

Article

Not peer-reviewed version

Intelligent Learning on Multidimensional Data Streams: A Bibliometric Analysis of Research Evolution and Future Directions

[Gary Reyes](#)*, [Roberto Tolozano-Benites](#), [Laura Lanzarini](#), [Waldo Hasperué](#), [Julio Barzola-Monteses](#)

Posted Date: 3 November 2025

doi: 10.20944/preprints202511.0005.v1

Keywords: data streams; intelligent learning; bibliometric analysis; knowledge discovery; information processing; research trends



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

Intelligent Learning on Multidimensional Data Streams: A Bibliometric Analysis of Research Evolution and Future Directions

Gary Reyes ^{1,2,*}, Roberto Tolozano-Benites ¹, Laura Lanzarini ³, Waldo Hasperué ³
and Julio Barzola-Monteses ^{1,2}

- ¹ Carrera de Sistemas Inteligentes, Universidad Bolivariana del Ecuador, Campus Durán Km 5.5 vía Durán Yaguachi, Durán 092405, Ecuador
- ² Facultad de Ciencias Matemáticas y Físicas, Universidad de Guayaquil, Cdla. Universitaria Salvador Allende, Guayaquil 090514, Ecuador
- ³ Instituto de Investigación en Informática LIDI (Centro CICIPBA), Facultad de Informática, Universidad Nacional de La Plata, Buenos Aires CP1900, Argentina
- * Correspondence: gxreyesz@ube.edu.ec

Abstract

Intelligent learning applied to multidimensional data streams has established itself as a rapidly expanding field, driven by the growth of ubiquitous computing and the Internet of Things. The complexity of these streams, characterized by their high dimensionality, variability, and continuous nature, poses significant challenges for traditional approaches to analysis. This study presents a bibliometric analysis of scientific output indexed in Scopus between 2015 and 2025, with the aim of identifying trends, challenges, and opportunities in this field. The results show sustained growth in publications, a marked interdisciplinary orientation, and a diversity of applications including transportation, biomedicine, energy, and information systems. Likewise, there is a geographical concentration in certain leading countries and uneven development in terms of international collaboration. This work contributes to mapping the current state of the field and points to future lines of research aimed at its consolidation.

Keywords: data streams; intelligent learning; bibliometric analysis; knowledge discovery; information processing; research trends

1. Introduction

The rise of ubiquitous computing and the Internet of Things (IoT) has generated a data ecosystem characterized by continuous and heterogeneous streams that challenge traditional approaches to analysis [1,2]. These streams are highly dimensional, have complex temporal patterns, and arrival rates that can reach millions of instances per second [3–5].

Against this backdrop, intelligent learning on multidimensional streams is emerging as an interdisciplinary field that integrates machine learning, real-time processing, and high-performance distributed systems [5]. This field faces key challenges such as the curse of dimensionality, concept drift, and the need for computational scalability [6,7].

Intelligent learning in high-dimensional streams consists of extracting useful knowledge and adaptive decisions in real time in response to dynamic changes in data [8,9]. It is based on four critical aspects: multidimensionality, continuous temporality, dynamic adaptability, and computational efficiency [10–13].

One of the main problems with dimensionality is that it severely affects algorithms in temporal contexts, where data dispersion and time window limitations reduce predictive effectiveness [14–16]. Concept drift can occur simultaneously across multiple dimensions and scales, requiring specialized detection and adaptation mechanisms [17–19]. Added to this is increasing computational complexity, which limits the viability of traditional methods in very high-dimensional scenarios [20,21].

Advances in this field have an impact on multiple areas. In IoT and cyber-physical systems, it facilitates anomaly detection and resource optimization [22,23]. In financial markets, it enables adaptive trading strategies and real-time risk management [24]. In social media and digital platforms, it allows for dynamic content personalization and trend detection [25,26]. In digital health, it contributes to continuous monitoring and personalized treatments [27,28].

This paper critically reviews the current state of learning about multidimensional streams, characterizing recent trends, identifying technical challenges, evaluating innovative solutions, and pointing out research gaps. It also proposes a systematic framework for classifying and comparing approaches, providing a synthesis of developments and future directions.

2. Related Works

During 2000-2010, the foundations of streaming learning were established with algorithms such as VFDT [29] and characterizations of the main challenges [30,31]. Between 2010-2018, methodologies were consolidated with frameworks such as MOA and techniques for concept drift detection and ensemble methods [32–35]. Starting in 2018, the focus expanded to high dimensionality, with approaches to online feature selection, adaptive dimensionality reduction, and early integrations with deep learning [36–42].

Recent trends include Continuous Learning Streaming, which integrates memory and synaptic consolidation mechanisms to mitigate catastrophic forgetting [1,43–45]; specialized Transformer architectures for tensor data that reduce processing complexity [2,3]; and temporal tensor factorization techniques that improve prediction and compression of multidimensional streams [4,6,7,46,47]. Also noteworthy are advances in online feature selection with norm-based approaches $\ell_{1,2}$, reinforcement learning, and adaptive sparsity [8,11,12,48], as well as specialized computational architectures for scalability [49,50].

The main challenges today include the problem of dimensionality in temporal contexts, which degrades neighborhood-based methods [15,16,51,52]; multidimensional concept drift, which requires adaptive data structures and models [14]; and computational scalability, limited by the combinatorial explosion in tensor and attention operations [13,53].

Among the most relevant applications are IoT systems with anomaly detection in sensor streams [10,54,55], optimization of service quality metrics in distributed systems, and high-frequency financial analysis, where multiple data sources are integrated for real-time risk management.

Current approaches face theoretical limitations such as Hughes' paradox and the No-Free-Lunch theorems, which reinforce the need for domain specialization and advanced regularization methods [56,57]. Practical challenges of technological heterogeneity, interpretability, and observability also persist. Emerging directions include Quantum-Enhanced Streaming, with quantum acceleration potential [58], and Neuromorphic Stream Processing, which promises extreme energy efficiency [59]. Likewise, paradigms such as Causal Stream Learning and Continuous Meta-Learning are emerging, aimed at robustness and knowledge transfer in non-stationary environments.

3. Materials and Methods

To analyze scientific production in the field of “Intelligent Learning on Multidimensional Data Streams”, a comprehensive methodology was designed that integrated bibliometric analysis with advanced data visualization tools.

Information was collected from articles indexed exclusively in the Scopus database, considering those that addressed this topic using algorithms, models, or applications related to intelligent learning.

In the first phase, the metadata of the retrieved records was refined, limiting the collection to publications between 2015 and September 2025. In order to ensure the relevance of the set, a strict filtering process was applied, manually discarding works not directly related to the topic of study. The criteria established were that the selected articles must contain the words “intelligent”, “learning”, and “multidimensional” simultaneously, as well as incorporate at least one of the complementary

terms “data flows” or “data streams” in their description. As a result, 594 articles were identified, covering both theoretical contributions and applied developments on intelligent learning in complex and multidimensional data streams.

The analysis of bibliographic networks was carried out using VOSviewer (version 1.6.20), specialized software that enables the representation of co-citation networks, collaboration between authors, and thematic distribution [60]. This tool facilitated the construction of structural graphs that show the interconnections between the most influential works, journals, and authors in the field. Bibliometrix (version 4.1.4) and its biblioshiny interface, developed in R, were also used to explore the evolution of topics, the use of keywords, and the identification of emerging trends. Both applications, which are open source and free, offer advantages in terms of replicability and enable other researchers to apply this approach in similar studies.

Finally, Shannon entropy was used to evaluate the concentration and diversity of variables such as the distribution of authors, countries, and research areas. This measure allowed us to estimate the degree of homogeneity in the dispersion of the data and provided a solid basis for interpreting patterns of concentration in authorship, the dynamics of international collaborations, and the diversification of lines of research related to intelligent learning applied to multidimensional data streams.

4. Results

The analysis was carried out based on the bibliographic metadata of the documents indexed in the Scopus database, restricting the selection to those works related to the topic of Intelligent Learning on Multidimensional Data Streams. In total, 594 publications were identified, distributed across 276 sources (journals, books, among others) in the period between 2015 and 2025. These contributions were produced by 1,588 authors, of whom only 41 signed documents individually, while the majority were co-authored works, with an average of 4.07 authors per publication. The set analyzed shows an annual growth rate of 39.38% and an average of 9.25 citations per document, reflecting sustained and expanding interest in the subject. In addition, the publications included 4,771 terms in Keywords Plus and 5,730 keywords provided by the authors, demonstrating the diversity of approaches associated with the field of study.

The results presented in Table 1 show that scientific production on Intelligent Learning on Multidimensional Data Streams is mainly concentrated in Computer Science (30.33%) and Engineering (19.98%), confirming the predominance of approaches focused on the design of computational models and the development of algorithms applied to spatial data processing. However, the presence of other disciplines such as Mathematics (11.44%), Physics and Astronomy (6.03%), and Decision Sciences (5.49%) reveals a growing interest in addressing the topic from complementary perspectives that incorporate theoretical foundations, mathematical modeling, and decision-making support criteria.

Table 1. Main areas of research assigned to the sample papers. Source: Scopus.

Research Areas	Records	% of 1276
Computer Science	387	30.33%
Engineering	255	19.98%
Mathematics	146	11.44%
Physics and Astronomy	77	6.03%
Decision Sciences	70	5.49%
Total of the 5 main research areas	935	73.28%

Together, these five areas represent 73.28% of the records analyzed, highlighting a significant thematic concentration, but also the existence of considerable scope for strengthening interdisciplinary research. In particular, the contribution of mathematics and physical sciences opens up the possibility of delving deeper into advanced analytical methods, while the incorporation of decision sciences highlights the potential for application in areas related to mobility, logistics, and strategic planning.

Table 2 shows that scientific output on intelligent learning applied to multidimensional data streams has experienced sustained growth over the last decade. Since 2015, there has been a progressive increase in the number of publications, with notable values in 2019 (59.09%), 2022 (70.27%), and 2025 (42.20%). This trend demonstrates a consolidation of interest in the academic community, particularly in the last five years, when the volume of articles reached a remarkable growth rate.

Table 2. Number of articles published per year. Source: Scopus.

Years	Items	Annual Growth Rate
2015	6	100.00%
2016	12	100.00%
2017	19	58.33%
2018	22	15.79%
2019	35	59.09%
2020	41	17.14%
2021	37	-9.76%
2022	63	70.27%
2023	84	33.33%
2024	109	29.76%
2025	166	42.20%
Total	594	39.38%

Although there were slight declines in some periods, such as in 2021 with a decrease of 9.76%, the overall trend reflects a rapidly expanding field, with a significant increase from 6 articles in 2015 to 166 in 2025. These results suggest that the subject is maturing, driven by the development of more advanced methodologies, the availability of large volumes of data, and the relevance of its application in various scientific and technological fields. In this sense, the growth dynamics confirm that this is a constantly evolving domain with an increasing potential to impact interdisciplinary research.

4.1. Geographical Distribution of the Corresponding Authors

Table 3 shows that China is the country with the highest scientific output in the field of intelligent learning on multidimensional data streams, accounting for 55.2% of total publications. It is followed by India and the United States with 7.2% and 3.9%, respectively. Together, the top ten countries account for 74.8% of the articles published in this line of research. In terms of the type of collaboration, Single-Country Publications (SCP) greatly predominate over Multi-Country Publications (MCP). Canada stands out as the country with the highest proportion of internationally co-authored publications (77.8%), followed by Australia (50.0%) and Germany (40.0%). In contrast, Italy, Ukraine, and Spain have no publications in collaboration with other countries.

Table 3. Ten countries of corresponding authors. Source: Scopus.

Country	Articles	Frequency	SCP	MCP	MCP Ratio
China	328	55.2%	282	46	14.0%
India	43	7.2%	35	8	18.6%
USA	23	3.9%	21	2	8.7%
Australia	10	1.7%	5	5	50.0%
Germany	10	1.7%	6	4	40.0%
Canada	9	1.5%	2	7	77.8%
Italy	6	1.0%	6	0	0.0%
Ukraine	6	1.0%	6	0	0.0%
Korea	5	0.8%	4	1	20.0%
Spain	5	0.8%	5	0	0.0%
Total 10 countries	445	74.8%	372	73	22.9%

Table 4 shows the leading countries in terms of impact measured by citations. The overall average number of citations per article is 14.93. The United States (34.0) and Canada (55.4) stand out with averages well above this figure, reflecting greater recognition of their contributions in the scientific community. Germany (37.5), Japan (42.3), and Bangladesh (46.5) also far exceed the average. In contrast, China and India, despite being the countries with the most articles published, have an average of 6.8 citations per publication, one of the lowest among the countries with the highest output, suggesting a lower relative visibility of their work compared to their high productivity.

Table 4. Top ten total citations by country. Source: Scopus.

Country	Total Citations	Average Citations of Articles
China	2218	6.80
USA	781	34.00
Canada	499	55.40
Germany	375	37.50
India	291	6.80
Spain	157	31.40
Japan	127	42.30
Australia	114	11.40
Bangladesh	93	46.50
Saudi Arabia	85	28.30
Total (all countries)	5299	14.93

In general terms, the data show that China and the United States lead scientific production in this field, albeit with different profiles: China accounts for the largest number of publications, while the United States has a much higher average number of citations per article, which demonstrates the greater influence and visibility of its contributions. These differences could be related to factors such as the perceived quality of the work, access to high-impact journals, or the strength of scientific collaboration networks. Likewise, the proportion of international publications varies significantly: countries such as Canada and Germany stand out for their high level of cooperation, while others, such as Italy, Ukraine, and Spain, have no collaborations in this field. The fact that more than half of the articles are concentrated in only ten countries confirms that production on intelligent learning about multidimensional data streams is not yet evenly distributed globally, which opens up opportunities to strengthen research in emerging regions through international cooperation networks.

4.2. Main Publication Sources

The analysis of the main sources of publication in the field of intelligent learning on multidimensional data streams, presented in Table 5, shows a clear predominance of book series and scientific journals over conference proceedings. In particular, the Lecture Notes in Computer Science collection ranks first with 23 articles, confirming the close relationship between the topic and computer science and artificial intelligence. It is followed by other series such as Communications in Computer and Information Science (15 articles), Smart Innovation, Systems and Technologies (11), and Advances in Intelligent Systems and Computing (9), which stand out for their role in disseminating recent advances related to technological innovation and intelligent systems.

Likewise, high-impact journals such as the IEEE Internet of Things Journal, IEEE Access, IEEE Transactions on Industrial Informatics, and IEEE Transactions on Intelligent Transportation Systems, with between nine and six publications, show that the field is also consolidating in peer-reviewed spaces focused on practical applications in the Internet of Things, industrial computing, and intelligent transportation systems. Finally, the presence of publications such as the Proceedings of SPIE and the Journal of Image and Graphics reinforces the multidisciplinary nature of the area, where perspectives from engineering, image processing, and applied sciences converge.

Table 5. The ten most relevant sources. Source: Scopus.

Sources	Articles	Type
Lecture Notes in Computer Science	23	Book Series
Communications in Computer and Information Science	15	Journal
Smart Innovation, Systems and Technologies	11	Book series
Advances in Intelligent Systems and Computing	9	Book series
IEEE Internet of Things Journal	9	Journal
IEEE Access	8	Journal
Proceedings of Spie - The International Society for Optical Engineering	7	Conference Proceedings
IEEE Transactions on Industrial Informatics	6	Journal
IEEE Transactions on Intelligent Transportation Systems	6	Journal
Journal of Image and Graphics	6	Journal

Taken together, these results suggest that the dissemination of research combines mechanisms for the rapid dissemination of preliminary advances through collections and conferences with the consolidation of findings in internationally indexed journals, reflecting a field that is expanding and has high interdisciplinary potential.

4.3. Most Cited Articles

An analysis of the most cited articles highlights the thematic diversity and interdisciplinary breadth that research on intelligent learning applied to multidimensional data streams has achieved, as can be seen in Table 6. The most influential work, “Advancing Biosensors with Machine Learning” with 551 citations, integrates artificial intelligence and biomedicine, showing how machine learning methods, especially convolutional and recurrent neural networks, can enhance detection and analysis using electrochemical, fluorescent, and spectral biosensors, as well as the fusion of data from multiple sensors for more accurate diagnoses. This article highlights the expansion of chemometrics into intelligent and automated applications in the biomedical field.

Table 6. The ten most cited articles, arranged in descending order by number of citations. Source: Scopus.

Author (Year) and Title	Source	Citations
Cui F. (2020) Advancing Biosensors with Machine Learning [61].	ACS Sensors	551
Su Z. (2022) Secure and Efficient Federated Learning for Smart Grid With Edge-Cloud Collaboration [62].	IEEE Xplore	216
Wei X. (2022) A high-accuracy, real-time, intelligent material perception system with a machine-learning-motivated pressure-sensitive electronic skin [63].	Matter	197
Brenner B. (2017) Digital Twin as Enabler for an Innovative Digital Shopfloor Management System in the ESB Logistics Learning Factory at Reutlingen - University [64].	Procedia Manufacturing	190
Chen X. (2019) Missing traffic data imputation and pattern discovery with a Bayesian augmented tensor factorization model [65].	Transportation Research Part C: Emerging Technologies	145
Zhu J. (2022) Application of recurrent neural network to mechanical fault diagnosis: a review [66].	Springer Nature Link	144
Galan E. (2020) Intelligent Microfluidics: The Convergence of Machine Learning and Microfluidics in Materials Science and Biomedicine [67].	Matter	129
Boquet G. (2020) A variational autoencoder solution for road traffic forecasting systems: Missing data imputation, dimension reduction, model selection and anomaly detection [68].	Transportation Research Part C: Emerging Technologies	126
Moussalli S. (2020) Intelligent personal assistants: can they understand and be understood by accented L2 learners? [69].	Computer Assisted Language Learning	113
Princy J. (2020) Prediction of Cardiac Disease using Supervised Machine Learning Algorithms [70].	IEEE Xplore	99

In the field of energy and smart systems, “Secure and Efficient Federated Learning for Smart Grid With Edge-Cloud Collaboration” with 216 citations stands out for its federated learning proposal that allows artificial intelligence models to be shared without compromising the privacy of users’ energy data. Its approach combines edge computing, cloud computing, and reinforcement learning to optimize the quality of local models and communication efficiency, addressing non-IID data problems and user participation limitations. Complementarily, “A high-accuracy, real-time, intelligent material perception system” with 197 citations applies machine learning to intelligent perception using hybrid e-skins capable of recognizing multidimensional materials and stimuli in real time, opening up possibilities for physical interfaces and touch-sensitive robotics.

In the industrial and transportation sector, publications such as “Digital Twin as Enabler for an Innovative Digital Shopfloor Management System” with 190 citations and articles in Transportation Research Part C (145 and 126 citations) demonstrate the importance of intelligent learning in manufacturing, logistics, and transportation. This research highlights the management of digital twins and the imputation of missing data using Bayesian models or variational autoencoders to optimize traffic prediction, scenario simulation, and production and vehicle flow planning in complex and multidimensional environments.

Finally, other highly cited works demonstrate the breadth of applications of intelligent learning. For example, studies on mechanical fault diagnosis using recurrent neural networks and intelligent microfluidics illustrate its impact on advanced manufacturing and biomedicine. Likewise, research on intelligent personal assistants for language learning and heart disease prediction using supervised algorithms reflects its ability to improve human interactions and the analysis of large medical datasets, consolidating the relevance of these methodologies in practical and multidisciplinary environments.

The high number of citations of these articles confirms that they are fundamental references in their respective fields, either for their theoretical contributions or for their applicability in real-world scenarios. Furthermore, the breadth of areas involved, ranging from biomedicine and energy to transportation and education, highlights the cross-cutting nature of intelligent learning, consolidating it as a key driver for the development of innovative solutions in diverse domains.

Table 7 shows the most productive authors in the field of study. According to the results obtained, the most prominent authors are Wang Yaoze and Zhang Yushuang, who top the list with 21 and 20 articles, respectively. They are followed by Wang Xiuwen with 16 publications, and Li Xiuzheng and Li Yonghui, both with 15 papers. This pattern of productivity suggests the existence of researchers who act as key references within the field, consolidating stable lines of research and promoting the continuous generation of knowledge. In addition, the concentration of publications by these authors indicates that their scientific leadership influences the orientation of studies and the consolidation of collaborative networks in the area.

Table 7. Most productive authors. Source: Scopus.

Authors	Institution	Articles
Wang Yaoze	Kunming University of Science and Technology, Kunming, China	21
Zhang Yushuang	Beijing Polytechnic University, Beijing, China	20
Wang Xiuwen	Dalian Minzu University, Dalian, China	16
Li Xiuzheng	School of Economics and Management, China	15
Li Yonghui	Anhui Xinhua University, Hefei, China	15

The analysis of the most productive authors allows us to identify the main contributors in this area of research, as well as the institutions that have promoted greater scientific production. The presence of multiple prominent researchers at specific universities shows a concentration of scientific activity in these centers, which could be related to the existence of specialized research clusters that favor the development and advancement of the discipline.

Likewise, these results reflect trends in institutional and geographical collaboration, mainly in Chinese universities, suggesting that efforts in intelligent learning applied to multidimensional data

streams are being led by well-established and well-structured research groups capable of generating a significant impact on scientific production in the area.

4.4. Main Keywords

The analysis of the most frequent keywords, presented in Table 8, reveals the centrality of concepts associated with machine learning and its applications in multidimensional data processing. Among the author keywords, “deep learning” and “machine learning” stand out with 247 and 193 articles respectively, followed by “learning systems” with 136 articles, which shows the predominance of approaches based on advanced artificial intelligence techniques. Other notable terms include “artificial intelligence” (79), “machine learning” (71), “learning algorithms” (68), and “intelligent systems” (64), reflecting the consolidation of specialized vocabulary linked to the design of predictive and optimization models.

Table 8. Main keywords. Source: Scopus.

Author Keywords	Articles	Keywords-Plus	Articles
deep learning	247	deep learning	158
machine learning	193	learning systems	135
learning systems	136	machine learning	82
artificial intelligence	79	machine-learning	70
machine-learning	71	learning algorithms	68
learning algorithms	68	intelligent systems	63
intelligent systems	64	forecasting	50
data mining	62	data mining	47
forecasting	51	data handling	43
big data	45	multidimensional data	43

As for the automatically generated plus keywords, “deep learning” (158) and “learning systems” (135) appear repeatedly, along with “machine learning” (82), “machine-learning” (70), and “learning algorithms” (68). In addition, terms associated with specific applications appear, such as “intelligent systems” (63), “forecasting” (50), “data mining” (47), and “multidimensional data” (43), which underscore the importance of prediction, large-volume information management, and complex data analysis.

The overlap between both types of keywords indicates a strong thematic convergence, in which deep and automatic learning are positioned as the backbone of the research. This pattern confirms that the field is oriented towards the integration of intelligent algorithms with prediction and analysis capabilities in environments characterized by large scale, heterogeneity, and multidimensionality of data.

Additionally, the analysis suggests that research is not only focused on the development of learning models, but also on their practical application to specific problems, such as data mining, information management, and the prediction of complex phenomena. This trend reflects a dual approach: on the one hand, the refinement of algorithmic techniques and, on the other, their implementation in real environments, which demonstrates the maturity of the field and the relevance of its theoretical and applied contributions.

4.5. Keyword Strategy Diagram

The strategic keyword diagram allows us to identify the most relevant subject areas within the field of intelligent learning applied to multidimensional data streams, differentiating between those that are well established and those that are emerging or in decline. This analysis is based on two parameters: density, which reflects the degree of internal cohesion of each topic, and centrality, which indicates its level of connection and influence with other topics in the field [71]. Therefore, the strategic diagram distinguishes research topics according to their degree of development (density) and degree of

relevance (centrality). Figure 1 shows four quadrants that allow us to interpret the status and evolution of topics within the field.

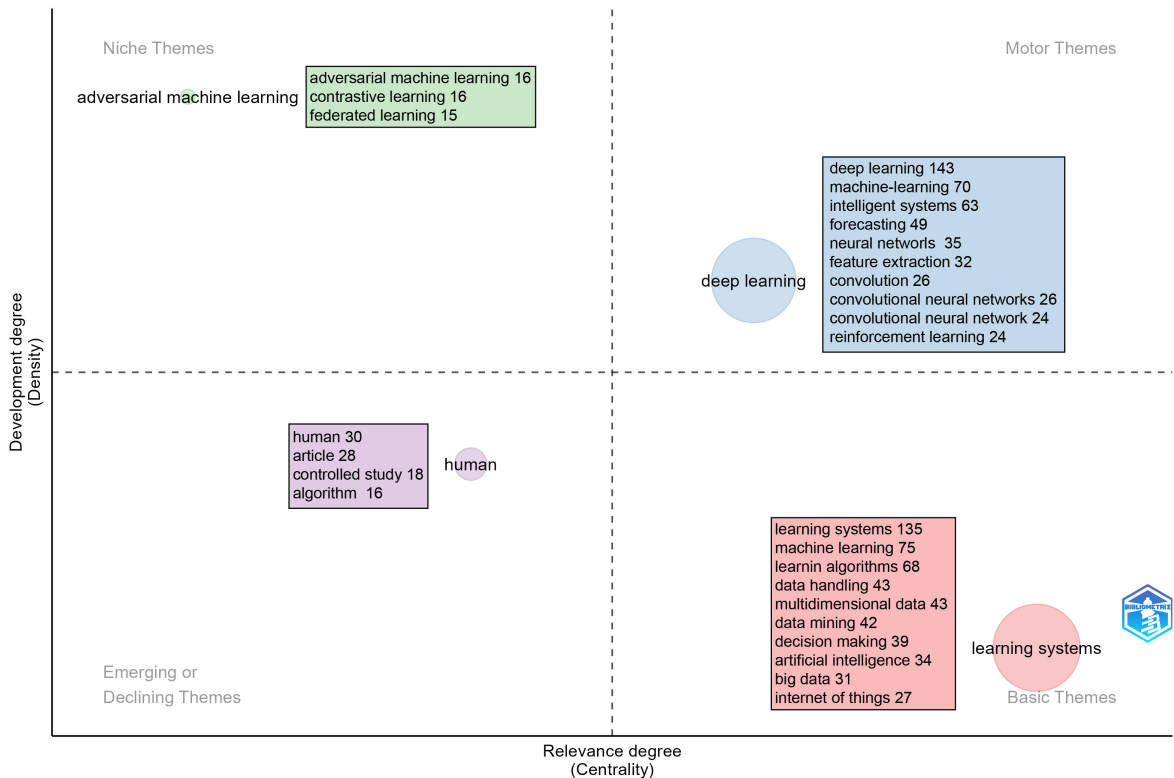


Figure 1. Strategic diagram of KeyWords Plus generated with bibliometrix.

In the upper right quadrant is “learning systems”, a well-developed and highly central topic. This indicates that it is a structuring axis of the area and also maintains strong links with other topics, becoming a driver of research. Its size in the figure confirms its weight within the field. The upper left quadrant contains the topic “adversarial machine learning”, a topic with high density but low centrality. Although it is well defined and developed internally, its contribution is more marginal within the research as a whole. This positions it as a specialized niche. The topic “human” is located in the lower left quadrant, with low density and low centrality. Its location indicates that it is an underdeveloped topic with little influence in the field, which can be interpreted as an emerging but still incipient line or a declining trend. Finally, in the lower right quadrant, we see “deep learning”, characterized by strong centrality but low density. This positioning makes it an essential topic that is widely connected to others, although it does not yet have such a consolidated internal development. It is a basic topic that supports other lines of research and has the potential to evolve into driving topics.

Overall, the diagram reflects that the field is structured around “learning systems” as a driving topic, while “deep learning” acts as a cross-cutting foundation. The topics “adversarial machine learning” and “human” show more specific trajectories: the former as a highly specialized line and the latter as a possible emerging trend.

Figure 2, which corresponds to the analysis of the authors’ keywords, shows patterns that complement and, in some cases, contrast with those identified in Figure 1. In the upper right quadrant, “learning systems” stand out. With their high density and centrality, they are consolidated as driving themes in the discipline, showing a well-cohesive structure and strong influence on the evolution of the field. In the upper left quadrant are “federated learning” and “adversarial machine learning”, specialized topics that, although highly developed internally, have less connection to the central cores of research, similar to what occurred in Figure 1 with the more peripheral topics.

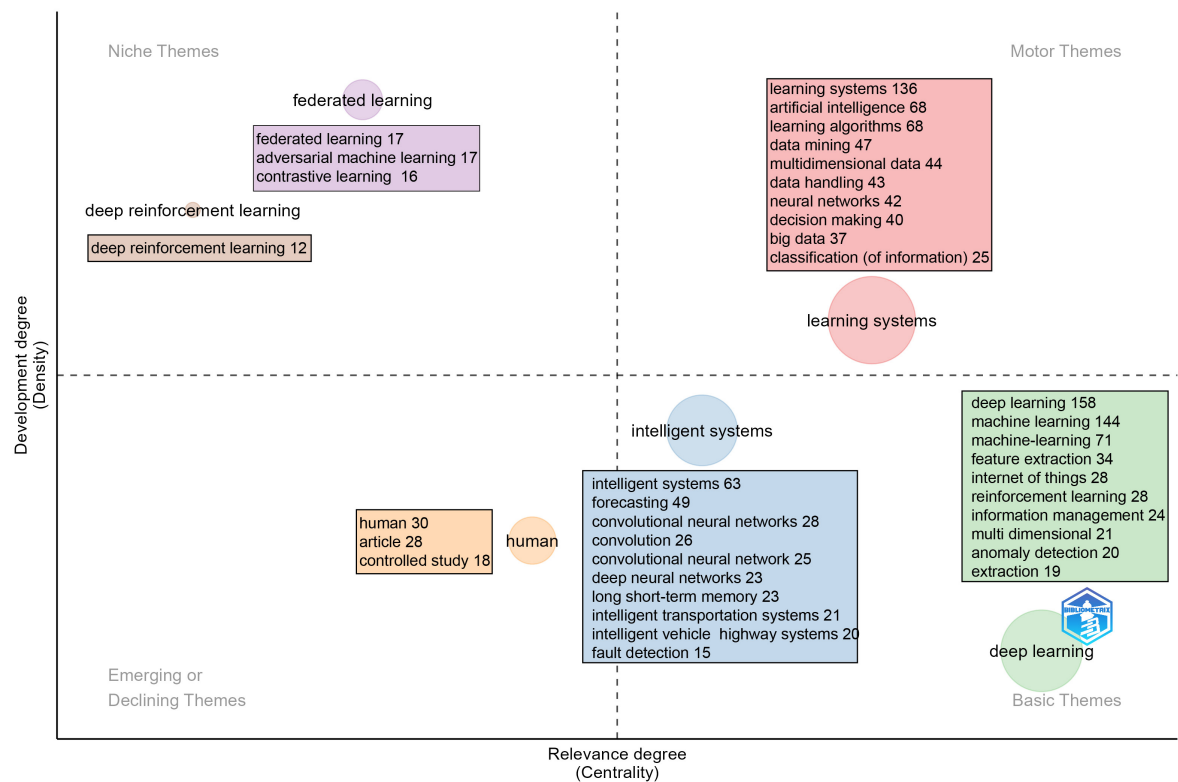


Figure 2. Strategic diagram of the authors' keywords generated with bibliometrix.

In contrast, the lower left quadrant contains terms such as “human” and “controlled study”, which reflect emerging or marginal areas with low density and little influence on the articulation of the main advances in the field. Finally, in the lower right quadrant are “deep learning” and “machine learning”, which, as in Figure 1, appear as highly central cross-cutting axes, although not yet fully developed, confirming their role as methodological and conceptual foundations that support other lines of research.

Comparatively, while Figure 1 placed “deep learning” as a driving theme, in Figure 2 it is repositioned as a basic theme, suggesting that from the authors’ perspective it constitutes a fundamental and broadly transversal axis, rather than a core of specialization in itself. Thus, both figures together allow us to appreciate the coexistence of consolidated approaches, specialized areas, and emerging topics, highlighting the diversity and dynamism of the field of intelligent learning applied to multidimensional data streams.

This diagram shows how the literature has structured the thematic contributions of the field, revealing a balance between established, specialized, and emerging areas. The prominence of approaches such as deep learning and intelligent systems, together with the emergence of specialized topics such as federated learning, reflects the trend toward the integration of advanced and collaborative techniques to address the challenges of managing large volumes of multidimensional data.

4.6. Thematic Evolution of Keywords

For the thematic analysis of keyword evolution, we used the Bibliometrix package with its Biblioshiny graphical interface, establishing ranges of years that allow us to observe changes in the dynamics of the topics. Based on the keyword plus, Figure 3 shows how the topics have evolved between the periods 2015–2020 and 2021–2025.

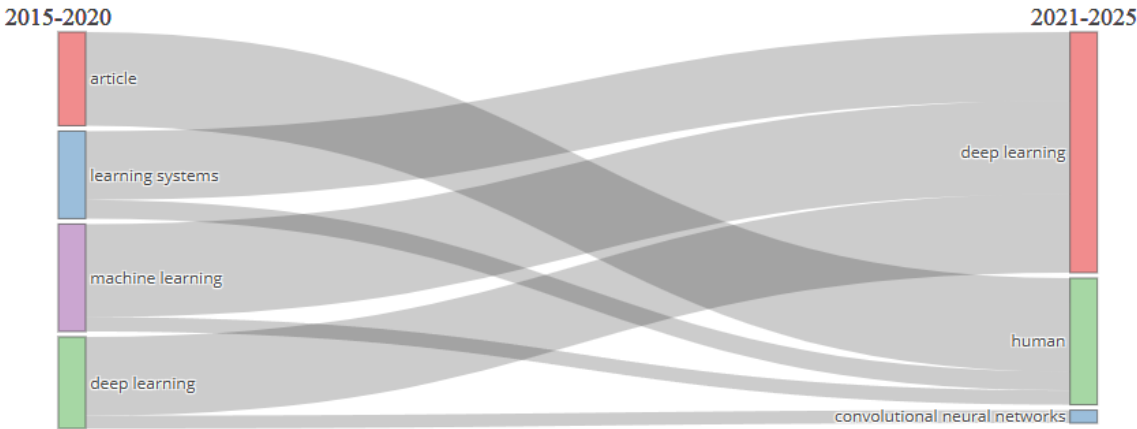


Figure 3. Thematic evolution of the authors’ keyWord Plus generated with bibliometrix.

In the first period, the main topics were “deep learning”, “machine learning”, “learning systems”, and “article”, which formed the conceptual basis of the initial studies. In the second period, some of these topics were transformed or integrated into new orientations. For example, “learning systems” remains a central theme, consolidating itself as a persistent and expanding topic within the topic of “deep learning”.

Likewise, new topics such as “human” and “convolutional neural networks” are emerging, reflecting both a shift towards human-machine interaction and the incorporation of specific architectures in the field of machine learning. On the other hand, “deep learning” retains its relevance, albeit with a more limited scope, and “machine learning” tends to be redistributed towards these new areas of research.

Overall, the thematic evolution analysis reveals a shift from more general notions of artificial intelligence and learning toward more specific and applied approaches, which demonstrates the maturation of the field and its progressive diversification into emerging lines of research.

Figure 4 shows the thematic evolution of author keywords between the periods 2015–2020 and 2021–2025. In the first stage, as with keyword plus, general topics such as “article”, “deep learning”, “learning systems” and “machine learning” predominate, reflecting the initial interest in laying the conceptual and methodological foundations for research in the field.

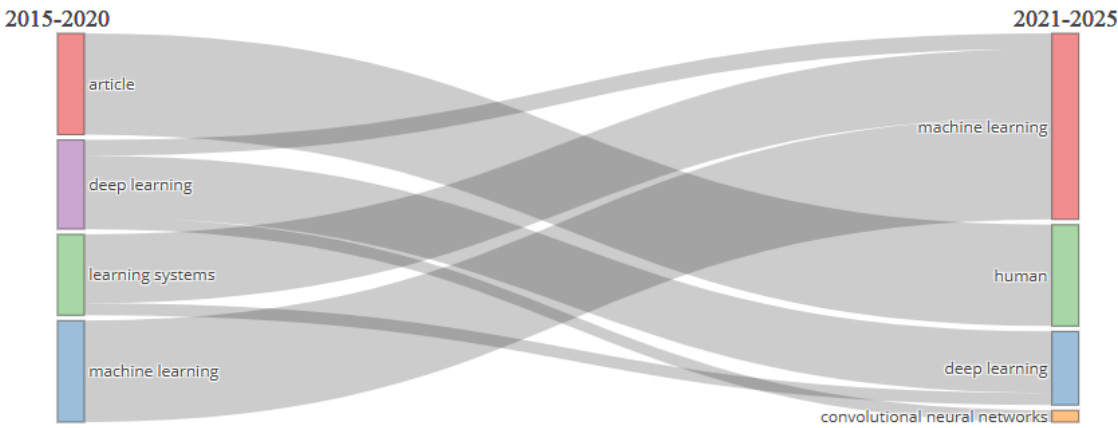


Figure 4. Thematic evolution of the authors’ keyWords generated with bibliometrix.

In the second period, there is a shift towards more specialized and applied areas. Machine learning and deep learning remain central, although integrated with new lines of research. At the same time, topics such as “human” emerge, reflecting the growing attention to human-machine interaction, and

“convolutional neural networks”, denoting the incorporation of specific architectures in the analysis of complex data. Notably, “deep learning” remains a topic with an independent evolution, which provides opportunities for expansion toward more autonomous and adaptive learning approaches.

This thematic shift suggests a process of maturation in the field, in which general concepts of machine learning have evolved toward more robust and specialized models. The diversification of topics highlights methodological sophistication and a focus on more practical applications, in line with advances in artificial intelligence and large-scale data processing.

4.7. Degree of Concentration of Selected Variables

To evaluate the distribution of scientific output in this field, the degree of concentration of different bibliometric variables was analyzed using Shannon entropy, a fundamental measure in information theory [72]. This metric quantifies the uncertainty or diversity in a data distribution: values close to 1 indicate that the elements are evenly distributed, while values close to 0 reflect a high concentration in a few elements. Mathematically, for a discrete probability distribution $P = \{p_j; j = 1, \dots, N\}$ that satisfies $\sum_{j=1}^N x_i p_j = 1$, entropy is defined in Equation 1.

$$S[P] = - \sum_{j=1}^N p_j \ln(p_j) \tag{1}$$

In bibliometrics, entropy is used to study the distribution of equity or concentration of relevant variables, such as research topics and authors. To facilitate interpretation, normalized Shannon entropy is used, defined in Equation 2.

$$H[P] = \frac{S[P]}{S_{MAX}} = \frac{- \sum_{j=1}^N p_j \ln(p_j)}{\ln N} \tag{2}$$

where $0 \leq H \leq 1$, with $H = 1$ indicating a uniform distribution, i.e., without concentration, and $H = 0$ when the entire distribution is concentrated at a single point.

The normalized entropic concentration index was calculated for the distribution of authors, sources, countries, research areas, and article citations, with the results presented in Table 9. It can be seen that authors show a highly homogeneous distribution with $H = 0.9539$, as shown in the Table 7, as do sources with $H = 0.9359$ shown in the Table 5. In contrast, countries show a high concentration with $H = 0.4950$, based on Table 3, indicating that scientific production is geographically concentrated in a few countries. However, the distribution of authors within these countries is balanced, suggesting that individual contributions are not dominated by a small group. Research areas show moderate concentration with $H = 0.7188$ this can be seen in the Table 1), with a predominance of disciplines such as computer science, engineering, decision sciences, and mathematics, which account for 73.28% of publications. Finally, article citations show moderate concentration ($H=0.8154$, Table 6), indicating the existence of multiple influential articles without centralization in a few works.

Table 9. Entropic concentration index (H) of the selected variables. Source: Scopus.

Variable	H
Authors	0.9539
Sources	0.9359
Countries	0.4950
Areas of research	0.7188
Article citations	0.8154

These results reflect that scientific production in the area is heterogeneous in terms of authors and sources, which points to a field open to new contributions and collaborations. In geographical terms, although some countries account for a large part of the production, the participation of authors within these countries is balanced. In terms of research areas and the influence of articles, it is evident

that certain key disciplines and publications are more relevant, but without generating excessive dominance, suggesting a diverse and dynamic scientific ecosystem.

Another way to examine how authors are distributed according to their productivity is through Lotka’s law. According to Lotka’s empirical finding [73], this law states that the number of authors who publish n articles follows a relationship similar to Zipf’s law. Originally, Lotka analyzed a database limited to physics and chemistry, and the law is expressed in Equation 3.

$$a_n = \frac{a_1}{n^2}, n = 1, 2, \dots, N \tag{3}$$

where a_n represents the number of authors who publish n articles and a_1 represents the number of authors who publish a single article. To generalize this relationship and better adjust it to other fields, it can be expressed as Equation 4.

$$a_n = \frac{a_1}{n^c}, n = 1, 2, \dots, N \tag{4}$$

where c is a parameter estimated to optimize the fit to the observed data. In this study, the value of $c = 2.52$ and $R^2 = 0.96$, indicating a very accurate fit.

Table 10 presents the observed distribution of authors according to the number of articles published, together with the frequency adjusted according to Lotka’s law. It can be seen that most authors publish only one article (1287 authors, 81% of the total), while authors with two or more publications represent much smaller proportions. In some cases, such as authors with seven articles, the observed frequency is slightly higher than predicted, reflecting the presence of particularly productive researchers.

Table 10. TObserve d distribution of the number of authors who wrote a given number of articles and adjusted values of Lotka’s law. Source: Scopus.

Number of Articles	Authors	Observed Frequency	Adjusted Frequency
1	1287	0.8100	0.8110
2	154	0.0970	0.0970
3	60	0.0380	0.0378
4	38	0.0240	0.0239
5	12	0.0080	0.0076
6	5	0.0030	0.0032
7	7	0.0040	0.0044
8	4	0.0030	0.0025
9	3	0.0020	0.0019
10	3	0.0020	0.0019
12	2	0.0010	0.0013
13	4	0.0030	0.0025
14	2	0.0010	0.0013
15	2	0.0010	0.0013
16	2	0.0010	0.0013
20	1	0.0010	0.0006
21	1	0.0010	0.0006
53	1	0.0010	0.0006

Overall, the results confirm that authorship is not evenly distributed: a small group of authors contribute multiple publications, while the majority publish only once. This concentration of productivity is consistent with what Lotka’s law predicts and evidence of the existence of a small core of highly active researchers within the field of study, which could influence the direction and impact of research.

4.8. Charts of Citations, Sources, and Authors

The visualization presented in Figure 5 was generated using VOSviewer software, a tool specialized in bibliometric analysis and the graphical representation of relationships between terms. This application, developed by Van Eck and Waltman (2010), allows the identification of co-occurrence patterns based on keywords extracted from titles, abstracts, and descriptors of scientific publications [60].

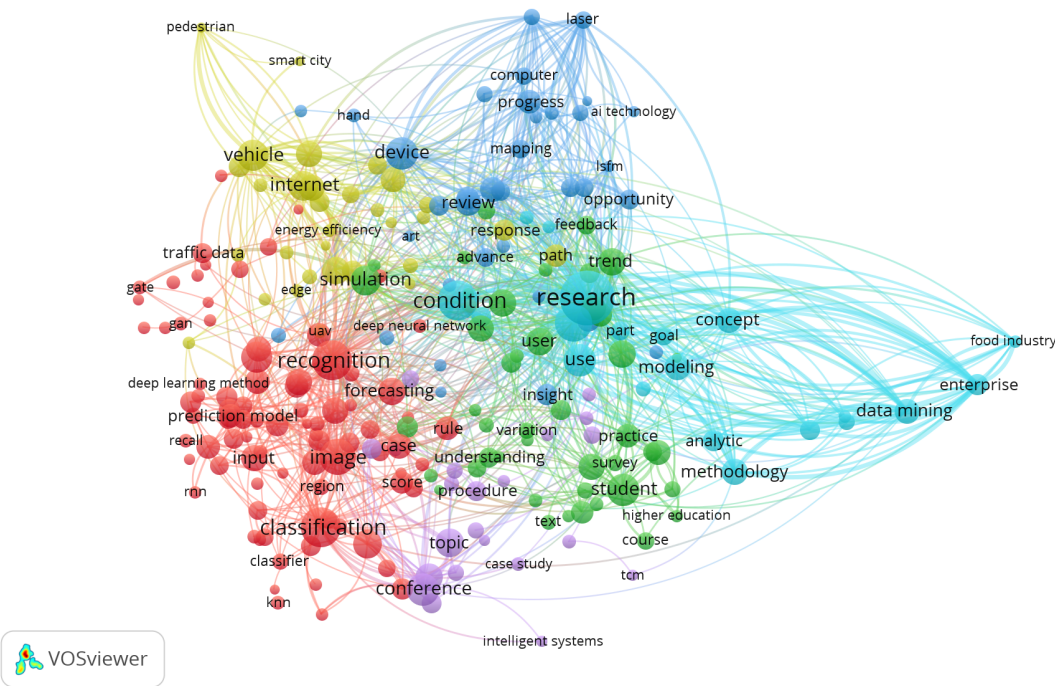
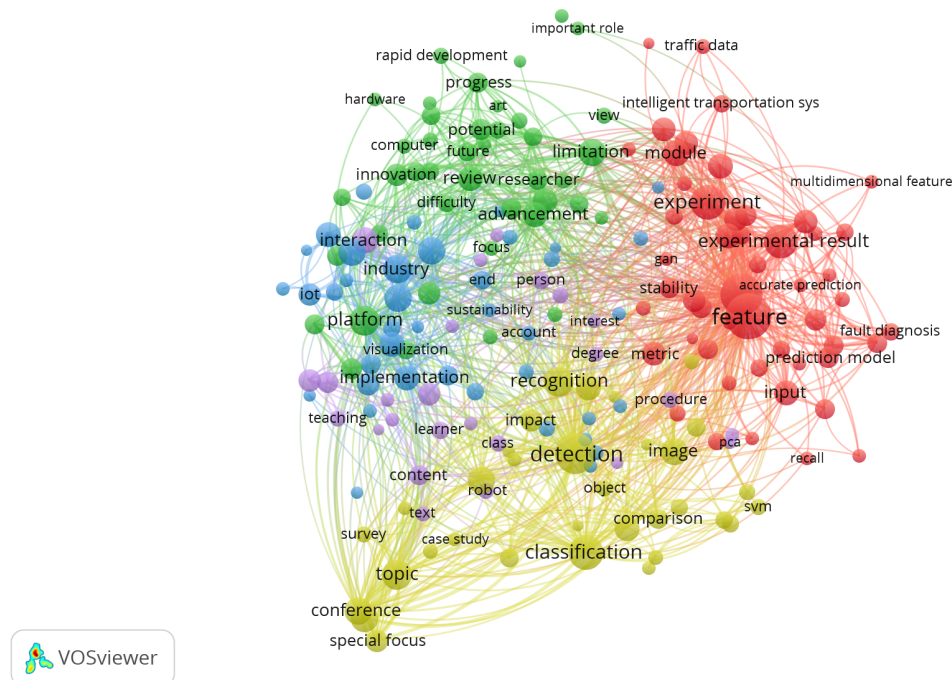


Figure 5. Map of word clouds in titles and abstracts (full count), generated with VOSviewer.

The figure shows a density and semantic connection map that groups key concepts related to the field of intelligent learning on multidimensional data streams. Each color in the diagram represents a thematic cluster, that is, a subset of terms that share high levels of co-occurrence within the analyzed documents.

The light blue cluster focuses on terms such as research, use, modeling, concept, and data mining. This group represents an orientation toward the development of analytical models and data mining techniques applied in organizational contexts and complex systems. In the green cluster, words such as simulation, user, practice, student, course, and higher education stand out, suggesting a focus on designing user-centered (student-centered) educational experiences to improve teaching and learning processes. The red cluster groups terms such as recognition, classification, image, prediction model, and input. This segment is clearly linked to visual data processing, pattern recognition, and prediction using machine learning models. The yellow cluster contains words such as internet, vehicle, smart city, and energy efficiency. This group points to technological applications in the field of the Internet of Things (IoT), smart vehicles, and energy efficiency. The dark blue cluster includes terms such as computer, progress, AI technology, and mapping, which refer to advances in emerging technologies in artificial intelligence and digital cartography. Finally, the purple cluster is composed of terms such as conference, topic, case study, and intelligent systems, reflecting a focus on academic production and the dissemination of knowledge in scientific forums.

Central terms such as research, recognition, data mining, classification, and modeling function as bridges between the different clusters, highlighting their integrative role in the development of research. This type of analysis allows us to identify key areas of interest, as well as possible future lines of exploration in the field of intelligent learning about data streams.



In this new co-occurrence map, significant changes can be observed in the structure of thematic clusters. One of the most notable aspects is the merging of the yellow cluster with key terms such as classification, topic, recognition, and detection. This grouping suggests a more focused approach to the analysis of classification and detection strategies applied to intelligent systems and machine learning, marking a difference from the previous map.

For their part, the green and blue clusters maintain their relevance, albeit with slight changes in their connections. The green cluster groups terms such as progress, innovation, limitation, view, and advancement, reflecting a critical and forward-looking view of technological advances in the field. The blue cluster, on the other hand, includes words such as platform, industry, IoT, and interaction, suggesting the growing integration of these technologies in industrial environments and interconnected systems.

This type of visualization allows for clearer detection of the thematic diversity of the field of study, as well as the interconnections between technical, methodological, and applied areas. In addition, it provides a useful tool for identifying emerging trends and possible future lines of research.

Figure 7 presents a cloud map of publication sources, distinguishing the main academic media that have contributed to the dissemination of research on algorithms, trajectory clustering methods, GPS trajectories, urban planning, and traffic. As also shown in Table 5, the source with the highest number of contributions is Lecture Notes in Computer Science, with 23 articles, consolidating its position as the most representative space for the dissemination of work in this area. It is followed by Communications in Computer and Information Science and Smart Innovation, Systems and Technologies, with 15 and 11 publications respectively, as well as Advances in Intelligent Systems and Computing and the IEEE Internet of Things Journal, with 9 articles each.

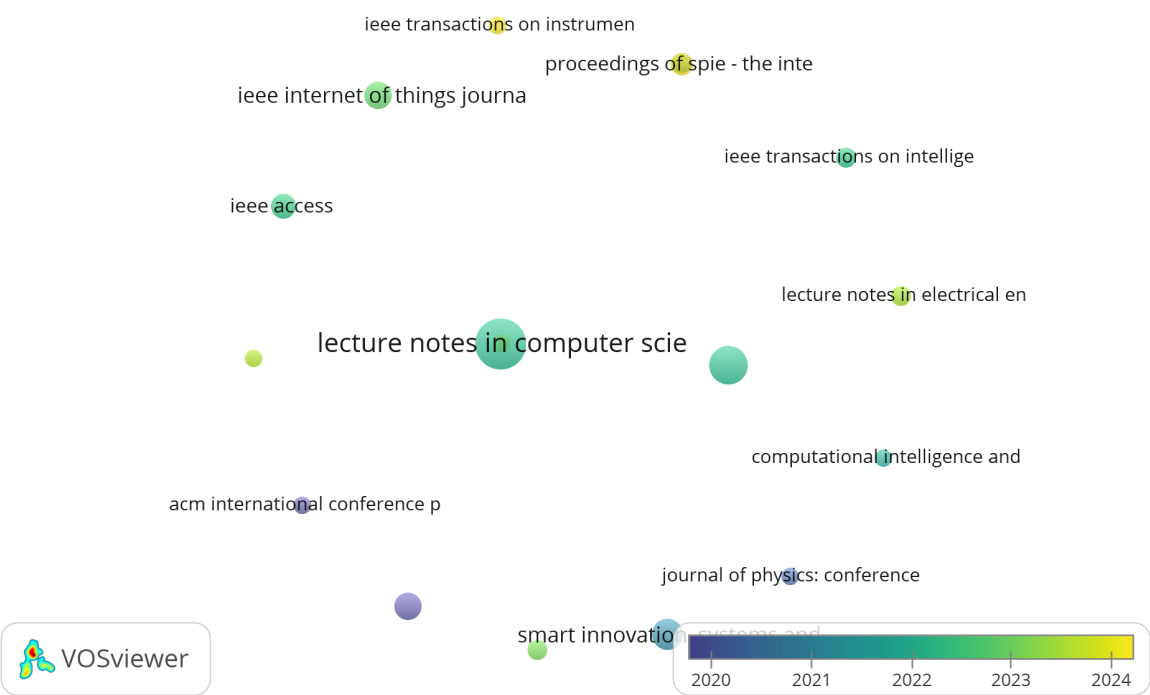


Figure 7. Cloud map of journals where articles on "Intelligent Learning on Multidimensional Data Streams" are published, generated with VOSviewer.

The map allows us to visualize the frequency of appearance of each source, reflected in the size of the nodes, and its evolution over time using a color scale. We can see that journals and conferences such as IEEE Internet of Things Journal and IEEE Access have had a more recent participation, while series such as Lecture Notes in Computer Science show sustained continuity over time, reaffirming their central role in the dissemination of knowledge. Together, these sources demonstrate the diversity of publication spaces, ranging from specialized conferences to high-impact journals in artificial intelligence and the Internet of Things, confirming the interdisciplinary and dynamic nature of the field.

Figure 8 presents the cloud map constructed from the most cited articles listed in Table 6, where the size of each node depends on the number of citations received. This visualization, similar to the previous figure, graphically highlights the influence of certain works on the development of the topic Intelligent Learning on Multidimensional Data Streams.

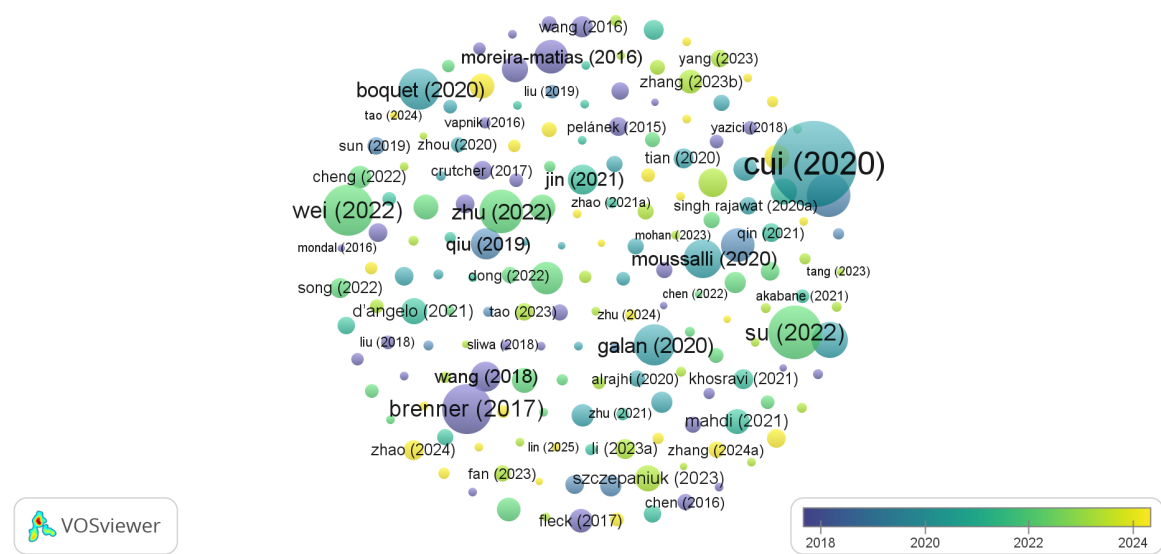


Figure 8. Cloud map created from authors with journal papers on "Intelligent Learning on Multidimensional Data Streams", generated with VOSviewer.

The most prominent node corresponds to the article by Cui et al. (2020), published in ACS Sensors, which, with 551 citations, leads the most influential production in applying machine learning techniques to the advancement of biosensors [61]. It is followed by Su et al. (2022), with a proposal for federated learning for smart grids [62], and Wei et al. (2022), who introduces a real-time material perception system based on machine learning [63]. These results reflect a shift towards concrete applications in highly relevant domains such as health, energy, and smart materials.

In contrast to the previous figure—where the most cited articles focused on vehicle trajectories and precision mapping—the current visualization shows greater thematic diversity. Works such as Brenner and Hummel (2017), on Digital Twin applied to smart factory management [64], and Zhu et al. (2022), focused on traffic data imputation using Bayesian tensor factorization, consolidate this breadth of approaches [66]. In this way, the most cited articles not only mark trends of impact in different disciplines, but also confirm the cross-cutting nature of intelligent learning applied to multidimensional data streams.

5. Conclusions

The bibliometric analysis carried out made it possible to characterize the current state and evolution of scientific production on intelligent learning applied to multidimensional data streams during the period 2015–2025. The results show sustained and accelerated growth in the literature, with an annual increase of 39.38%, which shows that this is a rapidly expanding field. The volume of publications, the diversity of sources, and the consolidation of thematic lines reflect the progressive maturity of this area, which combines both methodological developments and practical applications in various domains.

In disciplinary terms, the concentration in computer science and engineering confirms the leading role of these areas in the formulation of algorithms and models applied to the processing of spatial and multidimensional data. However, the growing participation of mathematics, physical sciences, and decision sciences reveals an interdisciplinary openness, in which theoretical foundations, mathematical modeling, and applications oriented toward strategic planning and decision-making converge. This diversity suggests fertile ground for research that integrates complementary perspectives, particularly in emerging fields such as mobility, logistics, and complex systems management.

Geographically, the results show an uneven picture: China leads in terms of quantity of production, while countries such as the United States, Canada, and Germany stand out for their impact as measured

by citations, reflecting different contribution profiles. Likewise, international cooperation is still limited, although cases such as Canada and Australia show a higher degree of collaboration. This finding points to the need to strengthen international research networks, especially in emerging regions, to promote greater diversity and global visibility.

A review of the most cited articles revealed the cross-cutting nature of intelligent learning applied to multidimensional data streams, with applications in biomedicine, energy, manufacturing, transportation, and human-machine interfaces. This research not only constitutes key references in theoretical terms, but also demonstrates a high degree of applicability in real-world scenarios. The variety of topics identified confirms that the field is establishing itself as a catalyst for innovative solutions in different scientific and technological domains.

Analysis of keywords and their thematic evolution shows that research has shifted from general approaches, such as deep learning and machine learning, to more specific and applied areas, such as federated learning, convolutional neural networks, and topics related to human-machine interaction. This transition reflects a process of diversification and maturation in the field, in which established topics, specialized areas, and emerging lines of research coexist. Overall, the results allow us to conclude that intelligent learning on multidimensional data streams is an area of growing academic and practical relevance, with high potential for interdisciplinary impact and multiple open challenges in terms of quality, international collaboration, and consolidation of new areas of research.

Future research in intelligent learning on multidimensional data streams should focus on the design of more efficient and adaptive algorithms capable of processing heterogeneous data in real time. In addition, it is essential to strengthen interdisciplinary and international collaboration to broaden the impact of findings and diversify applications in areas such as mobility, biomedicine, and complex systems management. Finally, progress is needed on ethical and sustainability issues, including privacy protection and the energy efficiency of models.

Author Contributions: Conceptualization, G.R.; methodology, G.R.; validation, L.L. and W.H.; formal analysis, R.T.-B. and J.B.-M.; investigation, G.R.; data curation, G.R.; writing—original draft preparation, G.R.; writing—review and editing, R.T.-B., L.L., W.H. and J.B.-M.; supervision, R.T.-B. and J.B.-M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Acknowledgments: During the preparation of this manuscript, the authors used ChatGPT (GPT-5, OpenAI) for the purposes of language editing, text organization, and improvement of clarity and readability. The authors have reviewed and edited the output and take full responsibility for the content of this publication.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Giannini, F.; Ziffer, G.; Cossu, A.; Lomonaco, V. Streaming Continual Learning for Unified Adaptive Intelligence in Dynamic Environments. *IEEE Intelligent Systems* **2024**, *39*, 81–85. <https://doi.org/10.1109/MIS.2024.3479469>.
2. Omranpour, S.; Rabusseau, G.; Rabbany, R. Higher Order Transformers: Efficient Attention Mechanism for Tensor Structured Data, 2024. <https://doi.org/10.48550/ARXIV.2412.02919>.
3. Picón, G.C.; Oleksienko, I.; Hedegaard, L.; Bakhtiarnia, A.; Iosifidis, A. Continual Low-Rank Scaled Dot-product Attention, 2024. <https://doi.org/10.48550/ARXIV.2412.03214>.
4. Qiu, R.; Jang, J.G.; Lin, X.; Liu, L.; Tong, H. TUCKET: A Tensor Time Series Data Structure for Efficient and Accurate Factor Analysis over Time Ranges. *Proceedings of the VLDB Endowment* **2024**, *17*, 4746–4759. <https://doi.org/10.14778/3704965.3704980>.
5. Lanzarini, L.C.; Hasperué, W.; Villa Monte, A.; Jimbo Santana, P.; Reyes Zambrano, G.; Corvi, J.P.; Fernández Bariviera, A.; Olivas Varela, J.Á. Minería de Datos, Minería de Textos y Big Data. In Proceedings of the XXI Workshop de Investigadores En Ciencias de La Computación (WICC 2019, Universidad Nacional de San Juan), 2019.

6. Shu, H.; Li, J.; Jin, Y.; Wang, H. Guaranteed Multidimensional Time Series Prediction via Deterministic Tensor Completion Theory, 2025. <https://doi.org/10.48550/ARXIV.2501.15388>.
7. Hou, Y.; Tang, P. Multi-Head Self-Attending Neural Tucker Factorization, 2025. <https://doi.org/10.48550/ARXIV.2501.09776>.
8. Chen, Z.; He, Y.; Wu, D.; Zuo, L.; Li, K.; Zhang, W.; Deng, Z. $\ell_{1,2}$ -Norm and CUR Decomposition Based Sparse Online Active Learning for Data Streams with Streaming Features. In Proceedings of the 2024 IEEE International Conference on Big Data (BigData), Washington, DC, USA, 12 2024; pp. 384–393. <https://doi.org/10.1109/BigData62323.2024.10825278>.
9. Gary Reyes Zambrano.; Lissette Ochoa Vera. Reference Architecture for an Intelligent Transportation System. *International Journal of Innovation and Applied Studies* **2016**, 15, 175–182.
10. Mutambik, I. An Entropy-Based Clustering Algorithm for Real-Time High-Dimensional IoT Data Streams. *Sensors* **2024**, 24, 7412. <https://doi.org/10.3390/s24227412>.
11. Chen, F.; Wu, D.; Yang, J.; He, Y. An Online Sparse Streaming Feature Selection Algorithm, 2022. <https://doi.org/10.48550/ARXIV.2208.01562>.
12. Wang, A.; Yang, H.; Mao, F.; Zhang, Z.; Yu, Y.; Liu, X. Anti-Drifting Feature Selection via Deep Reinforcement Learning (Student Abstract). *Proceedings of the AAAI Conference on Artificial Intelligence* **2023**, 37, 16356–16357. <https://doi.org/10.1609/aaai.v37i13.27038>.
13. Lu, C.; Shi, L.; Chen, Z.; Wu, C.; Wierman, A. Overcoming the Curse of Dimensionality in Reinforcement Learning Through Approximate Factorization, 2024. <https://doi.org/10.48550/ARXIV.2411.07591>.
14. Yuan, Z.; Sun, Y.; Shasha, D. Forgetful Forests: High Performance Learning Data Structures for Streaming Data under Concept Drift, 2022. <https://doi.org/10.48550/ARXIV.2212.07876>.
15. Laila Ab Ghani, N.; Abdul Aziz, I.; Jadid AbdulKadir, S. Subspace Clustering in High-Dimensional Data Streams: A Systematic Literature Review. *Computers, Materials & Continua* **2023**, 75, 4649–4668. <https://doi.org/10.32604/cmc.2023.035987>.
16. Heusinger, M.; Schleif, F.M. Random Projection in the Presence of Concept Drift in Supervised Environments. In *Artificial Intelligence and Soft Computing*; Rutkowski, L.; Scherer, R.; Korytkowski, M.; Pedrycz, W.; Tadeusiewicz, R.; Zurada, J.M., Eds.; Springer International Publishing: Cham, 2020; Vol. 12415, pp. 514–524. https://doi.org/10.1007/978-3-030-61401-0_48.
17. Gama, J.; Žliobaitė, I.; Bifet, A.; Pechenizkiy, M.; Bouchachia, A. A Survey on Concept Drift Adaptation. *ACM Computing Surveys* **2014**, 46, 1–37. <https://doi.org/10.1145/2523813>.
18. Žliobaitė, I.; Pechenizkiy, M.; Gama, J. An Overview of Concept Drift Applications. In *Big Data Analysis: New Algorithms for a New Society*; Japkowicz, N.; Stefanowski, J., Eds.; Springer International Publishing: Cham, 2016; Vol. 16, pp. 91–114. https://doi.org/10.1007/978-3-319-26989-4_4.
19. Ditzler, G.; Roveri, M.; Alippi, C.; Polikar, R. Learning in Nonstationary Environments: A Survey. *IEEE Computational Intelligence Magazine* **2015**, 10, 12–25. <https://doi.org/10.1109/MCI.2015.2471196>.
20. Gaber, M.M.; Zaslavsky, A.; Krishnaswamy, S. Mining Data Streams: A Review. *ACM SIGMOD Record* **2005**, 34, 18–26. <https://doi.org/10.1145/1083784.1083789>.
21. Sheth, A.; Henson, C.; Sahoo, S.S. Semantic Sensor Web. *IEEE Internet Computing* **2008**, 12, 78–83. <https://doi.org/10.1109/MIC.2008.87>.
22. Atzori, L.; Iera, A.; Morabito, G. The Internet of Things: A Survey. *Computer Networks* **2010**, 54, 2787–2805. <https://doi.org/10.1016/j.comnet.2010.05.010>.
23. Gubbi, J.; Buyya, R.; Marusic, S.; Palaniswami, M. Internet of Things (IoT): A Vision, Architectural Elements, and Future Directions. *Future Generation Computer Systems* **2013**, 29, 1645–1660. <https://doi.org/10.1016/j.future.2013.01.010>.
24. Taylor, S.J.; Letham, B. Forecasting at Scale. *The American Statistician* **2018**, 72, 37–45. <https://doi.org/10.1080/00031305.2017.1380080>.
25. Kwak, H.; Lee, C.; Park, H.; Moon, S. What Is Twitter, a Social Network or a News Media? In Proceedings of the Proceedings of the 19th International Conference on World Wide Web, Raleigh North Carolina USA, 04 2010; pp. 591–600. <https://doi.org/10.1145/1772690.1772751>.
26. Boyd, D.M.; Ellison, N.B. Social Network Sites: Definition, History, and Scholarship. *Journal of Computer-Mediated Communication* **2007**, 13, 210–230. <https://doi.org/10.1111/j.1083-6101.2007.00393.x>.
27. Pantelopoulos, A.; Bourbakis, N. A Survey on Wearable Sensor-Based Systems for Health Monitoring and Prognosis. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* **2010**, 40, 1–12. <https://doi.org/10.1109/TSMCC.2009.2032660>.

28. Patel, S.; Park, H.; Bonato, P.; Chan, L.; Rodgers, M. A Review of Wearable Sensors and Systems with Application in Rehabilitation. *Journal of NeuroEngineering and Rehabilitation* **2012**, *9*, 21. <https://doi.org/10.1186/1743-0003-9-21>.
29. Domingos, P.; Hulten, G. Mining High-Speed Data Streams. In Proceedings of the Proceedings of the Sixth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Boston Massachusetts USA, 08 2000; pp. 71–80. <https://doi.org/10.1145/347090.347107>.
30. Gama, J.; Medas, P.; Castillo, G.; Rodrigues, P. Learning with Drift Detection. In *Advances in Artificial Intelligence – SBIA 2004*; Hutchison, D.; Kanade, T.; Kittler, J.; Kleinberg, J.M.; Mattern, F.; Mitchell, J.C.; Naor, M.; Nierstrasz, O.; Pandu Rangan, C.; Steffen, B.; et al., Eds.; Springer Berlin Heidelberg: Berlin, Heidelberg, 2004; Vol. 3171, pp. 286–295. https://doi.org/10.1007/978-3-540-28645-5_29.
31. Babcock, B.; Babu, S.; Datar, M.; Motwani, R.; Widom, J. Models and Issues in Data Stream Systems. In Proceedings of the Proceedings of the Twenty-First ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems, Madison Wisconsin, 06 2002; pp. 1–16. <https://doi.org/10.1145/543613.543615>.
32. Bifet, A.; Gavaldà, R. Learning from Time-Changing Data with Adaptive Windowing. In Proceedings of the Proceedings of the 2007 SIAM International Conference on Data Mining, Society for Industrial and Applied Mathematics, 04 2007, pp. 443–448. <https://doi.org/10.1137/1.9781611972771.42>.
33. Widmer, G.; Kubat, M. Learning in the Presence of Concept Drift and Hidden Contexts. *Machine Learning* **1996**, *23*, 69–101. <https://doi.org/10.1007/BF00116900>.
34. Oza, N. Online Bagging and Boosting. In Proceedings of the 2005 IEEE International Conference on Systems, Man and Cybernetics, Waikoloa, HI, USA, 2005; Vol. 3, pp. 2340–2345. <https://doi.org/10.1109/ICSMC.2005.1571498>.
35. Gomes, H.M.; Bifet, A.; Read, J.; Barddal, J.P.; Enembreck, F.; Pfharinger, B.; Holmes, G.; Abdessalem, T. Adaptive Random Forests for Evolving Data Stream Classification. *Machine Learning* **2017**, *106*, 1469–1495. <https://doi.org/10.1007/s10994-017-5642-8>.
36. Krawczyk, B. Learning from Imbalanced Data: Open Challenges and Future Directions. *Progress in Artificial Intelligence* **2016**, *5*, 221–232. <https://doi.org/10.1007/s13748-016-0094-0>.
37. Zhang, P.; Zhu, X.; Tan, J.; Guo, L. Classifier and Cluster Ensembles for Mining Concept Drifting Data Streams. In Proceedings of the 2010 IEEE International Conference on Data Mining, Sydney, Australia, 12 2010; pp. 1175–1180. <https://doi.org/10.1109/ICDM.2010.125>.
38. Katakis, I.; Tsoumakas, G.; Vlahavas, I. Tracking Recurring Contexts Using Ensemble Classifiers: An Application to Email Filtering. *Knowledge and Information Systems* **2010**, *22*, 371–391. <https://doi.org/10.1007/s10115-009-0206-2>.
39. Aggarwal, C.C.; Yu, P.S. Online Analysis of Community Evolution in Data Streams. In Proceedings of the Proceedings of the 2005 SIAM International Conference on Data Mining, Society for Industrial and Applied Mathematics, 04 2005, pp. 56–67. <https://doi.org/10.1137/1.9781611972757.6>.
40. Zhao, H.; Yuen, P.C. Incremental Linear Discriminant Analysis for Face Recognition. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* **2008**, *38*, 210–221. <https://doi.org/10.1109/TSMCB.2007.908870>.
41. Hochreiter, S.; Schmidhuber, J. Long Short-Term Memory. *Neural Computation* **1997**, *9*, 1735–1780. <https://doi.org/10.1162/neco.1997.9.8.1735>.
42. Reyes, G.; Lanzarini, L.; Hasperué, W.; Bariviera, A.F. GPS Trajectory Clustering Method for Decision Making on Intelligent Transportation Systems. *Journal of Intelligent & Fuzzy Systems* **2020**, *38*, 5529–5535. <https://doi.org/10.3233/JIFS-179644>.
43. McCloskey, M.; Cohen, N.J. Catastrophic Interference in Connectionist Networks: The Sequential Learning Problem. In *Psychology of Learning and Motivation*; Elsevier, 1989; Vol. 24, pp. 109–165. [https://doi.org/10.1016/S0079-7421\(08\)60536-8](https://doi.org/10.1016/S0079-7421(08)60536-8).
44. Kirkpatrick, J.; Pascanu, R.; Rabinowitz, N.; Veness, J.; Desjardins, G.; Rusu, A.A.; Milan, K.; Quan, J.; Ramalho, T.; Grabska-Barwinska, A.; et al. Overcoming Catastrophic Forgetting in Neural Networks. *Proceedings of the National Academy of Sciences* **2017**, *114*, 3521–3526. <https://doi.org/10.1073/pnas.1611835114>.
45. Reyes, G.; Estrada, V.; Tolozano-Benites, R.; Maquilón, V. Batch Simplification Algorithm for Trajectories over Road Networks. *ISPRS International Journal of Geo-Information* **2023**, *12*, 399. <https://doi.org/10.3390/ijgi12100399>.

46. Kolda, T.G.; Bader, B.W. Tensor Decompositions and Applications. *SIAM Review* **2009**, *51*, 455–500. <https://doi.org/10.1137/07070111X>.
47. Tucker, L.R. Some Mathematical Notes on Three-Mode Factor Analysis. *Psychometrika* **1966**, *31*, 279–311. <https://doi.org/10.1007/BF02289464>.
48. Wright, J.; Yang, A.; Ganesh, A.; Sastry, S.; Yi Ma. Robust Face Recognition via Sparse Representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **2009**, *31*, 210–227. <https://doi.org/10.1109/TPAMI.2008.79>.
49. Jouppi, N.P.; Young, C.; Patil, N.; Patterson, D.; Agrawal, G.; Bajwa, R.; Bates, S.; Bhatia, S.; Boden, N.; Borchers, A.; et al. In-Datcenter Performance Analysis of a Tensor Processing Unit. In Proceedings of the Proceedings of the 44th Annual International Symposium on Computer Architecture, Toronto ON Canada, 06 2017; pp. 1–12. <https://doi.org/10.1145/3079856.3080246>.
50. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 06 2016; pp. 770–778. <https://doi.org/10.1109/CVPR.2016.90>.
51. Johnson, W.B.; Lindenstrauss, J. Extensions of Lipschitz Mappings into a Hilbert Space. In *Contemporary Mathematics*; Beals, R.; Beck, A.; Bellow, A.; Hajian, A., Eds.; American Mathematical Society: Providence, Rhode Island, 1984; Vol. 26, pp. 189–206. <https://doi.org/10.1090/conm/026/737400>.
52. Reyes, G.; Lanzarini, L.; Hasperu , W.; Bariviera, A.F. Proposal for a Pivot-Based Vehicle Trajectory Clustering Method. *Transportation Research Record: Journal of the Transportation Research Board* **2022**, *2676*, 281–295. <https://doi.org/10.1177/03611981211058429>.
53. Kingma, D.P.; Ba, J. Adam: A Method for Stochastic Optimization, 2014. <https://doi.org/10.48550/ARXIV.1412.6980>.
54. Bonomi, F.; Milito, R.; Zhu, J.; Addepalli, S. Fog Computing and Its Role in the Internet of Things. In Proceedings of the Proceedings of the First Edition of the MCC Workshop on Mobile Cloud Computing, Helsinki Finland, 08 2012; pp. 13–16. <https://doi.org/10.1145/2342509.2342513>.
55. Reyes, G.; Lanzarini, L.; Estrebou, C.; Fernandez Bariviera, A. Dynamic Grouping of Vehicle Trajectories. *Journal of Computer Science and Technology* **2022**, *22*, e11. <https://doi.org/10.24215/16666038.22.e11>.
56. Hughes, G. On the Mean Accuracy of Statistical Pattern Recognizers. *IEEE Transactions on Information Theory* **1968**, *14*, 55–63. <https://doi.org/10.1109/TIT.1968.1054102>.
57. Wolpert, D.; Macready, W. No Free Lunch Theorems for Optimization. *IEEE Transactions on Evolutionary Computation* **1997**, *1*, 67–82. <https://doi.org/10.1109/4235.585893>.
58. Shor, P. Algorithms for Quantum Computation: Discrete Logarithms and Factoring. In Proceedings of the Proceedings 35th Annual Symposium on Foundations of Computer Science, Santa Fe, NM, USA, 1994; pp. 124–134. <https://doi.org/10.1109/SFCS.1994.365700>.
59. Mead, C. Neuromorphic Electronic Systems. *Proceedings of the IEEE* **1990**, *78*, 1629–1636. <https://doi.org/10.1109/5.58356>.
60. Van Eck, N.J.; Waltman, L. Software Survey: VOSviewer, a Computer Program for Bibliometric Mapping. *Scientometrics* **2010**, *84*, 523–538. <https://doi.org/10.1007/s11192-009-0146-3>.
61. Cui, F.; Yue, Y.; Zhang, Y.; Zhang, Z.; Zhou, H.S. Advancing Biosensors with Machine Learning. *ACS Sensors* **2020**, *5*, 3346–3364. <https://doi.org/10.1021/acssensors.0c01424>.
62. Su, Z.; Wang, Y.; Luan, T.H.; Zhang, N.; Li, F.; Chen, T.; Cao, H. Secure and Efficient Federated Learning for Smart Grid With Edge-Cloud Collaboration. *IEEE Transactions on Industrial Informatics* **2022**, *18*, 1333–1344. <https://doi.org/10.1109/TII.2021.3095506>.
63. Wei, X.; Li, H.; Yue, W.; Gao, S.; Chen, Z.; Li, Y.; Shen, G. A High-Accuracy, Real-Time, Intelligent Material Perception System with a Machine-Learning-Motivated Pressure-Sensitive Electronic Skin. *Matter* **2022**, *5*, 1481–1501. <https://doi.org/10.1016/j.matt.2022.02.016>.
64. Brenner, B.; Hummel, V. Digital Twin as Enabler for an Innovative Digital Shopfloor Management System in the ESB Logistics Learning Factory at Reutlingen - University. *Procedia Manufacturing* **2017**, *9*, 198–205. <https://doi.org/10.1016/j.promfg.2017.04.039>.
65. Chen, X.; He, Z.; Chen, Y.; Lu, Y.; Wang, J. Missing Traffic Data Imputation and Pattern Discovery with a Bayesian Augmented Tensor Factorization Model. *Transportation Research Part C: Emerging Technologies* **2019**, *104*, 66–77. <https://doi.org/10.1016/j.trc.2019.03.003>.
66. Zhu, J.; Jiang, Q.; Shen, Y.; Qian, C.; Xu, F.; Zhu, Q. Application of Recurrent Neural Network to Mechanical Fault Diagnosis: A Review. *Journal of Mechanical Science and Technology* **2022**, *36*, 527–542. <https://doi.org/10.1007/s12206-022-0102-1>.

67. Galan, E.A.; Zhao, H.; Wang, X.; Dai, Q.; Huck, W.T.; Ma, S. Intelligent Microfluidics: The Convergence of Machine Learning and Microfluidics in Materials Science and Biomedicine. *Matter* **2020**, *3*, 1893–1922. <https://doi.org/10.1016/j.matt.2020.08.034>.
68. Boquet, G.; Morell, A.; Serrano, J.; Vicario, J.L. A Variational Autoencoder Solution for Road Traffic Forecasting Systems: Missing Data Imputation, Dimension Reduction, Model Selection and Anomaly Detection. *Transportation Research Part C: Emerging Technologies* **2020**, *115*, 102622. <https://doi.org/10.1016/j.trc.2020.102622>.
69. Moussalli, S.; Cardoso, W. Intelligent Personal Assistants: Can They Understand and Be Understood by Accented L2 Learners? *Computer Assisted Language Learning* **2020**, *33*, 865–890. <https://doi.org/10.1080/09588221.2019.1595664>.
70. Princy, R.J.P.; Parthasarathy, S.; Hency Jose, P.S.; Raj Lakshminarayanan, A.; Jeganathan, S. Prediction of Cardiac Disease Using Supervised Machine Learning Algorithms. In Proceedings of the 2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, India, 05 2020; pp. 570–575. <https://doi.org/10.1109/ICICCS48265.2020.9121169>.
71. Cobo, M.; López-Herrera, A.; Herrera-Viedma, E.; Herrera, F. An Approach for Detecting, Quantifying, and Visualizing the Evolution of a Research Field: A Practical Application to the Fuzzy Sets Theory Field. *Journal of Informetrics* **2011**, *5*, 146–166. <https://doi.org/10.1016/j.joi.2010.10.002>.
72. Shannon, C.E. A Mathematical Theory of Communication. *Bell System Technical Journal* **1948**, *27*, 379–423. <https://doi.org/10.1002/j.1538-7305.1948.tb01338.x>.
73. Lotka, A.J. The Frequency Distribution of Scientific Productivity. *Journal of Washington Academy Sciences* **1926**, *16*, 317–323.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.