DOI: https://doi.org/10.46354/i3m.2019.mas.007

MACHINE LEARNING TECHNIQUES APPLIED TO INDUSTRIAL ENGINEERING: A MULTI CRITERIA APPROACH

Fabio De Felice [(a)], Marta Travaglioni [(b)], Giuseppina Piscitelli [(b)], Raffaele Cioffi [(b)], Antonella Petrillo [(b)]

[(a)] Department of Civil and Mechanical Engineering - University of Cassino and Southern Lazio, Italy
[(a)] Department of Engineering - University of Naples "Parthenope", Italy

[(a)]defelice@unicas.it, [(b)]antonella.petrillo@uniparthenope.it, [(b)]marta.travaglioni@uniparthenope.it, [(b)]giuseppina.piscitelli@uniparthenope.it

ABSTRACT
With the Industry 4.0 (I4.0) beginning, the world is witnessing an important technological development. The success of I4.0 is linked to the implementation of enabling technologies, including Machine Learning, which focuses on the machines' ability to receive a series of data and learn on their own. The present research aims to systematically analyze the existing literature on the subject in various aspects, including publication year, authors, scientific sector, country, institution and keywords. Understanding and analyzing the existing literature on Machine Learning applied to predictive maintenance is preparatory to recommend policy on the subject.

Keywords: machine learning, predictive maintenance, literature review, industry 4.0

1. INTRODUCTION
With Industry 4.0, companies can exploit Artificial Intelligence applied as predictive analysis, to monitor the production process and have a more in-depth view of their activity in real time. The machines and components connected to the Internet of Things system send information any time, then they are historicized and analyzed with algorithms (Petrillo et al., 2019). The latter work on Machine Learning logics to predict system failures in advance and estimate the residual life of the machines. This means there is no longer the dilemma between waiting for the machine downtime or replacing the components ahead of time.
The challenge of proper scheduling grows with the complexity of machines. In this context, predictive can help to detect the anomalies and failure patterns and provide early warnings (Mobley 2002). It is based on the possibility of recognizing a progression anomaly through the notice and interpretation of weak signals that define a future failure. Therefore, it is not just talking about data generated by the interactions between machines and operators or by M2M (Machine to Machine) connection systems, but to process them with Analytics solutions.
The existence of a system that collects all the data coming from multiple sources allows to identify recursive patterns, that define the probability with which an event can happen in the future and to learn automatically from the feedback of the activities carried out. Thus, Machine Learning (ML) models are enabled to activate Predictive Maintenance (PdM).
With I4.0 and the consequent increase in data availability, the cases in which the ML is implemented for the PdM increased.
Given considerable importance of the subject, the present research will carry out a bibliographic survey that highlights the most important texts with respect to a research area, including year of publication, authors, scientific sector, country, institution and keyword. It not only offers a basis for future comparisons but prompts a number of new questions for investigations as well.
The rest of the paper is organized as follows: Section 2 presents a brief review of machine learning techniques; in Section 3 proposed bibliometric analysis method is explained; in Section 4 the sample building process is explained; in Section 5 results of the bibliometric analysis is developed; in Section 6 results of analysis are discussed. Finally, in Section 7 are summarized the main contribution of the research.

2. BRIEF REVIEW OF MACHINE LEARNING TECHNIQUES
One of the largest areas currently included in the label of artificial intelligence is Machine Learning. This term includes a large number of different techniques and approaches, but all of these share a clear and unambiguous property: solving problems through algorithms that have the ability to learn from what is entered as input. In these terms, the concept of knowledge changes. Knowledge is no longer the human one and transferred to the machine, but it is learned by the machine itself in an autonomous way from the activities that it carries out gradually. Consequently, the role of man also changes: he will only have to define, through programming, how the machine will have to learn. In some cases, the operator provides examples and information from which the machine must learn, so it can develop a "knowledge" that it will transfer or allow to automate activities.
The different approaches have been uniquely classified by scientific community, depending on their learning model. Each model works, in principle, on the basis of

two distinct approaches: supervised learning and unsupervised learning (Samuel A.L. 1959).
In the case of supervised learning, the computer is given concrete examples to be used in carrying out the required task, while in the case of unsupervised learning the software works without any kind of assistance.
Within these two categories there are subsets that allow us to classify Machine Learning even more in detail.
The first is *Supervised Learning*, the computer is given examples in the form of possible inputs and the respective desired outputs, to extract a general rule that associates the input with the correct output. It is usually used in applications where historical data can predict future events.
The second subset is *Unsupervised Learning*, where the system is provided only with data sets without indication of the result sought. The computer receives inputs in no way labeled, from which it must find a structure. This means the system is not given the "right" answer and therefore the algorithm must find out what is shown to it.
The third subset is *Reinforcement Learning*. The system (which can be a computer, software or algorithm) must be able to interact with a dynamic environment, from which to draw input data and reach a goal. If the goal is achieved, the system will be rewarded with a performance evaluation. The purpose is to learn an optimal policy that selects the subsequent actions to each state in order to maximize the rewards accumulated over time.
The fourth and final subset is the *Semi-Supervised Learning*. This, unlike the other subsets, is a hybrid model: incomplete data sets are supplied to the computer (some of the inputs are also output) as in Supervised Learning, and others lack them, as in Unsupervised Learning.

3. BIBLIOMETRIC ANALYSIS METHODS

The interest in using bibliometric techniques has increased thanks to the availability of large scientific databases. It involves a series of techniques that are used to quantify the process of written communication and identify patterns. The methodological used approach mixes bibliometric, content analysis and social network techniques. In this state-of-the-art study the research was initiated through the consultation of electronic databases, Scopus (SCP) and Web of Science (WoS), on April 9, 2019 and dividing it into the following phases.
In Phase 1, bibliometric data was collected from the SCP and WoS databases. This phase required three steps for the construction of the sample to be analyzed:

1. Step #1: identification;
2. Step #2: screening;
3. Step #3: inclusion.

Once Phase 1 is completed, the next phase is Phase 2, which is the analysis of the results.
The approaches used for the bibliometric analysis were:

- the use of indicators for the parameters studied;
- the SNA (Social Network Analysis) for the authors analysis.

The indicators chosen to perform the analysis are Total Papers (TP), which is the total number of publications and Total Citation (TC), which is the total number of citations. The SNA finds application in various social sciences, lately 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 an exponential growth in utilization of this mode of computable representation for complex and interdependent phenomena. For the purpose of this study, UCINET, NetDraw software was used, expressly designed for the creation and graphic processing of networks. It was used to represent the connection between publication and citation in authors' network, and Excel for data input. To analyze keywords, NVivo 12 software was used. At the end of the second phase, a third and final one follows, where the results will be discussed, and conclusions will be drawn. All analyzed documents are listed in Appendix (Table A).

4. BIBLIOMETRIC ANALYSIS: PHASE 1 - RESEARCH

In this paragraph, the results of the bibliometric research are analyzed, starting from Phase 1 in which the process of construction of the sample is described.

4.1. Step#1: Identification

First phase is dedicated to documents collection and the sample's construction.
It is divided into three steps; every step is useful for selecting the documents of interest and excluding ones that are not interesting for the purpose.
The first step is named "Identification".
In this step, Scopus (SCP) and Web of Science (WoS) databases were taken into consideration. In order to maintain the consistency of the results, the same keywords were used:

- Machine Learning;
- Predictive Maintenance.

In addition, a time period of 20 years was chosen, from 1999 to 2019, as shown in Table 1.

Table 1: Keywords And Time Period For Research

Keywords and Time Period		
Keywords	Machine Learning (ML)	Predictive Maintenance (PM)
Time Period	From 1999	To 2019

The results extracted by Scopus are numerically superior to Web of Science: 320 for the first database and just 173 for the second one (Table 2).

Table 2: Total Results Of Research On SCP And Wos

Data Extraction			
Source of research	SCP	WoS	Total
Results	320	173	493

Figure 2 shows the aggregated result for Scopus and Web of Science query; it points out that the results obtained are scarce in the period before I4.0. On the contrary, there is an increase in publications since the beginning of the fourth industrial revolution. Overall, this is a fairly poor operating result, just 493 documents, most of them were concentrated in year 2018, with 172 results (Figure 1).

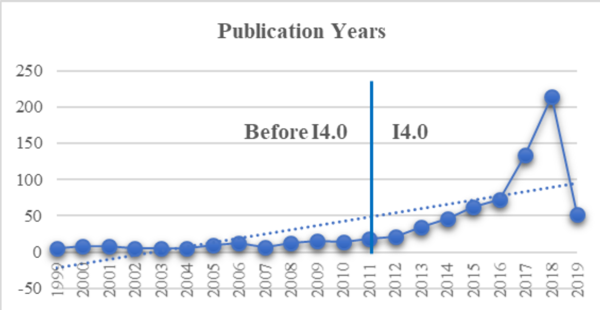

Figure 1: Research Growth

Wanting to analyze the numbers before and after I4.0 (Table 3), just before the fourth revolution, 23 documents were recorded in SCP and 19 in WoS. While after 2011 there are 297 documents in Scopus and 154 in WoS, as shown in Table 3.

Table 3: SCP And Wos Research Areas

Results			
-	Before I4.0	I4.0	Total
SCP	23	297	320
WoS	19	154	173
Total	42	451	-

With particular attention to year 2018, 122 and 50 documents are recorded on Scopus and Web of Science respectively.

The reason is not to be sought in the scarce interest, although Machine Learning is dated already to the last century, Predictive Maintenance has encountered an increase in development only with I4.0.

It follows that the analysis of the literature concerning the period 1999 - 2010 can be considered negligible, since it is an overall result equal to 42 documents out of 493.

On the contrary, it is useful to carry out an analysis starting from 2011 until the present. In this period the identified literature has a result equal to 451 documents, distributed unevenly along the time axis, with greater concentration in year 2018 (Figure 2).

4.2. Step#2: Screening

Starting from the result obtained in the first step, in the screening step, documents characterized by open access were chosen to be analyzed, excluding those with restrictions, from 2011 to 2019. With this step, the number of documents dropped dramatically. Thus, 38 documents are obtained from Scopus, while 29 documents are obtained from the Web of Science (Table 4).

Table 4: Results Of OA In Time Period 2011-2019

Results OA			
	OA	Time Horizon	
SCP	38	From 2011	To 2019
WoS	29		
Total	67		

This is a decrease of 87% and 83% respectively.

4.3. Step#3: Inclusion

The last step is dedicated to finalizing the final sample. Trying to give an overview of the topics and areas of the interface, filters have been applied in relation to the thematic areas to which they belong, shown in Table 5.

Table 5: Research Filters And Results Of Filters Application

Filters	
SCP	WoS
Engineering	Engineering Industrial
Computer Science	Engineering Manufacturing
Material Science	Computer Science Information System
Energy	Computer Science Interdisciplinary Applications
Business, Management and Accounting	Engineering Electrical Electronic
Chemical Engineering	Engineering Mechanical
Decision Science	Engineering Multidisciplinary
-	Telecommunication
-	Automation Control System
-	Computer Science AI
-	Engineering Civil
Results	
29 documents	18 documents

The result is another decrease in the documents number of our interest, obtaining 29 documents on Scopus and 18 on Web of Science (Table 5).

Finally, they were analyzed to exclude redundancies or document overlaps: at the end of phase 1, the sample to be analyzed consists of 37 total documents (Table A).

5. BIBLIOMETRIC ANALYSIS: PHASE 2 - ANALYSIS

This section presents and discusses the findings of this review.

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

5.1. Document Types

Before proceeding to the document's analysis, it is appropriate to establish their type. In fact, the 37 documents composing the sample are divided into two types: 17 are articles and 20 are conference papers, 46% and 54% respectively.

5.2. Publication by Years

The analysis of the years of publication shows that the research was not particularly intense in the years following the introduction of I4.0. In fact, it seems the research is peaking only recently, with documents concentration in year 2018, as shown in Table 6.

Table 6: Top Publications By Years

Results	
Year	TP
2011	1
2012	0
2013	0
2014	2
2015	3
2016	1
2017	6
2018	22
2019	2

In 2019 there is a small number, since the research was carried out at the end of the first four months. However, tracing a line representing the trend (Figure 2) it is plausible to think that there is a large number of documents.

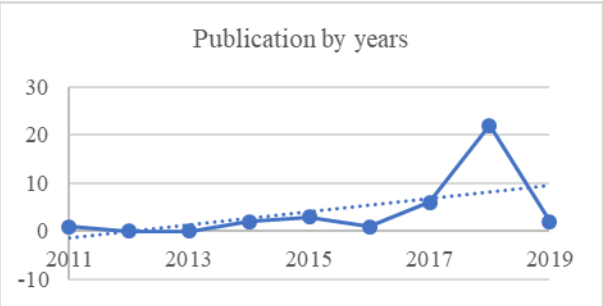

Figure 2: Publication By Years

5.3. Citation by Years

Compared to the years of publication, it is interesting to know the number of citations.
Table 7 shows the citations number for each year of publication and the graph relating to the table is shown in Figure 3.

Table 7: Top Citation By Years

Results	
Year	TC
2011	27
2012	0
2013	0
2014	9
2015	79
2016	3
2017	61
2018	10
2019	0

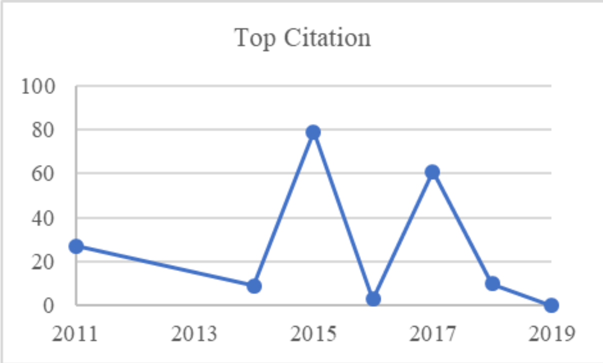

Figure 3: Top Citation By Years

The first aspect that emerges is the citations number in years 2015 and 2017, respectively with 79 and 61 citations. This can be interpreted as the source of the publications produced in year 2018, which is the most productive one (Figure 4).
One of the documents of 2015 is produced in Italy (Table A, ID4). Specifically, this includes 43 citations out of 79 totals. The document proposes a Predictive Maintenance methodology, that allows dynamical decision rules to be adopted for maintenance management and can be used with high-dimensional and censored data problems. Furthermore, the effectiveness of the methodology is demonstrated using a simulated example and a benchmark semiconductor manufacturing maintenance problem (Susto and Schirru and Pampuri and McLoone and Beghi 2015). Indeed, this is the document that counts the greatest number of citations of the whole sample (Table A).

5.4. Country Analysis

Referring to the previous paragraph, it is useful to understand which countries are more productive.
From the analysis, it is clear (Figure 5) that the countries most interested in the topic are China, USA and Italy, respectively with 7, 6 and 4 documents.

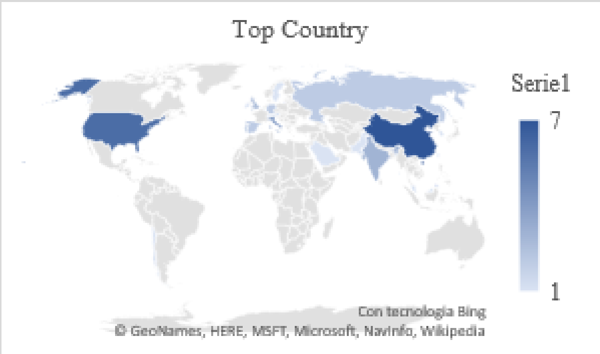

Figure 4: Top Country Analysis

Looking at Europe, it can be determined that it is the most productive, with an overall result of 19 documents, which means a 51% contribution.

5.5. Research Areas Analysis

According with the thematic areas used as a filter reported in Table 5 (paragraph 4.3) and coherently with the topic under examination, the analysis of the research areas revealed that the higher interest areas are "Engineering" and "Computer Science" (Table 8) with 24 and 16 documents respectively.

Table 8: Top Publications By Research Areas

Research	
Research Areas	TP
Engineering	24
Computer Science	16
Mathematics	7
Materials Science	5
Energy	4
Physics and Astronomy	4
Biochemistry, Genetics and Molecular Biology	2
Agricultural and Biological Sciences	1
Business, Management and Accounting	1
Decision Sciences	1
Economics, Econometrics and Finance	1
Social Sciences	1

The other areas also show interest, but with a much smaller number, as shown in Table 8.

5.6. Top Source Journal Analysis

In this section top 10 source or journals which are publishing most frequently are extracted. The total source journal detected from the document is 23 but, considering the top 10, given the source frequency distribution, Figure 6 shows that only the first 7 sources have more than one paper published with a percentage contribute of 57% on the total.

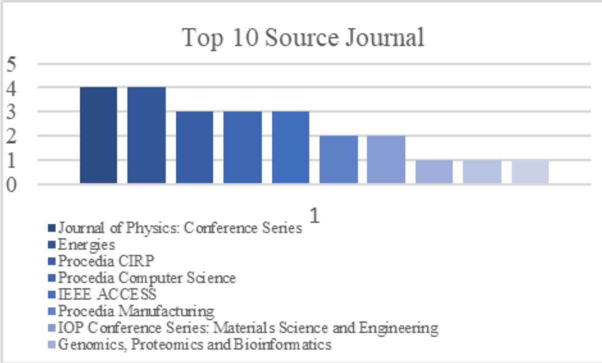

Figure 5: Top 10 Source Journal Analysis

Wanting to analyze the sources separately, the results obtained in the two databases are not the same. Referring to Table A, in Scopus the top source journal is "Journal of Physics: Conference Series" with 4 publications and a contribute of 11% of the total. In Web of Science the top source journal is "ENERGIES", with 3 publications and a contribute of 8%.

Aggregating the data collected from the two databases, the ranking moves to that obtained by Scopus, moving "ENERGIES" to second place (Figure 5). After, 3 source journals have 3 publications and 2 others have 2. The rest of the publications have a one-to-one relationship with the corresponding source journal: for this reason only some of them have been shown in Figure 6.

The low level of concentration of the sources suggests that there is a great deal of interest on these topics by several scientific journals.

5.7. Most Collaborative Authors

From the analysis carried out, it can be observed that all authors are authors or co-authors of a single article, except for Gao Robert X. (Table A, ID 8 and ID 26), who have published two documents.

The most useful analysis in this case is represented by a relationship between authors and citation. Only the top 20 authors are shown in Table 9, which represents a comparison between the number of publications (TP) with the number of citations (TC) associated with each author.

Table 9: Top 20 Authors Analysis

Results		
Top 20 Authors	TP	TC
Beghi, A.	1	43
Martins, T.	1	43
Oneto, L.	1	43
Pampuri, S.	1	43
Schirru, A.	1	43
He, Z.	1	27
Sun, C.	1	27
Zhang, Z.	1	27
Gao, R.X.	2	23
Jennings, C.	1	23
Kumara, S.	1	23
McLoone, S.	1	23

Miao, Q.	1	23
Susto, G.A.	1	23
Terpenny, J.	1	23
Triphathi, A.K.	1	23
Tsui, K.-L.,	1	23
Wang, C.	1	23
Wang, D.	1	23
Wu, D.	1	23

However, there are other researchers who are authors of two documents, but the number of citations associated with them is lower or even zero. For this reason they are not among the top 20 authors and therefore do not appear in Table 9. For completeness, the authors of two documents and their associated citations number are shown in Table 10.

Table 10: Authors Of 2 Documents

Results		
Authors	TP	TC
Gao, R.X.	2	23
Sahli, H.	2	14
Liu, Ling	2	1
Praça, I.	2	0

Starting from Table 9, it can represent the SNA, through which the link between publications number and citations number of each author is shown.
The graph shown in Figure 7, nodes and leaves can be identified. The nodes are represented by the blue square and correspond to publications number (TP) and citations number (TC). Red circles are the leaves and represent the authors' name. Each leaf is connected to both nodes with a different weight.
The weight of the connection between the authors and TP node is not showed, but the line that has a weight equal to 2.0 is highlighted. This line, colored with blue (Figure 7), is the connection between the author Gao R.X. and the number of documents produced by him. All other connections, consistent with Table 9, is equal to 1.0.
The weight that links the authors and the citations number, on the contrary, is reported in Figure 6. The highest weights among all the values shown in Table 9 are highlighted by the green lines and range from 43.0 to 27.0. Weight that is 23.0 is not highlighted.

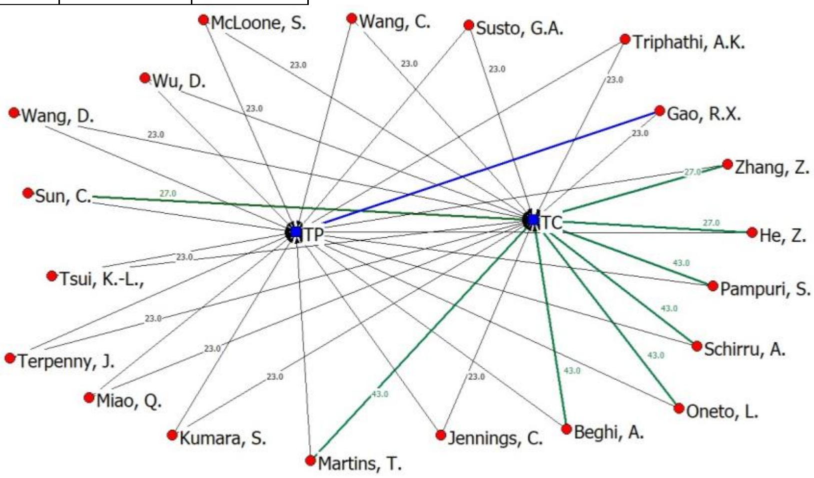

Figure 6: Publications And Citations' Connection For Each Authors

5.8. Top Most Keywords Analysis

Almost all the keywords are common in WoS and Scopus extracted documents and use one or more keywords. To analyze the most popular keywords, a software to create word clouds: NVivo 12, produced by QSR International Pty Ltd ® was used.
The word cloud in Figure 7 highlights the keywords most used by the authors to index their publications.
As shown in Figure 8, the largest words are "machine", "learning", "maintenance", "predictive", "system" and other words. These are the words that appear several times in the documents. For example, the word "learning" appears 14 times and "machine" appears 13 times. Top 20 keywords frequency (named F in Table 11) that appeared in the documents analyzed are shown in Table11.

Figure 7: Keywords Cloud Contribute By NVivo 12

Table 11: Top 20 Keywords Frequency

Results of Top 20 Keywords Frequency (F)			
Keywords	F	Keywords	F
Learning	14	Analytics	4
Machine	13	System	4
Data	11	Health	4
Maintenance	9	Prediction	3
Predictive	9	Industry	3
Monitoring	5	mining	3
Condition	5	fault	3
Analysis	5	identification	2
Prognostics	4	classification	2
Management	4	manufacturing	2

As the words frequency decreases, the size of words in the cloud also decreases. Therefore, the words that are very small in the cloud are those that appear only once.

6. BIBLIOMETRIC ANALYSIS: PHASE 3 - DISCUSSION

Interesting information emerged from the analysis carried out.

The first observation concerns the number of documents. The low number of documents available from 2011 to nowadays shows how the joint topic is still very young and unripe, but the sudden growth that has affected the last few years is indicative of an interest that has been made more alive. However, the technologies are not at all young (just think that the ML dates back to the late 1950s) but have not undergone such a rapid technological process as in recent years, with I4.0. In this regard, it is recalled that throughout history AI technology has undergone the so-called "AI winter" (this term first pronounced in 1984 by Roger Schank and Marvin Minsky and appeared as the topic of a public debate at the annual meeting of American Association of Artificial Intelligence).

Additional information that can be extracted from the analysis concerns the research area. Although Artificial Intelligence (and therefore ML) is a branch of Computer Science, the research area with the largest number of publications is Engineering. The reason is linked to the very high interest of Machine Learning for Predictive Maintenance interested in engineering applications.

Furthermore, it can be said that the countries most interested in scientific research are Europe, China and USA, which are the most industrialized countries.

With technological growth, the geographical spread of academic interest but also of applications is not excluded.

It is important to underline that this document was produced using only two databases. These are WoS and Scopus, in which only documents with open access were included.

7. CONCLUSION

This document focused on the study of the current state-of-the-art in Machine Learning for Predictive Maintenance topics. To date, with the tools available to scientific community, the literature available on any subject is very wide. As a result, complete coverage of all documents published in relation to a specific topic can be very difficult. Therefore, a selection of the most relevant literature was made and a document was produced, that provides a review of the applications in various scientific fields, using ML techniques. For the selection of documents, objective and clear investigation methods were used, independent of the researchers' experience. Among the objectives of the document there is not only the wish to provide a complete picture on research literature, but also a starting point for integrating knowledge through research in this field, suggesting future research paths. There are therefore many other documents with limited access and other indexing databases, such as Google Scholar, which could be integrated into the research in the near future.

ACKNOWLEDGMENTS

This work is the result of the Italian project "*Linee Guida per I4.0 – Campania*" - funded by Regione Campania within POR FSE 2014-2020 Asse IV "Capacità istituzionale e amministrativa" objectives 18 (RA) 11.3 and 21 (RA) 11.6

REFERENCES

Petrillo, A., De Felice, F., Zomparelli, F., 2019. Performance measurement for world class manufacturing: a model for Italian automotive industry. Total Quality Management & Business Excellence, Pages 908-935, Volume 30, 2019 - Issue 7-8.

Fernandes M., Canito A., Bolón-Canedo V., Conceição L., Praça I., Marreiros G., 2019. Data analysis and feature selection for predictive maintenance: a case-study in the metallurgic industry. International Journal of Information Management, 46, 252-262.

Lee C.Y., Huang T.S., Liu M.K., Lan C.Y., 2019. Data science for vibration heteroscedasticity and predictive maintenance of rotary bearings. Energies, 12 (5), art. no. 801.

Serio L., Antonello F., Baraldi P., Castellano A., Gentile U., Zio E., 2018. A smart framework for the availability and reliability assessment and management of accelerators technical facilities. Journal of Physics: Conference Series, 1067 (7), art. no. 072029.

Kleftogiannis D., Ashoor H., Bajic V.B., 2018. TELS: a novel computational framework for identifying motif signatures of transcribed enhancers. Genomics, Proteomics and Bioinformatics, 16 (5), 332-341.

Bedi J., Toshniwal D., 2018. Empirical mode decomposition based deep learning for electricity demand forecasting. IEEE Access, 6, 49144-49156.

Shafi U., Safi A., Shahid A.R., Ziauddin S., Saleem M.Q., 2018. Vehicle remote health monitoring and

prognostic maintenance system. Journal of Advanced Transportation, 2018, art. no. 8061514.
Pandit R., Infield D., 2018. Gaussian process operational curves for wind turbine condition monitoring. Energies, 11 (7), art. no. 1631.
Ullah I., Yang F., Khan R., Liu L., Yang H., Gao B., Sun K., 2017. Predictive maintenance of power substation equipment by infrared thermography using a machine-learning approach. Energies, 10 (12), art. no. 1987.
Wang D., Tsui K.-L., Miao Q., 2017. Prognostics and health management: a review of vibration based bearing and gear health indicators. IEEE Access, 6, 665-676.
Wilson D.L., Coyle J.R., Thomas E.A., 2017. Ensemble machine learning and forecasting can achieve 99% uptime for rural handpumps. PLoS ONE, 12 (11), art. no. e0188808.
Wu D., Jennings C., Terpenny J., Gao R.X., Kumara S., 2017. A comparative study on machine learning algorithms for smart manufacturing: tool wear prediction using random forests. Journal of Manufacturing Science and Engineering, Transactions of the ASME, 139 (7), art. no. 071018.
Zhou F., Gao Y., Wen C., 2017. A novel multimode fault classification method based on deep learning. Journal Of Control Science And Engineering, 2017, art. no. 3583610.
Wang C., Liu X., Liu H., Chen Z., 2016. A fault diagnostic method for position sensor of switched reluctance wind generator, Journal Of Electrical Engineering And Technology, 11 (1), 29-37.
Cartella F., Lemeire J., Dimiccoli L., Sahli H., 2015. Hidden semi-markov models for predictive maintenance. Mathematical Problems in Engineering, 2015, art. no. 278120.
Tulabandhula T., Rudin C., 2014. On combining machine learning with decision making. Machine Learning, 97 (1-2), 33-64.
Susto, G.A., Schirru A., Pampuri S., McLoone S., Beghi A., 2015. Machine learning for predictive maintenance: a multiple classifier approach, IEEE Transactions on Industrial Informatics, 11 (3), 812-820.
Leahy K., Gallagher C., O'Donovan P., Bruton K., O'Sullivan D.T.J., 2018. A robust prescriptive framework and performance metric for diagnosing and predicting wind turbine faults based on scada and alarms data with case study, Energies 11 (7), art. no. 1738.
Peres R.S., Rocha A.D., Leitao P., Barata J., 2018. IDARTS - Towards intelligent data analysis and real-time supervision for industry 4.0, Computers in Industry, 101, 138-146.
Wu Z.Y., Luo H., Yang Y.N., Lv P., Zhu X.N., Ji Y., Wu B.A., 2018. K-PdM: KPI-oriented machinery deterioration estimation framework for predictive maintenance using cluster-based hidden markov model. IEEE Access, 6, 41676-41687.
Mohamudally N., Peermamode-Mohaboob M., 2018. Building an anomaly detection engine (ade) for iot smart applications. Procedia Computer Science, 134, 10-17.
Uhlmann E., Pontes R.P., Geisert C., Hohwieler E., 2018. Cluster identification of sensor data for predictive maintenance in a Selective Laser Melting machine tool. Procedia Manufacturing, 24, 60-65.
Wegener K., Gittler T., Weiss L., 2018. Dawn of new machining concepts: compensated, intelligent, bioinspired. Procedia CIRP, 77, 1-17.
Zhang J., Wang P., Yan R., Gao R.X., 2018. Deep Learning for improved system remaining life prediction. Procedia CIRP, 72, 1033-1038.
Zuev D., Kalistratov A., Zuev A., 2018. Machine Learning in IT Service Management. Procedia Computer Science, 145 (2018) 675–679.
Schmidt B., Wang L., 2018. Predictive Maintenance of machine tool linear axes: a case from manufacturing industry. Procedia Manufacturing, 17, 118-125.
Pandey S.K., Mishra R.B., Triphathi A.K., 2018. Software bug prediction prototype using bayesian network classifier: a comprehensive model. Procedia Computer Science, 132, 1412-1421.
Chakravorti N., Rahman M.M., Sidoumou M.R., (...), Gosewehr F., Wermann J., 2018. Validation of PERFoRM reference architecture demonstrating an application of data mining for predicting machine failure. Procedia CIRP, 72, 1339-1344.
Cortes C., Vapnik V., 1995. Support-vector networks. Machine Learning, 20 (3), 273-297.
Kotsiantis S. B., Zaharakis I., Pintelas P., 2007. Supervised machine learning: a review of classification techniques. Emerging artificial intelligence applications in computer engineering, 160, 3-24.
Marasco D.E., Kontokosta C.E., 2016. Applications of machine learning methods to identifying and predicting building retrofit opportunities. Energy and Buildings, 128, 431-441.
Lamirel J.-C., Al Shehabi S., Hoffmann, M., François C., 2003. Intelligent patent analysis through the use of a neural network. Proceedings of the ACL-2003 Workshop on Patent Corpus Processing, 20, 7–23.
Markovic D., Petkovic D., Nikolic V., Milovancevic M., Petkovic B., 2017. Soft computing prediction of economic growth based in science and technology factors, Phys. Stat. Mech. Appl. 465, 217–220.
Choi J., Jang D., Jun S., Park S., 2015. A predictive model of technology transfer using patent analysis. Sustainability, 7 (12), 16175-16195.
Searle, J. R., 1980. Minds, brains, and programs. Behavioral and Brain Sciences 3 (3), 417-457.
Samuel A. L., 1959. Some studies in machine learning using the game of checkers. BM Journal of

Research and Development archive, 3 (3), 210-229.

Praveenkumar T., Saimurugan M., Krishnakumar P., Ramachandran K.I., 2014. Fault diagnosis of automobile gearbox based on machine learning techniques. 12th Global Congress on Manufacturing and Management (GCMM), pp. 2092-2098. December 08-10, Vellore (India).

Fumeo E., Oneto L., Anguita D., 2015. Condition based maintenance in railway transportation system based on Big Data streaming analysis. INN Conference on Big Data, pp. 437-446. August 08-10, San Francisco (California, USA).

Ali S.M., Hui K.H., Hee L.M., Salman Leong M., Al-Obaidi M.A., Ali Y.H., Abdelrhman A.M., 2018. A comparative experimental study on the use of machine learning approaches for automated valve monitoring based on acoustic emission parameters. 3rd International Conference on Mechanical, Manufacturing and Process Plant Engineering (ICMMPE), UNSP 012032. November 22-23, Penang (Malaysia).

Li X., Wei L., He J., 2018. Design and implementation of equipment maintenance predictive model based on machine learning. 2nd annual international conference on cloud technology and communication engineering, UNSP 012001. August 17-19, Nanjing (China).

Koltsidopoulos P.A., Dawood T., Thies P.R., 2018. Data insights from an offshore wind turbine gearbox replacement. EERA DeepWind'2018, 15th Deep Sea Offshore Wind R&D Conference, art. no. 012003. January 17-19, Trondheim (Norway).

Elshrif M., Rizzo S.G., Betz F.D., Margineantu D.D., Zaki M.J., Chawla S., 2018. Embeddings for the Identification of Aircraft Faults (MERIT). 2018 IEEE International Conference on Prognostics and Health Management (ICPHM), art. no. 8448927. June 11-13, Seattle (Washington, USA).

Zanelli D., López E., Pavez C., Pedreros J., Jain J., Avaria G., Moreno J., Bora B., Davis S., Soto L., 2018. Analysis of deterioration in a plasma focus device. XX Chilean Physics Symposium, November 30-December 2, Santiago (Chile).

Katona A., Panfilov P., 2018. Building predictive maintenance framework for smart environment application systems. 29th International DAAM Symposium “Intelligent Manufacturing & Automation”, October 24-27, Zadar (Croatia).

Morales F.J., Reyes A., Cáceres N., Romero L.M., Benitez F.G., Morgado J., Duarte E., Martins T., 2017. Historical maintenance relevant information road-map for a self-learning maintenance prediction procedural approach. Building up Efficient and Sustainable Transport Infrastructure 2017 (BESTInfra2017), September 21-11, Prague (Czech Republic).

Sun C., Zhang Z., He Z., 2011. Research on bearing life prediction based on support vector machine and its application. 9th International conference on demage assessment of structure (DAMAS), July 11-13, Oxford (UK).

Mobley R. K., 2002. An Introduction to Predictive Maintenance. 2nd ed. Butterworth-Heineman.

APPENDIX

Table A: Extracted Documents From Scopus And Web of Science

ID	Source	Title	Authors	Years	Source Title	Cit.	Country
1	SCP	Research on bearing life prediction based on support vector machine and its application	Sun C., Zhang Z., He Z.	2011	Journal Of Physics: Conference Series	27	China
2	WoS	Fault diagnosis of automobile gearbox based on machine learning techniques	Praveenkumar T., Saimurugan M., Krishnakumar P., Ramachandran K.I.	2014	12th Global Congress On Manufacturing And Management (GCMM - 2014)	9	India
3	WoS	On combining machine learning with decision making	Tulabandhula T., Rudin, C.	2014	Machine Learning	1	USA
4	WoS	Machine Learning for Predictive Maintenance: A Multiple Classifier Approach	Susto G.A., Schirru A., Pampuri S., McLoone S., Beghi A.	2015	IEEE Transactions On Industrial Informatics	43	Italy
5	WoS	Condition Based Maintenance in Railway Transportation Systems Based on Big Data Streaming Analysis	Fumeo E., Oneto L., Anguita D.	2015	Inns Conference On Big Data 2015 Program	22	Italy
6	SCP	Hidden semi-markov models for predictive maintenance	Cartella F., Lemeire J., Dimiccoli L., Sahli H.	2015	Mathematical Problems In Engineering	14	Belgium
7	WoS	A Fault Diagnostic Method for Position Sensor of Switched Reluctance Wind Generator	Wang C., Liu X., Liu H., Chen Z.	2016	Journal Of Electrical Engineering & Technology	3	South Korea
8	WoS	A Comparative Study on Machine Learning Algorithms for Smart Manufacturing: Tool Wear Prediction Using Random Forests	Wu D., Jennings C., Terpenny J., Gao R.X., Kumara S.	2017	Journal Of Manufacturing Science And Engineering-Transactions Of The Asme	23	USA
9	SCP	Prognostics and Health Management: A Review of Vibration Based Bearing and Gear Health Indicators	Wang D., Tsui K.L., Miao Q.	2017	IEEE Access	23	Hong Kong, China
10	SCP	A Novel Multimode Fault Classification Method Based on Deep Learning	Zhou F., Gao Y., Wen C.	2017	Journal Of Control Science And Engineering	11	China
11	SCP	Ensemble machine learning and forecasting can achieve 99% uptime for rural handpumps	Wilson D.L., Coyle J.R., Thomas E.A.	2017	PLoS ONE	3	USA
12	WoS	Predictive Maintenance of Power Substation Equipment by Infrared Thermography Using a Machine-Learning Approach	Ullah I., Yang F., Khan R., Liu L., Yang H., Gao B., Sun K.	2017	Energies	1	China
13	SCP	Historical maintenance relevant information road-map for a self-learning maintenance prediction procedural approach	Morales F.J., Reyes A., Cáceres N., (...), Duarte E., Martins T.	2017	IOP Conference Series: Materials Science And Engineering	0	Portugal, Spain
14	WoS	A Robust Prescriptive Framework and Performance Metric for Diagnosing and Predicting Wind Turbine Faults Based on SCADA and Alarms Data with Case Study	Leahy, Kevin, Gallagher, Colm, O'Donovan, Peter, Bruton, Ken, O'Sullivan, Dominic T. J.	2018	Energies	6	Ireland
15	SCP	Data Insights from an Offshore Wind Turbine Gearbox Replacement	Koltsidopoulos Papatzimos, A., Dawood, T., Thies, P.R.	2018	Journal Of Physics: Conference Series	1	UK
16	SCP	Empirical Mode Decomposition Based Deep Learning for Electricity Demand Forecasting	Bedi, J., Toshniwal, D.	2018	IEEE Access	1	India
17	WoS	IDARTS - Towards intelligent data analysis and real-time supervision for industry 4.0	Peres, Ricardo Silva, Rocha, Andre Dionisio, Leitao, Paulo, Barata, Jose	2018	Computers In Industry	1	Portugal
18	WoS	Vehicle Remote Health Monitoring and Prognostic Maintenance System	Shafi, Uferah, Safi, Asad, Shahid, Ahmad Raza, Ziauddin, Sheikh, Saleem, Muhammad Qaiser	2018	Journal Of Advanced Transportation	1	Pakistan
19	WoS	A Comparative Experimental Study on the Use of Machine Learning Approaches for Automated Valve Monitoring Based on Acoustic Emission Parameters	Ali, Salah M., Hui, K. H., Hee, L. M., Leong, M. Salman, Al-Obaidi, M. A., Ali, Y. H., Abdelrhman,	2018	3rd International Conference On Mechanical, Manufacturing And	0	Malesia

			Ahmed M.		Process Plant Engineering (ICMMPE 2017)		
20	SCP	A smart framework for the availability and reliability assessment and management of accelerators technical facilities	Serio, L., Antonello, F., Baraldi, P., (...), Gentile, U., Zio, E.	2018	Journal Of Physics: Conference Series	0	Italy
21	SCP	Analysis of Deterioration in a Plasma Focus Device	Zanelli, D., López, E., Pavez, C., (...), Davis, S., Soto, L.	2018	Journal Of Physics: Conference Series	0	Chile
22	SCP	Building An Anomaly Detection Engine (ADE) for IoT Smart Applications	Mohamudally, N., Peermamode-Mohaboob, M.	2018	Procedia Computer Science	0	Mauritius
23	SCP	Building predictive maintenance framework for smart environment application systems	Katona, A., Panfilov, P.	2018	Annals Of DAAAM And Proceedings Of The International DAAAM Symposium	0	Austria, Russian Federation
24	SCP	Cluster identification of sensor data for predictive maintenance in a Selective Laser Melting machine tool	Uhlmann, E., Pontes, R.P., Geisert, C., Hohwieler, E.	2018	Procedia Manufacturing	0	Germany
25	SCP	Dawn of new machining concepts: Compensated, intelligent, bioinspired	Wegener, K., Gittler, T., Weiss, L.	2018	Procedia CIRP	0	Switzerland
26	SCP	Deep Learning for Improved System Remaining Life Prediction	Zhang, J., Wang, P., Yan, R., Gao, R.X.	2018	Procedia CIRP	0	USA
27	SCP	Design and Implementation of Equipment Maintenance Predictive Model Based on Machine Learning	Li, X., Wei, L., He, J.	2018	IOP Conference Series: Materials Science And Engineering	0	China
28	SCP	Embeddings for the Identification of Aircraft Faults (MERIT)	Elshrif, M., Rizzo, S.G., Betz, F.D., (...), Zaki, M.J., Chawla, S.	2018	2018 IEEE International Conference On Prognostics And Health Management, ICPHM 2018	0	Qatar, Italy, USA
29	WoS	Gaussian Process Operational Curves for Wind Turbine Condition Monitoring	Pandit, Ravi, Infield, David	2018	Energies	0	Scotland
30	WoS	K-PdM: KPI-Oriented Machinery Deterioration Estimation Framework for Predictive Maintenance Using Cluster-Based Hidden Markov Model	Wu, Zhenyu, Luo, Hao, Yang, Yunong, Lv, Peng, Zhu, Xinning, Ji, Yang, Wu, Bian	2018	IEEE Access	0	China
31	SCP	Machine Learning in IT Service Management	Zuev, D., Kalistratov, A., Zuev, A.	2018	Procedia Computer Science	0	China, Russian Federation
32	SCP	Predictive Maintenance of Machine Tool Linear Axes: A Case from Manufacturing Industry	Schmidt, B., Wang, L.	2018	Procedia Manufacturing	0	Sweden
33	SCP	Software Bug Prediction Prototype Using Bayesian Network Classifier: A Comprehensive Model	Pandey, S.K., Mishra, R.B., Triphathi, A.K.	2018	Procedia Computer Science	0	India
34	SCP	TELS: A Novel Computational Framework for Identifying Motif Signatures of Transcribed Enhancers	Kleftogiannis, D., Ashoor, H., Bajic, V.B.	2018	Genomics, Proteomics And Bioinformatics	0	UK, USA, Saudi Arabia
35	SCP	Validation of PERFoRM reference architecture demonstrating an application of data mining for predicting machine failure	Chakravorti, N., Rahman, M.M., Sidoumou, M.R., (...), Gosewehr, F., Wermann, J.	2018	Procedia CIRP	0	UK, Germany
36	SCP	Data analysis and feature selection for predictive maintenance: A case-study in the metallurgic industry	Fernandes, M., Canito, A., Bolón-Canedo, V., (...), Praça, I., Marreiros, G.	2019	International Journal Of Information Management	0	Portugal, Spain
37	SCP	Data science for vibration heteroscedasticity and predictive maintenance of rotary bearings	Lee, C.-Y., Huang, T.-S., Liu, M.-K., Lan, C.-Y.	2019	Energies	0	Taiwan