

چکیده توسعه یافته

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

حامد برنگی: گروه مهندسی کامپیوتر، دانشگاه پیام نور، تهران، ایران

علی اکبر خاصه: دانشیار گروه علم اطلاعات و دانش‌شناسی، دانشگاه پیام نور، تهران، ایران

khasseh@pnu.ac.ir

مقدمه: 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 پایش علم و فناوری جهان اسلام (آی‌اس‌سی) طی سال‌های ۱۳۹۲ تا ۱۴۰۳ است.

روش‌شناسی: 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. 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. پس از آماده‌سازی داده‌ها، اطلاعات در محیط نرم‌افزار بیبلوشاینی و ووز-ویور تحلیل شد. برای ترسیم روابط همکاری میان دانشگاه‌ها، ۴۱ دانشگاه که دارای حداقل ۳ مقاله بودند وارد تحلیل نهایی شدند. همچنین در بخش تحلیل هم‌واژگانی، پس از بررسی فراوانی کلیدواژه‌ها، تعداد ۹۶ کلیدواژه با حداقل فراوانی ۳ وارد تحلیل نهایی شدند تا میزان هم‌رخدادی آنان در مقاله‌ها به دست آید و سپس شبکه هم‌واژگانی و نمودار راهبردی حوزه داده کاوی با استفاده از بیبلوشاینی ترسیم شد.

یافته های اصلی: نتایج نشان داد تعداد مقالات منتشر شده 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 دانشگاه که دارای مقاله در حوزه یادگیری ماشین بودند، دانشگاه آزاد اسلامی، دانشگاه تهران و دانشگاه تربیت مدرس به ترتیب با ۱۰۱، ۶۵ و ۲۳ مقاله، فعال ترین مراکز علمی در این حوزه بودند. علاوه بر این، از میان ۱۸۷۳ کلیدواژه شناسایی شده، کلیدواژه های «ماشین بردار پشتیبان»، «جنگل تصادفی» و «شبکه عصبی مصنوعی» به ترتیب با فراوانی ۶۰، ۵۹ و ۴۳، پرتکرارترین موضوعات پژوهشی محسوب می شوند. بررسی سیر تحول موضوعات بیانگر حرکت تدریجی پژوهش ها از مفاهیم پایه ای و الگوریتم های کلاسیک به سمت کاربردهای تخصصی تر و موضوعاتی نظیر یادگیری عمیق، مدل سازی و کاربردهای محیط زیستی و مخاطرات طبیعی حرکت کرده است. همچنین تحلیل شبکه هم رخدادی کلیدواژه ها چهار خوشه اصلی شامل کاربردهای محیط زیستی، توسعه الگوریتم ها، کاربردهای میان رشته ای و مطالعات مبتنی بر سنجش از دور را آشکار ساخت. نتایج نمودار راهبردی نیز نشان داد موضوعات «دسته بندی»، «پیش بینی» و «هوش مصنوعی» در زمره مضامین بنیادی قرار دارند، در حالی که خوشه های مرتبط با «ماشین بردار پشتیبان»، «جنگل تصادفی» و «سنجش از دور» به عنوان موضوعات تخصصی شناخته می شوند.

بحث و نتیجه گیری: به طور کلی، نتایج نشان داد که یادگیری ماشین در ایران طی یک دهه گذشته از یک حوزه نوظهور و محدود به کاربردهای خاص، به حوزه ای با رشد سریع، ساختار دانشی نسبتاً تثبیت شده و دامنه ای گسترده از کاربردهای میان رشته ای تبدیل شده است. این تحول نه تنها در افزایش کمی تولیدات علمی، بلکه در تغییر ساختار مفهومی، گسترش شبکه های همکاری و تنوع کاربردهای پژوهشی نیز قابل مشاهده است. همچنین، یافته ها بیانگر آن است که اگرچه الگوریتم های کلاسیک همچنان جایگاه محوری در ساختار مفهومی این حوزه دارند، برخی روش ها و کاربردهای جدید نیز در سال های اخیر وارد پژوهش های فارسی شده اند. با این حال، حضور محدود این موضوعات در مقایسه با روندهای جهانی نشان می دهد که توسعه این حوزه همچنان نیازمند توجه بیشتر به موضوعات نوظهور و گسترش همکاری های علمی است.

داده های پژوهش صرفاً از پایگاه آی اس سی استخراج شده اند و مدارک فارسی نمایه شده در سایر پایگاه های اطلاعاتی نظیر مرکز اطلاعات علمی جهاد دانشگاهی (اس آید) و مگیران در تحلیل حاضر وارد نشده اند که می تواند یکی از محدودیت های پژوهش از نظر جامعیت داده ها محسوب شود. در نتیجه، تصویر به دست آمده الزاماً نماینده تمامی تولیدات علمی فارسی حوزه یادگیری ماشین در ایران نیست. برای پژوهش های آینده پیشنهاد می شود داده های پایگاه آی اس سی با سایر پایگاه های نمایه کننده منابع علمی فارسی تلفیق شوند تا تصویر جامع تری از تولیدات علمی فارسی حوزه یادگیری ماشین به دست آید.

کلیدواژه ها: یادگیری ماشین، هوش مصنوعی، ساختار دانش، علم سنجی، تحلیل هم رخدادی کلیدواژه ها، تکامل موضوعی

Extended Abstract

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)

Introduction: 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 bibliometric 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.

Methodology: The present study is a bibliometric study with a descriptive-analytical approach that aims to outline the status of Persian scientific productions in the field of machine learning and analyse the conceptual structure of this field. To achieve this goal, bibliometric indicators and scientific network analysis methods were used, including university-level co-occurrence analysis, keyword co-occurrence analysis, and scientific mapping.

The source of the research data is the Islamic World Science Citation Centre (ISC). Data retrieval was performed in May 2025 in the title field of the ISC advanced search. The search term included “machine learning” and all indexed Persian research articles from 2013 to 2014 were retrieved. Finally, 399 articles were selected as the final data set for analysis. In order to increase the validity of the analyses, the data were subjected to a filtering and standardization process before performing bibliometric processing. After data preparation, the information was analysed in the BiblioShiny and Vosviewer software environments. To outline the collaborative relationships between universities, 41 universities that had at least 3 articles were included in the final analysis. Also, in

the co-word analysis section, after examining the frequency of keywords, 96 keywords with a minimum frequency of 3 were included in the final analysis to obtain their co-occurrence rate in the articles, and then the co-word network and strategic diagram of the data mining field were drawn using Biblioshiny.

Main findings: The results showed that the number of published articles increased from one article in 2013 to 106 articles in 2023, which indicates the rapid growth of this field in Iran. Among the 136 universities that had articles in the field of machine learning, Islamic Azad University, University of Tehran, and Tarbiat Modares University were the most active scientific universities in this field with 101, 65, and 23 articles, respectively. In addition, among the 1873 identified keywords, the keywords “support vector machine”, “random forest” and “artificial neural network” are considered the most frequent research topics with frequencies of 60, 59 and 43, respectively. The study of the evolution of the topics indicates a gradual movement of research from basic concepts and classical algorithms towards more specialized applications and topics such as deep learning, modelling and environmental applications and natural hazards. Also, the analysis of the keyword co-occurrence network revealed four main clusters including environmental applications, algorithm development, interdisciplinary applications and remote sensing-based studies. The results of the strategic diagram also showed that the topics “classification”, “forecasting” and “artificial intelligence” are among the fundamental topics, while the clusters related to “support vector machine”, “random forest” and “remote sensing” are known as specialized topics.

Discussion and conclusions: Overall, the results showed that machine learning in Iran has transformed from an emerging field limited to specific applications over the past decade into a rapidly growing field with a relatively established knowledge structure and a wide range of interdisciplinary applications. This transformation is visible not only in the quantitative increase in scientific production, but also in the change in conceptual structure, expansion of collaboration networks and diversity of research applications. Also, the findings indicate that although classical algorithms still have a central place in the conceptual structure of this field, some new methods and applications have also entered Persian research in recent years. However, the limited presence of these topics compared to global trends indicates that the development of this field still requires more attention to emerging topics and the expansion of scientific collaborations.

The research data were extracted solely from the ISC database, and Persian documents indexed in other databases such as The Scientific Information

Database (SID) and Magiran were not included in the present analysis, which can be considered one of the limitations of the research in terms of data comprehensiveness. As a result, the obtained picture is not necessarily representative of all Persian scientific productions in the field of machine learning in Iran. For future research, it is suggested that the ISC database data be combined with other databases indexing Persian scientific resources to obtain a more comprehensive picture of Persian scientific productions in the field of machine learning.

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