IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 7, April 2024

Artificial Intelligence and Machine Learning Applications in Smart Production: Progress, Trends and Directions

Vaibhav Sharma[1] and Dr. Ashima Mehta[2]
Student Researcher, Department of Computer Science Engineering[1]
HOD, Department of Computer Science Engineering[2]
Dronacharya College of Engineering, Gurugram, India
vaibhavsharma251098@gmail.com and ashima.mehta@ggnindia.dronacharya.info

Abstract*: Adaptation and innovation are extremely important to the manufacturing industry. This development should lead to sustainable manufacturing using new technologies. To promote sustainability, smart production requires global perspectives of smart production application technology. In this regard, thanks to intensive research efforts in the field of artificial intelligence (AI), a number of AI-based techniques, such as machine learning, have already been established in the industry to achieve sustainable manufacturing. Thus, the aim of the present research was to analyze, systematically, the scientific literature relating to the application of artificial intelligence and machine learning (ML) in industry. In fact, with the introduction of the Industry 4.0, artificial intelligence and machine learning are considered the driving force of smart factory revolution. The purpose of this review was to classify the literature, including publication year, authors, scientific sector, country, institution, and keywords. The analysis was done using the Web of Science and SCOPUS database. Furthermore, UCINET and NVivo 12 software were used to complete them. A literature review on ML and AI empirical studies published in the last century was carried out to highlight the evolution of the topic before and after Industry 4.0 introduction, from 1999 to now. Eighty-two articles were reviewed and classified. A first interesting result is the greater number of works published by the USA and the increasing interest after the birth of Industry 4.0.*.

Keywords: artificial intelligence; machine learning; systematic literature review; applications; Industry 4.0; smart production; sustainability

I. INTRODUCTION

Smart production systems require innovative solutions to increase the quality and sustainability of manufacturing activities while reducing costs. In this context, artificial intelligence (AI)-driven technologies, leveraged by I4.0 Key Enabling Technologies (e.g., Internet of Thing, advanced embedded systems, cloud computing, big data, cognitive systems, virtual and augmented reality), are ready to generate new industrial paradigms .

In this regard, it is interesting to remember that the father of artificial intelligence, John McCarthy ,in the 1990s, defined artificial intelligence as "artificial intelligence is the science and engineering of making intelligent machines, especially intelligent computer programs". Generally, the term "AI" is used when a machine simulates functions that humans associate with other human minds, such as learning and problem solving .

On a very broad account, the areas of artificial intelligence are classified into 16 categories [4–8]. These are reasoning, programming, artificial life, belief revision, data mining, distributed AI, expert systems, genetic algorithms, systems, knowledge representation, machine learning, natural language understanding, neural networks, theorem proving, constraint satisfaction, and theory of computation [9– 11].

In the 21st century, AI has become an important area of research in all fields: Engineering, science, education, medicine, business, accounting, finance, marketing, economics, stock market, and law, among others [12–18]. The range of AI has grown enormously since the intelligence of machines with machine learning capabilities has created profound impacts on business, governments, and society [19]. They also influence the larger trends in global sustainability. Artificial intelligence can be useful to solve critical issue for sustainable manufacturing (e.g., optimization of energy resources, logistics, supply chain management, waste management, etc.). In this context, in smart production, there is a trend to incorporate AI into green manufacturing processes for stricter environmental policies [20]. In fact, as said in March 2019 by Hendrik Fink, head of Sustainability Services at PricewaterhouseCoopers, “If we properly incorporate artificial intelligence, we can achieve a revolution with regard to sustainability. AI will be the driving force of the fourth industrial revolution” [21].

Thus, subfields of AI, such as machine learning, natural language processing, image processing, and data mining, have also become an important topic for today’s tech giants. The subject of AI generates considerable interest in the scientific community, by virtue of the continuous evolution ofthe technologies available today.

The development of ML as a branch of AI is now very fast. Its usage has spread to various fields, such as learning machines, which are currently used in smart manufacturing, medical science, pharmacology, agriculture, archeology, games, business, and so forth.

According to the above considerations, in this work, a systematic literature review of research from 1999 to 2019 was performed on AI and the ML technique. Therefore, it is considered necessary to create a classification system that refers to the articles that jointly treat the two topics, in order to have greater variance and reflection. Furthermore, to gain a deeper understanding, the influence of other variables was explored, such as the thematic areas and the sectors in which the technologies are most influential. The main contribution of this work is that it provides an overview of the research carried out to date.

A number of impressive documentations of established research methods and philosophy have been discussed for several years. Unfortunately, little comparison and integration across studies exists. In this article, a common understanding of AI and ML research and its variations was created.

This paper is not attempting to provide an all-encompassing framework on the literature on AI and ML research. Rather, it attempts to provide a starting point for integrating knowledge across research in this domain and suggests paths for future research. It explores studies in certain novel disciplines: Environmental pollution, medicine, maintenance, manufacturing, etc.

Further research is needed to extend the present boundary of knowledge in AI by integrating principles and philosophies of some traditional disciplines into the existing AI frameworks [22–24].

The target that this document would like to assume is not the trigger of a sudden proliferation of an already consolidated sector, but it is hoped that this research could be an important intellectual tool for both the refocusing of the work and creating new intellectual opportunities. This paper presents valuable ideas and perspectives for undergoing research on AI and ML.

The final aim was to anticipate the transformation of the discipline in the future age. This would be a journey that may experience change in its course as new generations of scholars contribute to the dialogue and to the action. As noted earlier, this work presents a review, hence it lays a foundation for future inquiry. It not only offers a basis for future comparisons but prompts a number of new questions for investigations as well. While topics that might be considered as results of this work are numerous, some are of particularly broad interest or impact.

The paper is organized as follows. Section 2 presents the proposed methodology and details the research methodology adopted for the literature survey. Section 3 analyzes the main results of the bibliometric analysis. Finally, in Section 4, the main contribution of the research is summarized.

II. METHODOLOGY

The methodological approach used mixes bibliometric, content analysis, and social network techniques. In this study, a state-of-the-art research was conducted through the SCOPUS and Web of Science databases. For the publication time span, the time from 1999 to 2019 was considered with the intent to understand how the level of attention towards the topic has changed before and afterthe introduction of Industry 4.0. The research methodology chosen for this study was a systematic literature review [25]. The main phases of the study were as follows:

IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Impact Factor: 7.53

Volume 4, Issue 7, April 2024

Phase 1: Research and Classification.

The present phase was divided into three steps:

Step 1: Identification;

Step 2: Screening; and

Step 3: Inclusion.

In phase 1, bibliometric data was collected (step 1). Then, a screening of the overall result was carried out to identify which documents can be taken into consideration, in line with the research areas deemed interesting and relevant (step 2). At the end of this step, the last step (step 3) aimed to select the documents to be analyzed in detail.

Phase 2: Analysis.

Once phase 1 was completed, the next phase was phase 2, which was the analysis of the results. The approach used for the bibliometric analysis included:

The use of indicators for the parameters studied; and SNA (social network analysis) for the keywords.

The indicators chosen to perform the analysis were total papers (TPs), which is the total number of publications, and total citations (TCs), which is the total number of citations.

SNA finds application in various social sciences, and has lately been employed in the study of various phenomena, such as international trade, information dissemination, the study of institutions, and the functioning of organizations. The analysis of the use of the term SNA in the scientific literature has undergone exponential growth in the use of this mode of computable representation of complexand interdependent phenomena. For the purpose of the study, UCINET, NetDraw software was used, which was expressly designed for the creation and graphic processing of networks, and was used torepresent the keywords in the network, and Excel for data input.

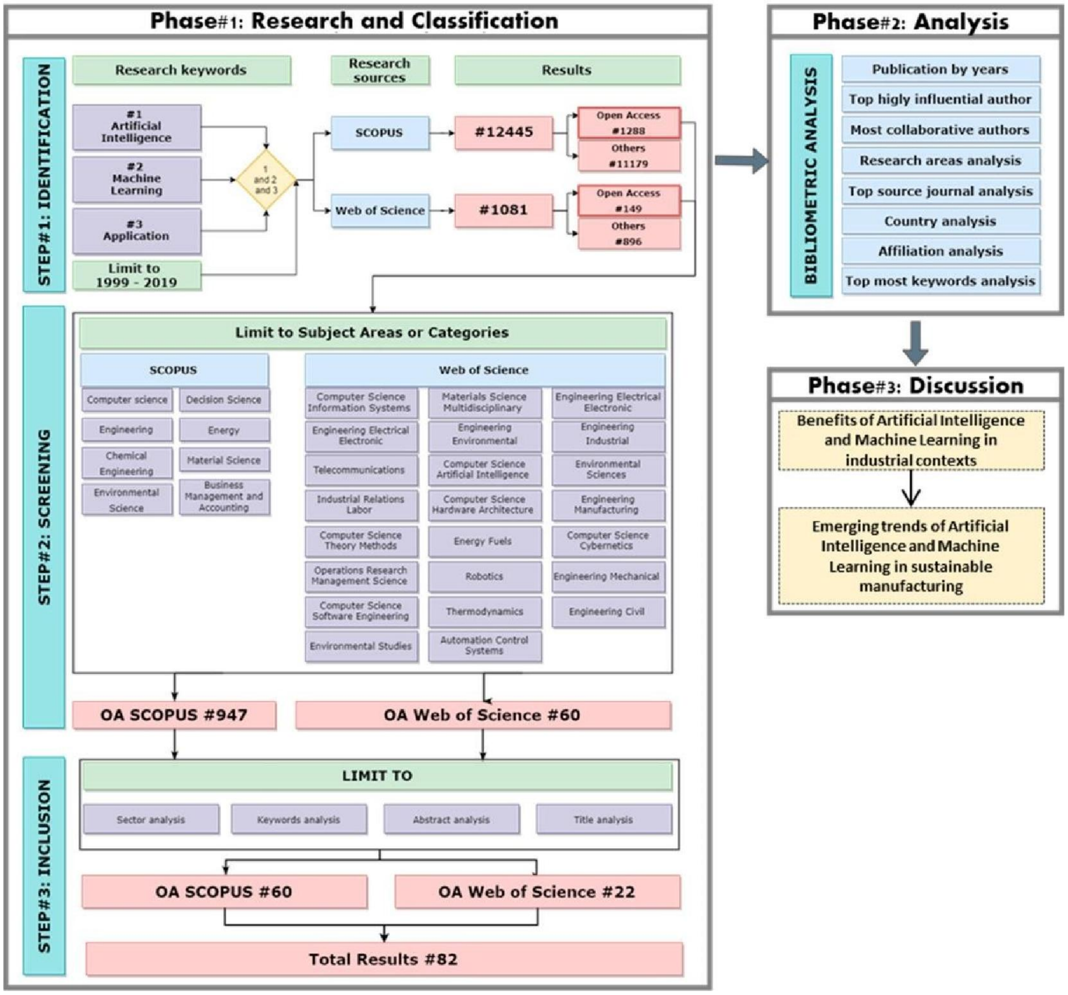
The software UCINET, NetDraw returned a sociometric network that describes the relationships between the classes, that is, data entered as input.

Furthermore, NVivo 12 software, the leading program for computer-assisted qualitative analysis (CAQDAS), was used to analyze keywords of all documents. In this specific case, it was used to identify the possible links between the keywords of the various documents examined, developing conceptual schemes from which to make interpretative hypotheses.

Phase 3: Discussion.

At the end of the second phase, a third and final one followed, where theresults were discussed, and conclusions were drawn.

In Figure 1, the main phases and steps followed for the analysis are shown



III. RESULTS OF THE BIBLIOMETRIC ANALYSIS

Phase 1: Research and Classification

The first phase consisted of the search for documents, which included the activities of collecting thematerial belonging to the academic universe. This first phase was divided into three steps as follows.

Identification (Step 1)

For a comprehensive survey of the phenomenon, an investigation on the Scopus (SCP) and Web of Science (WoS) databases was carried out using Boolean operators. We began by making a search query on the Scopus and WoS databases with the general keywords “artificial intelligence” AND “machine learning” AND “application”, as shown in Table 1.

In order to maintain the consistency of the results, the same keywords were used in both databases and a time horizon of 20 years was chosen, from 1999 to 2019.

The choice of keywords for performing the survey was based on the awareness that AI and MLcan be an important tool in the effort to adopt responsible business practices in the context of smart production. In this regard, it is worthy to note that with the increasingly urgent discussions of climate change, it seemed appropriate to focus our research on the topic of sustainability. Thus, the selection of papers also considered applications on sustainability

IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Impact Factor: 7.53

Volume 4, Issue 7, April 2024

Table 1. Keywords and time period

Keywords	Time Period
Artificial Intelligence	1999–2019
Machine Learning	
Application	

The search returned in total 13,512 documents.

The results extracted by Scopus are numerically superior to Web of Science (WoS): 12,445 for the first and only 1081 for the second one (Table 2).

Table 2. Total results of research on Scopus and WoS.

Research Carried out on 2019		
Source of research	Scopus	Web of Science
Results	12,445	1081

The result is not entirely unexpected, and the reason is to be found in the fact that Scopus, beingan Elsevier product, collects data from all the other databases, in particular Science Direct and thosequeried by the Scirus search engine, while Web of Science (WoS) collects fewer documents.

From the documents extracted in Scopus, it was found that most of them are conference papers (57.28%) and, subsequently, articles (33.85%).

On the contrary, the research on Web of Science (WoS) underlines that most of the documents are articles (46.12%) and, subsequently, proceedings papers (42.86%).

All the document types are filled in Table 3.

Table 3. Distribution of document types in Scopus and Web of Science.

Web of Science			**Scopus**		
Document Types	**Records**	**Contribute %**	**Document Types**	**Records**	**Contribute %**
Article	481	46.12	Conference Paper	7128	57.28
Proceedings paper	447	42.86	Article	4212	33.85
Review	133	12.76	Review	412	3.31
Editorial material	16	1.53	Article in Press	194	1.56
Meeting abstract	2	0.19	Book Chapter	177	1.42
Book chapter	1	0.1	Conference Review	177	1.42
Retracted publication	1	0.1	Book	90	0.72
-	-	-	Editorial	27	0.22
-	-	-	Note	10	0.08
-	-	-	Letter	9	0.07
-	-	-	Short Survey	9	0.07

AI began working in the 1940s and researchers showed strong expectations until the 1970s when they began to encounter serious difficulties and investments were greatly reduced.

Since then, a long period began, known as the "AI winter" [26]: Despite some great successes, such as IBM's Deep Blue system, which in the late 1990s defeated the then chess world champion Garri Kasparov, the study of solutions for AI has only come back for a few years. The push for a new technological development has been given by the I4.0, which considered AI as one of the primary key enabling technologies (KETs).

From this period onwards, the literature has been enriched with documents, as shown in Figure 2. Growth is apparent after 2011 when new technologies began to be implemented more frequently. In fact, the Industry 4.0 term first appeared at Hannover Messe in 2011 when Professor Wolfgang

Wahlster, Director and CEO of the German Research Center for Artificial Intelligence, addressed the opening ceremony audience.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 7, April 2024

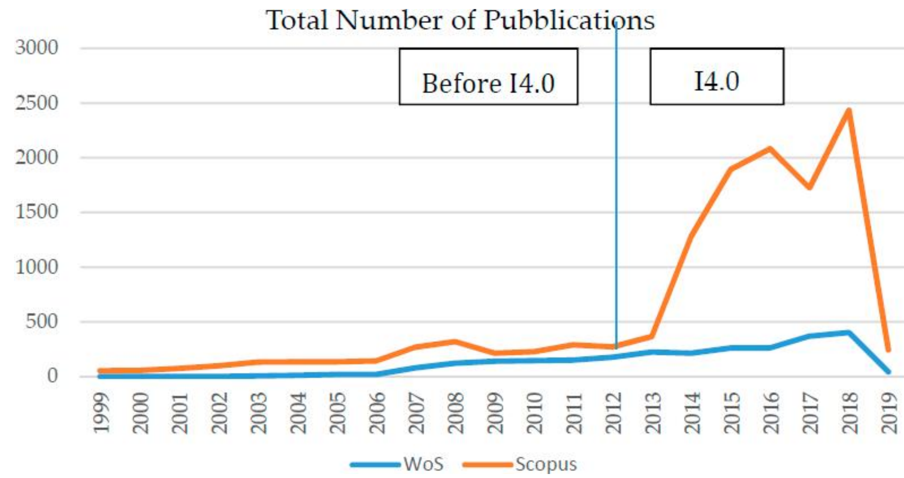


Figure 2. Research growth on Scopus and Web of Science.

In fact, this research indicates that over the time period considered (1999–2019), the number of published articles remains almost constant until 2013, from which it undergoes an increase.

Subsequently, the increase in the adoption of these ones has led researchers to keep pace with the growth of I4.0 [27].

Screening (Step 2)

Trying to give an overview of the topics and areas interface, in the screening phase, an analysis of documents characterized by free access was chosen, excluding those that have restrictions, and to restrict the field to the thematic areas of scientific interest.

With this in mind, the number of open access items has been drastically reduced (1288 results for Scopus and 149 for WoS) and, also applying the filter related to the thematic areas (Table 4), it determined a further reduction: 947 for Scopus and 60 for WoS

Table 4. Subject area filter on Scopus and WoS

Subject Area				
Scopus		**Web of Science (WoS)**		
Computer Science	Chemical Engineering	Computer Science Information Systems	Computer Science Artificial Intelligence	Automation Control Systems
Engineering	Energy	Materials Science Multidisciplinary	Environmental Sciences	Environmental Studies
Materials Science	Decision Science	Engineering Electrical Electronic	Computer Science Hardware Architecture	Operations Research Management Science
Environmental Science	Business Management and accounting	Telecommunications	Industrial Relations Labor	Robotics
		Engineering Environmental	Engineering Manufacturing	Thermodynamics
		Engineering Industrial	Computer Science Theory Methods	Energy Fuels
		Engineering Civil	Engineering Mechanical	Computer Science Cybernetics
		Computer Science Software Engineering	Multidisciplinary Sciences	

Note how the number of filters applied is different. The databases, in fact, offer the same search options, but, in the specific case of the thematic areas, the latter are more numerous and structured on Web of Science (WoS) compared to Scopus.

Inclusion (Step 3)

At the end of the screening process, the inclusion step was started, which consisted in the selection of documents, which was extracted from the last passage, destined to be included in the sample on which bibliometric analysis was performed. In this review step, for the purposes of eligibility, we examined the complete text of each document independently. For each article, we examined whether there was interest from the academic world, and if it contained case studies or real applications, proposals for new AI and ML algorithms, or possible future scenarios.

Therefore, the final sample to be analyzed consisted of 60 documents for Scopus and 22 for WoS.

Phase 2: Analysis

This section presents and discusses the findings of this review.

First, an overview of the selected studies is presented. Second, the review findings according to the research criteria, one by one in the separate subsections, are reported.

Top Highly Influential Analysis

This section lists the most highly cited documents in WoS and Scopus. The list is structured by research source, date, title, authors, source title, and top citation (TP) in WoS or Scopus, according to the research source. The whole list is available in the Appendix A. Looking into the Appendix A, it is possible underline that the document by Larrañaga, Calvo, Santana et al. in 2006 [28] has the highest citation count of 298. This article reviews machine learning methods for bioinformatics and it presents modelling methods. Moreover, the document year is 2006, so before I4.0 was introduced. Therefore, having more years than today has an advantage in terms of diffusion. This means that it is one of the most influential documents in the academic world, as it proposes some of the most useful techniques for modelling, giving the document the opportunity to become a pioneer in the computer science research area.

Obviously, all documents before I4.0, in general, have more citations than the most recent documents. However, it is significant to note that even recent documents have a very high number of citations compared to the year of publication. This denotes the interest in the topic from the scientific community.

The citation analysis revealed that the first article that we can identify among the most cited in the I4.0 period dates to 2016. The work, published by Krawczyk [29], proposes application models to further develop the field of unbalanced learning, to focus on computationally effective, adaptive, and real-time methods, and provides a discussion and suggestions on the lines of future research in the application subject of the study. It received 119 citations. Moreover, an article published by Wuest, Weimer, Irgens et al. [30] received much attention among the scientific community. It contributes by presenting an overview of the available machine learning techniques.

Finally, the citation analysis pointed out that the average number of citations of all documents is 16.58. This value is expected to increase rapidly considering the interest in the issues of ML and AI.

Publications by Years

Consistent with what is defined in Section 3.1.1., the study shows that the number of items included in the analysis is definitely low for the entire period before I4.0 and then suddenly increases, starting in 2012. The data shown in Figure 3 also show two holes in the 2001–2008 and 2008–2011 intervals. This means that the technological applications were limited before it became an enabling technology of I4.0 in all respects, only to have a peak of technological implementation, as was foreseeable.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 7, April 2024

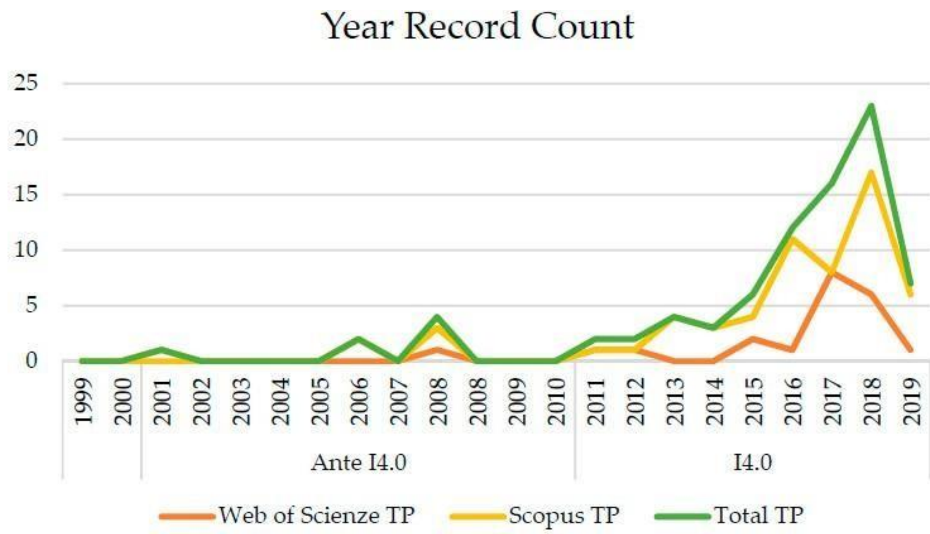


Figure 3. Years of publications.

With reference to 2019, the figure refers to the first months of the year, so it is plausible that during the year, there will be a further increase in the documents in the literature. Furthermore, an increase is expected in the coming years, in parallel with the growth of I4.0

Most Collaborative Authors

The analysis highlighted that most of publications have more than one author. From this point of view, it is possible to identify the number of authors for each document. As shown in Figure 4, most of the manuscripts were produced by groups ranging from two to five authors. The indicators chosen to perform the analysis were total papers (TPs), which is the total number of publications.

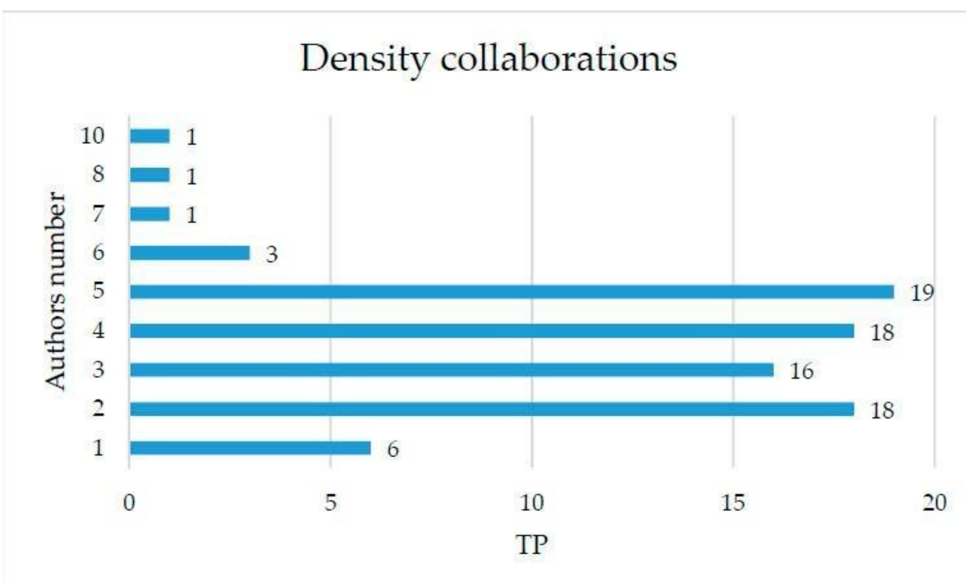


Figure 4. Collaborative groups

Research Areas Analysis

The total research area analysis collected from the 82 papers was 164 because each paper can be considered as more than one research area analysis. Given the small number of documents identifiedin the period before I4.0, the ranking refers mostly to the current industrial revolution. Also, in this case, the result is consistent with the introduction of paradigm 4.0, which has intensified research and the adoption of technology.

The first thematic areas and disciplines that are at the top of the ranking are computer science, engineering and biochemistry, genetics, and molecular Biology, respectively, with 29%, 23%, and 6% of publications. Furthermore, the other disciplines identified for which applicative findings are found are considered transversal to the first three disciplines and this is a consequence of I4.0. In terms of the percentage contribution, the first three areas cover about 60% of the papers considered.

Considering the top 20 research areas, given the frequency of the research areas' distribution, Figure 5 shows a higher level of concentration in the disciplines indicated above.

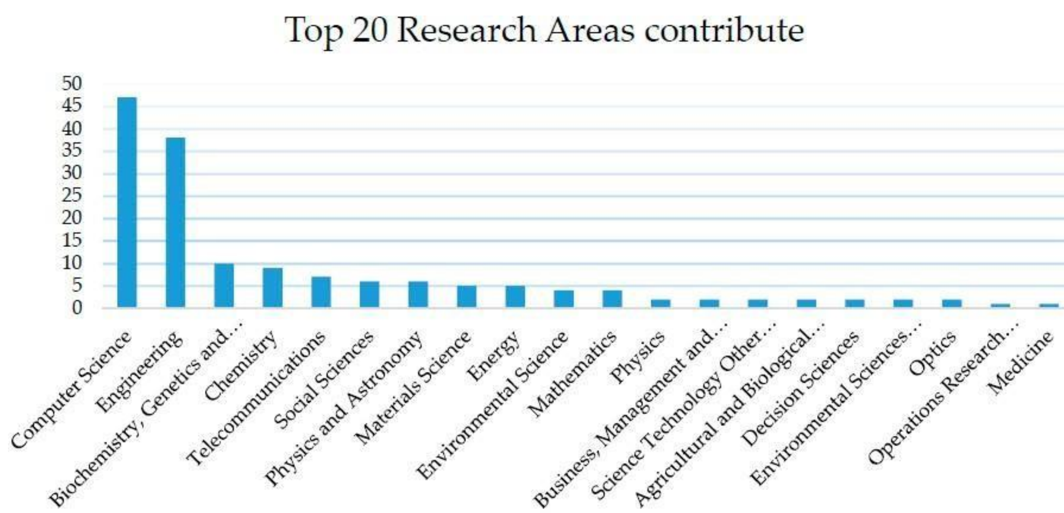


Figure 5. Top 20 research areas contributions.

In fact, in terms of the percentage contribution, the first five areas cover about 70% of the papers considered. Regardless, by only counting research areas found once, there is a total of 27.

This means two things:

The large number of fields in which this kind of research is involved; and

Most papers have a transversal approach, that is, the object of each research crosses more than one field of application, thus involving more research areas.

This confirms the wide interest in these subjects from several fields.

Top Source Journals Analysis

In this section, the top 20 sources or journals that were published most frequently were extracted.

A journal is a time-bound publication with the objective of promoting and monitoring the progress of the discipline it represents.

In this specific case, the total source journals detected from the documents is 74, but, considering the top 20, given the frequency of the source journals' distribution, only the first 13 sources have more than one paper published, with a total percentage contribution of 43% of the total.

After analyzing the sources separately, the results obtained in the two databases were found to not be the same. In WoS, the top source journal was *IEEE Access* with two publications while in Scopus the top source journals are *Procedia Computer Science*, *Matec Web of Conferences*, and *Machine Learning*
with four publications, which contribute 5% of the total.