی بر داده‌های محیطی و سنجش از دور هستند. در خوشه سوم، هم‌رخدادی واژگانی نظیر یادگیری عمیق، داده‌کاوی، انتخاب ویژگی، بیماری، پردازش تصویر، تشخیص نفوذ، اینترنت اشیا، پردازش زبان طبیعی، تحلیل احساسات و امنیت سایبری، بیانگر نفوذ روزافزون یادگیری ماشین در حوزه‌هایی مانند سلامت، امور مالی، امنیت اطلاعات و تحلیل داده‌های متنی است. در مقابل، خوشه چهارم با تمرکز بر مفاهیمی همچون سنجش از دور، نقشه‌برداری رقومی، رطوبت خاک، خشکسالی، مخاطرات و تغییرپذیری مکانی، نقش فزاینده فناوری‌های مشاهده زمین و داده‌های محیطی را در توسعه کاربردهای یادگیری ماشین برجسته می‌سازد.

برای آنکه بینش عمیق‌تری نسبت به خوشه‌های حاصل از پژوهش‌های فارسی یادگیری ماشین کسب نماییم، لازم است جایگاه آنها را در نمودار راهبردی ترسیم و تحلیل نماییم. نمودار راهبردی، موضوعات تحقیقاتی را به چهار ربع طبقه‌بندی می‌کند: ۱. موضوعات محرک (بالا سمت راست): که مرکزیت بالا و تراکم بالا دارند. این موضوعات به خوبی توسعه یافته و برای حوزه بسیار مهم هستند و آنها «موتورهای» حوزه تحقیقاتی هستند؛ ۲. موضوعات تخصصی (بالا سمت چپ): که مرکزیت پایین و تراکم بالا دارند. این خوشه‌ها بسیار توسعه یافته اما منزوی و حاشیه‌ای هستند (تحقیقات تخصصی)؛ ۳. موضوعات نوظهور یا رو به زوال (پایین سمت چپ): که مرکزیت پایین و تراکم پایین دارند. این موضوعات یا تازه شروع به ظهور کرده‌اند یا در حال محو شدن هستند؛ و ۴. مضامین پایه/عرضی (پایین سمت راست): که مرکزیت بالا و تراکم پایینی

دارند. این مضامین برای این حوزه مهم و مرکزی هستند، ولی ممکن است به خوبی در درون توسعه نیافته باشند (آنها به عنوان مفاهیم بنیادی عمل می‌کنند). نمودار راهبردی مربوط به نقشه موضوعی پژوهش‌های فارسی حوزه یادگیری ماشین در شکل ۹ نشان داده شده است که در ادامه، آن را به ترتیب بر اساس چهار ربع موجود در نمودار راهبردی بررسی می‌کنیم.



شکل ۹: نمودار راهبردی مربوط به نقشه موضوعی پژوهش‌های حوزه یادگیری ماشین

با توجه به موقعیت خوشه‌ها در نمودار راهبردی، دو خوشه در ربع پایین سمت راست و دو خوشه در ربع بالا سمت چپ جای گرفته‌اند. خوشه‌های واقع در ربع پایین سمت راست که شامل مفاهیم از قبیل «دسته‌بندی»، «پیش‌بینی» و «هوش مصنوعی» هستند جزو خوشه‌های بنیادی در حوزه یادگیری ماشین محسوب می‌شوند. این موضوعات با دارا بودن درجه مرکزیت بالا، ستون فقرات علمی این حوزه محسوب می‌شوند و به مثابه دانش پایه و زیربنایی، پیونددهنده‌ی مباحث مختلف در ادبیات یادگیری ماشین در ایران هستند. با توجه به اینکه این خوشه‌ها در میانه نمودار و در نزدیکی ناحیه‌ی موضوعات محرک (ربع بالا سمت راست) قرار دارد، به نظر می‌رسد این موضوعات در مرحله‌ای از پویایی قرار دارند که با افزایش مرکزیت و چگالی خود، در حال تبدیل شدن به پیشران‌های اصلی و محرک‌های توسعه‌ی دانش یادگیری

ماشین در ایران هستند.

دو خوشه نیز در ربع بالا سمت چپ قرار گرفته‌اند. یکی از این خوشه‌ها عمدتاً شامل الگوریتم‌ها و روش‌های یادگیری ماشین مانند «ماشین بردار پشتیبان»، «جنگل تصادفی» و «رگرسیون» است و خوشه دیگر بیشتر بر حوزه‌های کاربردی همچون «سنجش از دور»، «مدل‌سازی» و «نقشه‌برداری رقومی» تمرکز دارد. قرارگیری این دو خوشه در ناحیه موضوعات تخصصی نشان می‌دهد که هم برخی روش‌های یادگیری ماشین و هم برخی حوزه‌های کاربردی آن، از انسجام درونی و توسعه‌یافتگی مناسبی برخوردار شده‌اند، اما با توجه به مرکزیت کمتر، هنوز نقش محوری در کل ساختار مفهومی این حوزه ایفا نمی‌کنند. یافته‌های این تحلیل حاکی از آن است که در بازه‌ی زمانی مذکور، هیچ خوشه‌ای در ناحیه‌ی موضوعات محرک (بالا سمت راست) و موضوعات نوظهور یا رو به زوال (پایین سمت چپ) قرار نگرفته‌اند.

به طور کلی، این ساختار نشان می‌دهد که فضای پژوهشی یادگیری ماشین در ایران در وضعیت «گذار علمی» قرار دارد؛ به گونه‌ای که پژوهش‌ها به‌جای تمرکز در یک هسته‌ی محرک واحد، به دو بخش مجزا تقسیم شده‌اند: بخشی که بر اصول بنیادین هوش مصنوعی متمرکز است (ربع پایین سمت راست) و بخشی که با تخصصی‌سازی عمیق، بر حل مسائل محیطی و زمین‌شناسی از طریق الگوریتم‌های پیشرفته تمرکز دارد (ربع بالا سمت چپ). این الگو گویای آن است که حوزه‌ی یادگیری ماشین در ایران، حوزه‌ای پویا و کاربردی است که مسیر رشد و تکامل خود را از طریق پیوند روش‌های بنیادی با کاربردهای خاص تخصصی طی می‌کند.

۶. بحث و نتیجه‌گیری

هدف این پژوهش ترسیم ساختار دانش، الگوهای همکاری علمی و سیر تحول موضوعی پژوهش‌های فارسی حوزه یادگیری ماشین در ایران بر اساس مدارک نمایه‌شده در پایگاه آی‌اس‌سی طی سال‌های ۱۳۹۲ تا ۱۴۰۳ بود. یافته‌ها نشان داد که یادگیری ماشین در ایران طی یک دهه گذشته از یک حوزه نوظهور و محدود به کاربردهای خاص، به حوزه‌ای با رشد سریع، ساختار دانشی نسبتاً تثبیت‌شده و دامنه‌ای گسترده از کاربردهای میان‌رشته‌ای تبدیل شده است. این تحول نه تنها در افزایش کمی تولیدات علمی، بلکه در تغییر ساختار مفهومی، گسترش شبکه‌های همکاری و تنوع کاربردهای پژوهشی نیز قابل مشاهده است.

نتایج نشان داد پژوهش‌های یادگیری ماشین در ایران از مرحله پذیرش اولیه عبور کرده و به یکی از حوزه‌های فعال پژوهشی کشور تبدیل شده است. این بخش از نتایج با یافته‌های مطالعات بین‌المللی همخوانی دارد. برای مثال، لاندبرگ و همکاران (۲۰۲۴) با تحلیل نزدیک به ۴۰۰ هزار

مدرک علمی نشان دادند که یادگیری ماشین از سریع ترین حوزه های در حال رشد در نظام علمی جهان است. همچنین رضایی زاده و عسکری (۱۴۰۰) رشد فزاینده تولیدات علمی هوش مصنوعی را در سطح جهانی گزارش کرده اند که می تواند بیانگر آن باشد که توسعه یادگیری ماشین بخشی از یک تحول جهانی مبتنی بر گسترش داده های دیجیتال، افزایش توان پردازشی و توسعه فناوری های هوش مصنوعی است.

در سطح ساختار مفهومی، نتایج پژوهش نشان داد که مفاهیمی نظیر «ماشین بردار پشتیبان»، «جنگل تصادفی»، «شبکه عصبی مصنوعی»، «دسته بندی» و «پیش بینی» در مرکز شبکه دانشی این حوزه قرار دارند. این یافته با مطالعات داخلی و خارجی همسو است. ابراهیمی ترکمانی و محمودی (۱۴۰۴) نیز یادگیری ماشین و شبکه های عصبی را از مهم ترین محورهای پژوهش های هوش مصنوعی در ایران معرفی کرده اند. در سطح بین المللی نیز چن و همکاران (۲۰۲۱) و فانگ و همکاران (۲۰۲۵) نشان دادند که الگوریتم های یادگیری ماشین و شبکه های عصبی از اصلی ترین هسته های دانشی در بسیاری از زیر حوزه های علمی محسوب می شوند.

با این وجود، یافته های پژوهش حاضر یک ویژگی مهم دیگر را نیز آشکار ساخت. برخلاف برخی مطالعات جهانی که رشد سریع موضوعاتی نظیر هوش مصنوعی مولد، مدل های زبانی بزرگ، هوش مصنوعی قابل تبیین و حکمرانی هوش مصنوعی را گزارش کرده اند (Lundberg et al., 2024)، ساختار مفهومی پژوهش های فارسی هنوز عمدتاً حول الگوریتم های کلاسیک و کاربردهای تثبیت شده شکل گرفته است. این وضعیت را می توان از دو منظر تفسیر کرد. از یک سو، این یافته بیانگر تداوم تمرکز پژوهش ها بر الگوریتم های کلاسیک و کاربردهای تثبیت شده است. از سوی دیگر، می تواند نشان دهنده فاصله زمانی میان مباحث پژوهشی داخلی و مرزهای نوظهور دانش جهانی باشد. بنابراین، اگرچه جامعه علمی ایران توانسته است در جذب و بومی سازی فناوری های اصلی یادگیری ماشین موفق عمل کند، اما برای حضور مؤثرتر در مرزهای دانش جهانی نیازمند ورود سریع تر به حوزه های نوظهور است.

یکی از یافته های قابل توجه پژوهش حاضر، غلبه کاربردهای محیط زیستی، منابع طبیعی، سنجش از دور و مخاطرات طبیعی در میان خوشه های موضوعی بود. این الگو با نتایج توسن و همکاران (۱۴۰۴) و سایر مطالعات کاربردی داخلی همخوانی دارد. از منظر تفسیری، این یافته نشان می دهد که توسعه یادگیری ماشین در ایران بیش از آنکه تحت تأثیر نیازهای صنایع دیجیتال

یا اقتصاد پلتفرمی باشد، در پاسخ به مسائل بومی و چالش‌های ملی شکل گرفته است. موضوعاتی نظیر خشکسالی، فرسایش خاک، زمین‌لغزش، کیفیت آب و مدیریت منابع طبیعی از جمله مسائل راهبردی کشور هستند و پژوهشگران ایرانی از یادگیری ماشین به‌عنوان ابزاری برای تحلیل و مدیریت این مسائل بهره گرفته‌اند. این ویژگی را می‌توان یکی از نقاط تمایز پژوهش‌های یادگیری ماشین در ایران نسبت به بسیاری از کشورهای توسعه‌یافته دانست؛ کشورهایی که سهم بیشتری از پژوهش‌های آنها به حوزه‌هایی مانند تجارت الکترونیک، سامانه‌های توصیه‌گر، و هوش مصنوعی مولد اختصاص دارد.

نتایج تحلیل تکامل موضوعی نشان داد که در سال‌های اخیر، در کنار تداوم حضور الگوریتم‌های یادگیری ماشین، کلیدواژه‌های مرتبط با کاربردهایی مانند «زمین‌لغزش» و «سنجش از دور» نیز پررنگ‌تر شده‌اند. همچنین، ظهور کلیدواژه «یادگیری عمیق» نشان‌دهنده ورود تدریجی برخی روش‌های جدیدتر به پژوهش‌های فارسی این حوزه است. این الگو بیانگر افزایش تنوع موضوعی و گسترش دامنه کاربردهای یادگیری ماشین در پژوهش‌های فارسی است. این یافته با نتایج کانتورویچ و کانتوروویچ (۲۰۲۵) همخوانی دارد که نشان دادند بسیاری از حوزه‌های علمی در فرایند تحول خود از پژوهش‌های نظری به سمت پژوهش‌های کاربردی و مسئله‌محور حرکت می‌کنند.

از منظر همکاری علمی، یافته‌های پژوهش نشان داد که دانشگاه آزاد اسلامی، دانشگاه تهران و دانشگاه تربیت مدرس نقش محوری در تولید دانش این حوزه دارند و هسته اصلی شبکه همکاری را تشکیل می‌دهند. این نتیجه با یافته‌های ظهوریان و نادعلی و همکاران (۱۴۰۱) و یزدی و رضایی شریف‌آبادی (۱۳۹۶) همخوانی دارد. با وجود شکل‌گیری شبکه‌ای نسبتاً پراکنده از همکاری‌های دانشگاهی، دانشگاه‌های آزاد اسلامی، تهران و تربیت مدرس جزو هسته‌های مرکزی این شبکه بوده و نقش محوری در شبکه همکاری و تولیدات علمی این حوزه ایفا می‌کنند. از این رو، توسعه همکاری‌های بین‌دانشگاهی، تقویت مشارکت دانشگاه‌های کمتر فعال و گسترش همکاری‌های بین‌المللی می‌تواند به گسترش ظرفیت‌های پژوهشی و تقویت شبکه همکاری علمی در این حوزه کمک کند.

از بُعد نظری، یافته‌های این پژوهش از دیدگاه تکامل حوزه‌های علمی حمایت می‌کند که بر اساس آن، رشد دانش صرفاً با افزایش تعداد انتشارات قابل تبیین نیست، بلکه تغییرات ساختار

مفهومی، شبکه‌های همکاری و سیر تحول موضوعی نیز باید مورد توجه قرار گیرد. در پژوهش حاضر، این سه بُعد به صورت مکمل بررسی شدند و نتایج نشان داد که پژوهش‌های یادگیری ماشین در ایران طی دوره مورد مطالعه، از نظر ساختار مفهومی، شبکه همکاری علمی و تنوع موضوعی دستخوش تغییر شده‌اند. همچنین، یافته‌ها بیانگر آن است که اگرچه الگوریتم‌های کلاسیک همچنان جایگاه محوری در ساختار مفهومی این حوزه دارند، برخی روش‌ها و کاربردهای جدید نیز در سال‌های اخیر وارد پژوهش‌های فارسی شده‌اند. با این حال، حضور محدود این موضوعات در مقایسه با روندهای جهانی نشان می‌دهد که توسعه این حوزه همچنان نیازمند توجه بیشتر به موضوعات نوظهور و گسترش همکاری‌های علمی است.

در مجموع، یافته‌های این پژوهش نشان داد که تولیدات علمی فارسی حوزه یادگیری ماشین در فاصله سال‌های ۱۳۹۲ تا ۱۴۰۳ رو به رشد بوده، اما ساختار مفهومی آن همچنان عمدتاً بر الگوریتم‌های کلاسیکی مانند ماشین بردار پشتیبان، جنگل تصادفی و شبکه‌های عصبی مصنوعی استوار است و موضوعات نوظهور سهم کمتری در پژوهش‌های فارسی دارند. همچنین، تحلیل شبکه همکاری علمی نقش محوری چند دانشگاه در تولید دانش این حوزه و تحلیل خوشه‌های موضوعی، تمرکز قابل توجه پژوهش‌ها بر کاربردهای محیط‌زیستی، منابع طبیعی و سنجش از دور را نشان داد. بر پایه این یافته‌ها، پیشنهاد می‌شود وزارت علوم، تحقیقات و فناوری و صندوق‌های حمایت از پژوهش و فناوری در تدوین برنامه‌های حمایتی، به توسعه پژوهش در موضوعات نوظهور شناسایی شده در این مطالعه و تقویت همکاری میان دانشگاه‌های پیشرو و سایر دانشگاه‌ها توجه بیشتری داشته باشند. همچنین، دانشگاه‌ها و مراکز پژوهشی می‌توانند با گسترش فعالیت‌های میان‌رشته‌ای و توسعه پژوهش در حوزه‌هایی که در نقشه راهبردی از جایگاه ضعیف‌تری برخوردار بوده‌اند، به متوازن‌تر شدن ساختار موضوعی پژوهش‌های یادگیری ماشین در کشور کمک کنند.

۷. محدودیت‌های پژوهش

پژوهش حاضر، همانند سایر مطالعات علم‌سنجی، با محدودیت‌هایی همراه است که در تفسیر یافته‌ها باید مورد توجه قرار گیرد. نخستین محدودیت به منبع داده‌ها بازمی‌گردد. داده‌های پژوهش صرفاً از پایگاه آی‌اس‌سی استخراج شده‌اند و مدارک فارسی نمایه‌شده در سایر پایگاه‌های

اطلاعاتی نظیر مرکز اطلاعات علمی جهاد دانشگاهی (اس‌آی‌دی) و مگیران در تحلیل حاضر وارد نشده‌اند که می‌تواند یکی از محدودیت‌های پژوهش از نظر جامعیت داده‌ها محسوب شود. رویکرد تک‌پایگاهی می‌تواند موجب عدم شمول برخی مدارکی شود که در سایر پایگاه‌ها نمایه شده‌اند و در نتیجه، تصویر به‌دست آمده الزاماً نماینده تمامی تولیدات علمی فارسی حوزه یادگیری ماشین در ایران نیست.

دومین محدودیت، ماهیت زمانی داده‌های پژوهش است. تحلیل حاضر وضعیت تولیدات علمی حوزه یادگیری ماشین را تا سال ۱۴۰۳ بررسی می‌کند و با توجه به سرعت بالای تحولات این حوزه، به‌ویژه در زمینه‌هایی مانند هوش مصنوعی مولد و مدل‌های زبانی بزرگ، ممکن است برخی موارد جدید که هنوز در قالب مدارک علمی فارسی نمایه‌شده در ISC منعکس نشده‌اند، در نتایج پژوهش مشاهده نشده باشند.

محدودیت دیگر پژوهش حاضر مربوط به ماهیت تفسیری برخی روش‌های به‌کاررفته است. تحلیل هم‌رخدادی کلیدواژه‌ها، روابط موضوعی میان مفاهیم را بر اساس حضور هم‌زمان آنها در اسناد نشان می‌دهد، اما این روابط لزوماً به معنای ارتباط علی یا مفهومی مستقیم میان موضوعات نیستند. همچنین، جایگاه یک مضمون در نمودار راهبردی به‌تنهایی نمی‌تواند نوظهور یا رو به افول بودن آن را اثبات کند و تفسیر این وضعیت نیازمند توجه هم‌زمان به شواهد زمانی و تغییرات موضوعی است. از این‌رو، نتایج پژوهش باید به‌عنوان بازنمایی ساختار و سیر موضوعی موجود در مجموعه اسناد مورد مطالعه تفسیر شوند، نه به‌عنوان سنج‌ای قطعی از میزان توسعه یک حوزه علمی. و محدودیت آخر آنکه، در این پژوهش به تاسی از اغلب پژوهش‌های علم‌سنجی در ایران، مدارک مربوط به دانشگاه‌های جامع (از قبیل دانشگاه آزاد اسلامی، دانشگاه پیام نور، ...) ذیل یک واحد یکپارچه در نظر گرفته شده است؛ این درحالیست که این دانشگاه‌ها از واحدهای زیادی تشکیل شده‌اند و اتخاذ چنین رویکردی ممکن است منجر به شکل‌گیری یک سوگیری به نفع این دانشگاه‌های جامع گردد.

۸. پیشنهادهایی برای پژوهش‌های آینده

نخست، پیشنهاد می‌شود در مطالعات آینده، داده‌های پایگاه آی‌اس‌سی با سایر پایگاه‌های نمایه‌کننده منابع علمی فارسی، نظیر مرکز اطلاعات علمی جهاد دانشگاهی (اس‌آی‌دی) و مگیران،

تلفیق شوند تا تصویر جامع‌تری از تولیدات علمی فارسی حوزه یادگیری ماشین به دست آید. همچنین، مقایسه ساختار مفهومی پژوهش‌های فارسی با پژوهش‌های بین‌المللی نمایه‌شده در پایگاه‌هایی نظیر اسکوپس و وب‌آف‌ساینس می‌تواند به شناسایی شباهت‌ها و تفاوت‌های مباحث پژوهشی داخلی و بین‌المللی کمک کند. همچنین، با توجه به برجستگی موضوعات محیط‌زیستی، منابع طبیعی و سنجش از دور در یافته‌های پژوهش حاضر، مطالعات آینده می‌توانند به‌طور اختصاصی ساختار موضوعی و سیر تحول این زیرحوزه‌ها را بررسی کرده و شکاف‌های پژوهشی و فرصت‌های میان‌رشته‌ای مرتبط با کاربرد یادگیری ماشین در مسائل محیط‌زیستی و منابع طبیعی ایران را شناسایی کنند.

فهرست منابع

- ابراهیمی ترکمانی، ف. و محمودی، م. ۱۴۰۴. ترسیم نقشه علمی هوش مصنوعی در برون‌داد پژوهشگران دانشگاه‌های شهر تهران: تحلیل بیبلیومتریک برون‌دادهای علمی منتشر شده در پایگاه استنادی اسکوپس در بازه ۲۰۱۵-۲۰۲۵. *مطالعات کاربردی علم سنجی*، ۱(۳): ۵۵-۷۶.
- <https://doi.org/10.22091/apss.2024.11834.1024>
- اکبری، م. و جلال‌نیا، ر. ۱۴۰۴. ارائه الگوی سیاست‌گذاری توسعه هوش مصنوعی در راستای برنامه هفتم توسعه. *سیاست‌گذاری عمومی*، ۱۱(۲): ۱۱۹-۱۳۶. <https://doi.org/10.22059/jppolicy.2025.102510>
- آقایاری، پ. و کریم‌میان، ز. ۱۴۰۳. حکمرانی و هوش مصنوعی: روایت علم سنجی از دو داستان در هم تنیده. *سیاست علم و فناوری*، ۱۷(۴): ۱-۱۷. <https://doi.org/10.22034/jstp.2025.11814.1837>
- باغانی، ا.، الهی، ش.، حسن‌زاده، ع. و رجب‌زاده قطری، ع. ۱۴۰۲. تاکسونومی شناسایی مشتریان صنعت بانکی با بکارگیری یادگیری ماشین: مروری نظام‌مند با رویکرد فراترکیب. *پژوهش‌نامه پردازش و مدیریت اطلاعات*، ۳۹(۱): ۴۲۹-۳۹۴. <https://doi.org/10.22034/jipm.2023.704230>
- توسن، م.، شمشیرگران، ر.، دستورانی، م. ۱۴۰۴. استفاده از ترکیب رویکردهای سنجش از دور و یادگیری ماشین در پیش‌بینی پارامترهای هیدرولوژیکی: یک مطالعه علم‌سنجی. *تحقیقات آب و خاک ایران*، ۵۶(۳): ۸۵۰-۸۲۵. <https://doi.org/10.22059/ijswr.2025.386927.669850>
- جواهری، م.، وکیلی‌مفرد، ح.، امیری، م. و خاصه، ع. ۱۴۰۰. ترسیم و تحلیل نقشه دانش حوزه پژوهش‌های زنان و زایمان با استفاده از تحلیل هم‌رخدادی واژگان. *پژوهش‌نامه علم‌سنجی*، ۷(۱۴): ۱۳۷-۱۵۶. <https://doi.org/10.22070/rsci.2020.5289.1359>

- خاصه، ع.، بانگ آور، م.، قاضی زاده، ح.، و مختاری، ح. ۴۰۰. یک دهه فعالیت فصلنامه تحقیقات اطلاع‌رسانی و کتابخانه‌های عمومی (۱۳۸۷-۱۳۹۶): تحلیل کتاب‌سنجی و دیداری‌سازی برونداد علمی. پژوهش‌نامه علم‌سنجی، ۷ (۱۳): ۵۳-۷۶. <https://doi.org/10.22070/rsci.2020.4724.1319>
- رحیمی، ص.، سهیلی، ف.، و خاصه، ع. ۱۴۰۳. تحلیلی بر پژوهش‌های حوزه سواد در ایران: ساختار مفهومی سواد در نمایه استنادی علوم ایران. کتابداری و اطلاع‌رسانی، ۲۷ (۴): ۱۹۵-۲۲۰.
- رستمی، م.، سهیلی، ف.، و خاصه، ع. ۱۳۹۹. ساختار دانش در پروانه‌های ثبت اختراع حوزه کشف دانش: مصورسازی با استفاده از تحلیل هم‌رخدادی واژگان. پژوهشنامه علم‌سنجی، ۶ (۲): ۴۱-۶۰.
- رضایی زاده، کریم، و عسکری، مجتبی ۱۴۰۰. هوش مصنوعی براساس پایگاه استنادی وب‌آو ساینس؛ فراتحلیل از نوع علم‌سنجی. پژوهش‌های نوین در مدیریت کارآفرینی و توسعه کسب و کار، ۲ (۱): ۲۸۶-۳۰۶.
- سهرابی، ب.، مانیان، ا.، و آرمان، م. ۱۳۹۵. ارزیابی کیفیت پر مشاهده‌ترین وب‌سایت‌های خبری در ایران مبتنی بر روش یادگیری ماشین. پژوهشنامه پردازش و مدیریت اطلاعات، ۳۲ (۲): ۵۵۱-۵۷۹.
- <https://doi.org/10.35050/JIPM010.2017.053>
- سهیلی، ف.، و کرانیان، پ. ۱۳۹۷. روند موضوعی مفاهیم حوزه علم اطلاعات و دانش‌شناسی ایران بر اساس تحلیل هم‌رخدادی واژگان. فصلنامه مطالعات ملی کتابداری و سازماندهی اطلاعات، ۲۹ (۲): ۱۷۱-۱۹۰.
- <https://doi.org/10.30484/nastinfo.2018.2233>
- سیاحی، ع.، علوی، ع.، و جادریان، م. ۱۴۰۵. مطالعه علم‌سنجی روش‌های یادگیری ماشین و یادگیری عمیق برای تشخیص بیماری کبد چرب غیرالکلی در تصویربرداری پزشکی. مجله علم‌سنجی کاسپین، ۱۳ (۱): ۱۲-۲۵.
- فرزین یزدی، م.، و رضایی شریف آبادی، س. ۱۳۹۶. بررسی تولیدات علمی حوزه موضوعی هوش مصنوعی در کشورهای خاورمیانه طی سال‌های ۱۹۹۶ تا ۲۰۱۴. پژوهش‌نامه علم‌سنجی، ۳ (۲): ۹۷-۱۱۴.
- <https://doi.org/10.22070/rsci.2017.512>
- فرهنگی، س.، خاصه، ع.، و ابراهیمی دینانی، آ. ۱۳۹۷. ترسیم ساختار دانش در پژوهش‌های عرفانی: تحلیل هم‌واژگانی. ادبیات عرفانی و اسطوره‌شناسی، ۱۴ (۵۱): ۲۰۷-۲۳۹.
- <https://sanad.iau.ir/Journal/jmmlq/Article/1049904>
- ظهوریان نادعلی، ا.، اجاقی، ح.، و سلیمانی روزبهانی، ف. ۱۴۰۱. بررسی مقایسه‌ای الگوی همکاری دانشگاه، صنعت و دولت در تحقیقات علمی هوش مصنوعی ایران، چین و آمریکا. علوم و فنون مدیریت اطلاعات، ۴ (۴): ۴۵-۷۴. <https://doi.org/10.22091/stim.2022.8360.1816>
- کاشانی، م.، و داستانی، م. ۱۴۰۳. تحلیل روند موضوعی تولیدات علمی پژوهشگران ایرانی در زمینه هوش مصنوعی در علوم پزشکی: مطالعه علم‌سنجی. مجله اطلاع‌رسانی پزشکی نوین، ۱۰ (۳): ۲۳۱-۲۴۶.

References

- Abdullah, A. A., Mohammed, N. S., Khanzadi, M., Asaad, S. M., Abdul, Z. K., & Maghddid, H. S. 2025. In-depth Analysis on Machine Learning Approaches: Techniques, Applications, and Trends. *ARO-The Scientific Journal of Koya University*, 13(1): 190-202. <https://doi.org/10.14500/aro.12038>

- Aghayari, P., & Karimmian, Z. 2024. Governance and Artificial Intelligence: A Scientometric Narrative of Two Interwoven Stories. *Journal of Science and Technology Policy*, 17(4): 1-17. <https://doi.org/10.22034/jstp.2025.11814.1837> [In Persian]
- Akbari, M., & Jalalniya, R. 2025. Presenting The Policy Model of Artificial Intelligence Development in Line with the 7th Development Plan. *Iranian Journal of Public Policy*, 11(2): 119-136. <https://doi.org/10.22059/jppolicy.2025.102510> [In Persian]
- Alnuaimi, A. F. A. H., & Albaldawi, T. H. K. 2024. Concepts of statistical learning and classification in machine learning: An overview. *BIO Web of Conferences*, 97, 00129. <https://doi.org/10.1051/bioconf/20249700129>
- Aria, M., & Cuccurullo, C. 2017. bibliometrix: An R-tool for comprehensive science mapping analysis. *Journal of Informetrics*, 11(4): 959-975. <https://doi.org/10.1016/j.joi.2017.08.007>
- Azadi Ahmadabadi, G. 2025. Predicting Scientific Research Impacts in Biotechnology by Machine Learning Algorithms. *Scientometrics Research Journal*, 11(1): 1-24. <https://doi.org/10.22070/rsci.2024.18868.1719> [In Persian]
- Baghani, E., Elahi, S., Hasanzadeh, A., & Rajabzadeh Ghatari, A. 2023. Taxonomy of Customers Identification in Banking Industry Using Machine Learning: a Systematic Review with a Meta-Synthesis Approach. *Iranian Journal of Information Processing and Management*, 39(1): 394-429. <https://doi.org/10.22034/jipm.2023.704230> [In Persian]
- Belmonte, J. L., Segura-Robles, A., Moreno-Guerrero, A. J., & Parra-González, M. E. 2020. Machine learning and big data in the impact literature. A bibliometric review with scientific mapping in web of science. *Symmetry*, 12(4): 495.
- Bishop, C. M. 2006. Pattern recognition and machine learning. Springer.
- Callon, M., Courtial, J. P., & Laville, F. 1991. Co-word analysis as a tool for describing the network of interactions between basic and technological research: The case of polymer chemistry. *Scientometrics*, 22(1): 155-205. <https://doi.org/10.1007/BF02019280>
- Chen, Z., Hu, G., Zheng, M., Song, X., & Chen, L. 2021. Bibliometrics of Machine Learning Research Using Homomorphic Encryption. *Mathematics*, 9(21): 2792. <https://doi.org/10.3390/math9212792>
- Cobo, M. J., López-Herrera, A. G., Herrera-Viedma, E., & Herrera, F. 2011. Science mapping software tools: Review, analysis, and cooperative study among tools. *Journal of the American Society for Information Science and Technology*, 62(7): 1382-1402. <https://doi.org/10.1002/asi.21525>
- De Felice, F., & Polimeni, A. 2020. Coronavirus disease (COVID-19): a machine learning bibliometric analysis. *in vivo*, 34(3 suppl): 1613-1617.
- Ebrahimi Torkamani, F., & Mahmoodi, M. 2024. Mapping the Scientific Landscape of Artificial Intelligence in the Publications of Researchers from Tehran-based Universities: A Bibliometric Analysis of Scholarly Outputs Indexed in Scopus from 2015 to 2025. *Applied Scientometric Studies*, 1(3): 55-76. <https://doi.org/10.22091/apss.2024.11834.1024> [In Persian]
- Elshenawy, A., Mohammed, A., & Hamouda, S. 2025. The Evolution of Deep Learning: Models, Applications, and Future Directions. 2025 *International Mobile, Intelligent, and Ubiquitous Computing Conference (MIUCC)*: 1-8. <https://doi.org/10.1109/MIUCC66482.2025.11196834>
- Ezugwu, A. E., Oyelade, O. N., Ikotun, A. M., Agushaka, J. O., & Ho, Y. S. 2023. Machine Learning Research Trends in Africa: A 30 Years Overview with Bibliometric Analysis Review: AE Ezugwu et al. *Archives of Computational Methods in Engineering*, 30(7): 4177-4207.
- Ezugwu, A. E., Ho, Y.-S., Egwuiche, O. S., Ekundayo, O. S., Van Der Merwe, A., Saha, A. K., & Pal, J. 2024. *Classical Machine Learning: Seventy Years of Algorithmic Learning Evolution* (Version 2). arXiv. <https://doi.org/10.48550/ARXIV.2408.01747>
- Fang, Y., Wu, Y., & Gao, L. 2025. Machine learning-based myocardial infarction bibliometric analysis. *Frontiers in Medicine*, Volume 12-2025. <https://doi.org/10.3389/fmed.2025.1477351>

- Farhangi, S., Khasseh, A. A., Ebrahimi Dinani, A. 2018. Mapping Knowledge Structure in Mystical Researches: A Co-Word Analysis. *Mytho-Mystic Literature Quarterly Journal*, 14(51): 207-239. <https://sanad.iau.ir/Journal/jmmlq/Article/1049904> [In Persian]
- Farzin Yazdi, M. and Rezaei Sharifabadi, S. 2017. Scientific publications in the subject area of Artificial Intelligence in Middle Eastern countries during 1996 to 2014. *Scientometrics Research Journal*, 3(2): 97-114. <https://doi.org/10.22070/rsci.2017.512> [In Persian]
- Higaki, A., Uetani, T., Ikeda, S., & Yamaguchi, O. 2020. Co-authorship network analysis in cardiovascular research utilizing machine learning (2009–2019). *International Journal of Medical Informatics*, 143, 104274. <https://doi.org/10.1016/j.ijmedinf.2020.104274>
- Jaramillo, A. M., Williams, H. T. P., Perra, N., & Menezes, R. 2023. The structure of segregation in co-authorship networks and its impact on scientific production. *EPJ Data Science*, 12(1): 47. <https://doi.org/10.1140/epjds/s13688-023-00411-8>
- Javaheri, M., Vakilimofrad, H., Amiri, M., & Khasseh, A. A. 2021. Mapping Knowledge Structure of Obstetrics and Gynecology studies: A Co-Word Analysis. *Scientometrics Research Journal*, 7(2): 137-156. <https://doi.org/10.22070/rsci.2020.5289.1359> [In Persian]
- Jordan, M. I., & Mitchell, T. M. 2015. Machine learning: Trends, perspectives, and prospects. *Science*, 349(6245): 255–260.
- Kashani M., & Dastani M. 2024. Analysis of thematic trends in scientific publications of Iranian researchers in artificial intelligence for medical sciences: A scientometric study. *Journal of Modern Medical Information Sciences*, 10 (3): 231-246.
- Khasseh, A. A., BangAvar Mashak, M., Ghazizadeh, H., & Mokhtari, H. 2021. A Decade of Performance of the Research on Information Science and Public Libraries Quarterly (2008-2017): A Bibliometric Analysis and Visualization of Scientific Output. *Scientometrics Research Journal*, 7(1): 53-76. <https://dor.isc.ac/dor/20.1001.1.24233773.1400.7.13.4.9> [In Persian]
- Khasseh, A. A., Soheili, F., Moghaddam, H. S., & Chelak, A. M. 2017. Intellectual structure of knowledge in iMetrics: A co-word analysis. *Information Processing & Management*, 53(3): 705-720.
- Kantorowicz-Reznichenko, E., & Kantorowicz, J. 2025. Law & Economics at sixty: Mapping the field with bibliometric and machine learning tools. *Journal of Economic Surveys*, 39(1): 201-245.
- LeCun, Y., Bengio, Y. & Hinton, G. 2015. Deep learning. *Nature* 521, 436–444. <https://doi.org/10.1038/nature14539>
- Leydesdorff, L. 2001. The challenge of Scientometrics: The Development, Measurement, and Self-Organization of Scientific Communications (2nd ed.). Universal Publishers. <https://www.amazon.com/Challenge-Scientometrics-Development-Self-Organization-Communications/dp/1581126816>
- Lundberg, L., Boldt, M., Borg, A., & Grahm, H. 2024. Bibliometric Mining of Research Trends in Machine Learning. *AI*, 5(1): 208–236. <https://doi.org/10.3390/ai5010012>
- Mehrad, J. & Ghane, M. 2021. 6.3 The Islamic World Science Citation Center (ISC): The Construction and Application. In R. Ball (Ed.), *Handbook Bibliometrics* (pp. 431-454). Berlin, Boston: De Gruyter Saur. <https://doi.org/10.1515/9783110646610-040>
- Mingers, J., & Leydesdorff, L. 2015. A review of theory and practice in scientometrics. *European Journal of Operational Research*, 246(1): 1–19. <https://doi.org/10.1016/j.ejor.2015.04.002>
- Mokhtarpour, R., & Khasseh, A. A. 2020. Who is who in library and information science research? The integrative application of scholarly influence indicators. *Journal of Librarianship and Information Science*, 52(4): 1186-1196. <https://doi.org/10.1177/0961000620907956>
- Patel, M. B., & Patel, M. M. 2025. Machine Learning Algorithms: An Extensive Survey, Applications and Future Research Directions. In C. S. In, N. S. Londhe, N. Bhatt, & M. Kitsing (Eds.),

- Information Systems for Intelligent Systems* (Vol. 430, pp. 609–619). Springer Nature Singapore. https://doi.org/10.1007/978-981-96-1206-2_47
- Rahimi, S., Soheili, F., & Khasseh, A. A. 2024. An Analysis of Literacy Research in Iran: The Conceptual Structure of Literacy in Islamic World Science Citation Center. *Library and Information Sciences*, 27(4): 195-220 [In Persian]
- Rath, S., Pandey, M., & Rautaray, S. S. 2025. Synergistic review of automation impact of big data, AI, and ML in current data transformative era. *F1000Research*, 14, 253. <https://doi.org/10.12688/f1000research.161477.2>
- Rezaeizadeh, K., & Askari, M. (2021). Artificial Intelligence Based on Web of Science Citation Database; A Meta-Analysis of Scientometric Type. *New Research in Entrepreneurship Management and Business Development*, 2(1): 286-306.
- Rostami, M., Soheili, F., & Khasseh, A. 2020. Knowledge Structure in Knowledge Discovery Patents: Visualization based on Co-Word Analysis. *Scientometrics Research Journal*, 6(2): 41-60. <https://doi.org/10.22070/rsci.2019.3841.1240> [In Persian]
- Sayyahi A, Alavi S E, Jaderian M. 2026. A Scientometric Study of Machine Learning and Deep Learning Methods for Non-Alcoholic Fatty Liver Disease Detection in Medical Imaging. *CJS* 13 (1) :12-25 [In Persian]
- SCImago Journal & Country Rank. 2025. Country Rankings in Artificial Intelligence – All Regions (1996-2025). Retrieved from <https://www.scimagojr.com/countryrank.php?category=1702&area=1700>
- Shokouhian, M., Asemi, A., Shabani, A and Cheshmesohrabi, M. 2020. Presenting a Thematic Model of Health Scientific Productions Using Text-Mining Methods. *Iranian Journal of Information Processing and Management*, 35(2): 553-574. <https://doi.org/10.35050/IJPM010.2020.061>
- Singh, K. A. 2025. The Role of Machine Learning in Various Sectors: A Comprehensive Review. *International Journal for Research in Applied Science and Engineering Technology*, 13(5): 3152–3159. <https://doi.org/10.22214/ijraset.2025.70952>
- Soheili, F., Khasseh, A. and koranian, P. 2018. Thematic trends of concepts in Knowledge and Information Science based on co-word analysis in Iran. *Librarianship and Information Organization Studies*, 29(2): 171-190. <https://doi.org/10.30484/nastinfo.2018.2233> [In Persian]
- Sohrabi, B., Manian, A., & Arman, M. 2017. Evaluation of Most Visited News Websites in Iran based on Machine learning. *Iranian Journal of Information Processing and Management*, 32(2): 551-579. <https://doi.org/10.35050/IJPM010.2017.053> [In Persian]
- Surur, F. M., Mamo, A. A., Gebresilassie, B. G., Mekonen, K. A., Golda, A., Behera, R. K., & Kumar, K. 2025. Unlocking the power of machine learning in big data: A scoping survey. *Data Science and Management*, 8(4): 519–535. <https://doi.org/10.1016/j.dsm.2025.02.004>
- Tosan, M., Shamshirgaran, R. and Dastourani, M. 2025. Application of the Combination of Remote Sensing and Machine Learning Approaches in Predicting Hydrological Parameters: A Bibliometric Analysis. *Iranian Journal of Soil and Water Research*, 56(3): 825-850. <https://doi.org/10.22059/ijswr.2025.386927.669850> [In Persian]
- Turing, A. M. 2009. Computing Machinery and Intelligence. In R. Epstein, G. Roberts, & G. Beber (Eds.), *Parsing the Turing Test* (pp. 23–65). Springer Netherlands. https://doi.org/10.1007/978-1-4020-6710-5_3
- van Eck, N.J., & Waltman, L. 2010. Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics* 84, 523–538. <https://doi.org/10.1007/s11192-009-0146-3>
- Zohoorian Nadali, I., Ojaghi, H., & Soleymani Roozbahani, F. 2022. Comparative Study of University, Industry, and Government Cooperation Model in Artificial Intelligence Scientific Research of Iran,

China, and America. *Sciences and Techniques of Information Management*, 8(4): 45-74.
<https://doi.org/10.22091/stim.2022.8360.1816> [In Persian]

Mapping the Conceptual Structure and Thematic Evolution of Persian-Language Machine Learning Research Indexed in the ISC database: A Bibliometric Study

Hamed Barangi: Department of Computer Engineering, Payame Noor University, Tehran, Iran. hbarangi@pnu.ac.ir (0000-0003-2444-3474)

Ali Akbar Khasseh*: Associated Professor, Department of Knowledge and Information Science, Payame Noor University, Tehran, Iran. khasseh@pnu.ac.ir (0000-0001-5664-4671)

Abstract

Machine learning has emerged as one of the most influential branches of artificial intelligence and has experienced remarkable growth in scientific production during the last decade. Despite the increasing volume of research in Iran in the field of machine learning, there are limited scientometric studies on the knowledge structure and thematic evolution of Persian-language research in this domain. Therefore, this study aims to map the knowledge structure and thematic evolution of Persian machine learning research indexed in the Islamic World Science Citation Center (ISC) during the period 2013–2024. This research adopts a descriptive-analytical scientometric approach. Data were retrieved from the ISC database, resulting in 399 relevant articles after data cleaning and content validation. To improve the accuracy of the analyses, keywords and institutional affiliations were standardized manually. The data were analyzed using Biblioshiny and VOSviewer. Scientometric indicators such as scientific production trends, three-field plots (authors–journals–keywords), university collaboration networks, keyword co-occurrence analysis, thematic evolution maps, and strategic diagrams were employed to investigate the intellectual and thematic structure of the field. The results reveal a substantial increase in scientific production, from only one article in 2013 to 106 articles in 2023, indicating

the rapid expansion of machine learning research in Iran. Among the 136 academic institutions that had articles in the field of machine learning, Islamic Azad University (101 articles), the University of Tehran (65 articles), and Tarbiat Modares University (23 articles) were identified as the most productive contributors. Furthermore, among the 1,873 identified keywords, "support vector machine," "random forest," and "artificial neural network" were the most frequent research topics, with frequencies of 60, 59, and 43, respectively. Thematic evolution analysis indicated a gradual transition from foundational machine learning concepts and traditional algorithms toward more specialized applications, including deep learning, modeling, landslide prediction, and geospatial analysis. Co-occurrence analysis of keywords led to the identification of four main thematic clusters in the field of machine learning: environmental and natural resource applications, methodological and algorithmic developments, interdisciplinary applications, and remote sensing-based studies. The strategic diagram identified "classification," "prediction," and "artificial intelligence" as fundamental themes with high centrality, whereas clusters related to "support vector machines," "random forests," and "remote sensing" were categorized as highly developed but specialized themes. The findings suggest that Persian-language machine learning research in Iran has entered a stage of relative growth characterized by a transition from algorithm-centered studies to application-oriented and interdisciplinary research. The field demonstrates a stable intellectual foundation while simultaneously expanding into new domains of application.

Keywords: Machine Learning, Artificial Intelligence, Knowledge Structure, Scientometrics, Co-word Analysis, Thematic Evolution.

حامد برنگی متولد ۱۳۶۴، دکتری تخصصی خود را در رشته مهندسی کامپیوتر از دانشگاه اصفهان دریافت کرده است. وی هم اکنون عضو هیئت علمی گروه مهندسی کامپیوتر دانشگاه پیام نور می باشد. مهندسی نرم افزار مدل رانده، سیستمهای آگاه به زمینه و اینترنت اشیا از جمله علایق پژوهشی وی است.	
علی اکبر خاصه دارای مدرک دکتری علم اطلاعات و دانش شناسی است. ایشان در حال حاضر دانشیار گروه علم اطلاعات و دانش شناسی دانشگاه پیام نور است. علم سنجی و ارزیابی علم از جمله علایق پژوهشی وی است.	