TECHNOLOGY IDENTIFICATION IN INDUSTRY 4.0 WITH THE HELP OF TEXT MINING

BYSANI SAI AMARNATH, B R VENKATESH, CH DIVESH

Objective: The point of this paper is to recognizable proof of highlighted innovation in Industry 4.0.

approach: The most common way of distinguishing the dynamic/logical text with the assistance of text mining to investigate in this field. Suppositions of own iterative strategies for examine the logical texts we propose, with the utilization of R language, tokenization, lemmatization, n-grams and correspondence examination. The suspicions of the proposed strategy were utilized to dissect the 40 most frequently cited articles filed in the Web of Science.

research: Based on the got results, 4 driving innovations were recognized. These are Cloud Computing, Internet of Things, Cyber-actual System and Big Data. Creativity: This paper tells a unique technique for distinguishing the highlighted advancements that are utilized in Industry 4.0. The strategies depends on text mining.

Keywords: Text mining, industry 4.0, data and correspondence innovation, classifying and design.

INTRODUCTION

The possibility of "Industry 4.0" is connected to World Vision, which will arrive at improvement stage, which will be characterized as the Fourth Industrial Revolution later on. The universal digitization constrained by market globalization will build the progressions in the worldwide economy. The influences are caused for clients of specific assistance and design of market. Changes in the help market and creation are as of now noticeable today and are supposed to increment later on. As per the sources, the most extreme changes can happen in information driven enterprises, for example, new advances, media, amusement, monetary administrations, broadcast communications and retail benefits. Simultaneously, incredible advancement open doors are being underlined for organizations that can adjust to the changing climate and foster items. In this way, an ever increasing number of chiefs are zeroing in on 's recently sent off IT administrations and arrangements that in a roundabout way affect the market. Source is an exchange fair, industry paper, well known logical articles, business gatherings, and meetings with specialists, makers, specialist co-ops, and researchers. , And corporate specialists share data, experience, and make dreams for local area wide turn of events. A greater amount of the gatherings perhaps associated direct (eg Future Manufacturing Technologies, Industries 4.0 and so forth).

DEFINITION AND PLANS OF INDUSTRY 4.0

Industry 4.0 has numerous definitions. Industry4.0 has been characterized as an overall idea that empowers producing with components of strategic insight utilizing innovations, for example, the web of things, distributed computing, and huge information. With the web and supporting advancements, for example, the implanted frameworks going about as the spine, actual articles, human entertainers, smart machines and so on process across product offerings and hierarchical limits to shape another kind of insightful, arranged and lithe worth chain. In the mean time, Kohler and Weisz have characterized industry 4.0 as another way to deal with controlling the creation cycle by giving constant synchronization of activities and empowering the development of unfiled and tweaked items in this multitude of definitions, the idea of industry 4.0 depends on the utilization of new innovation that empowers continuous data stream, control and coordination of cycles through the creation framework

Technology identification in industry 4.0
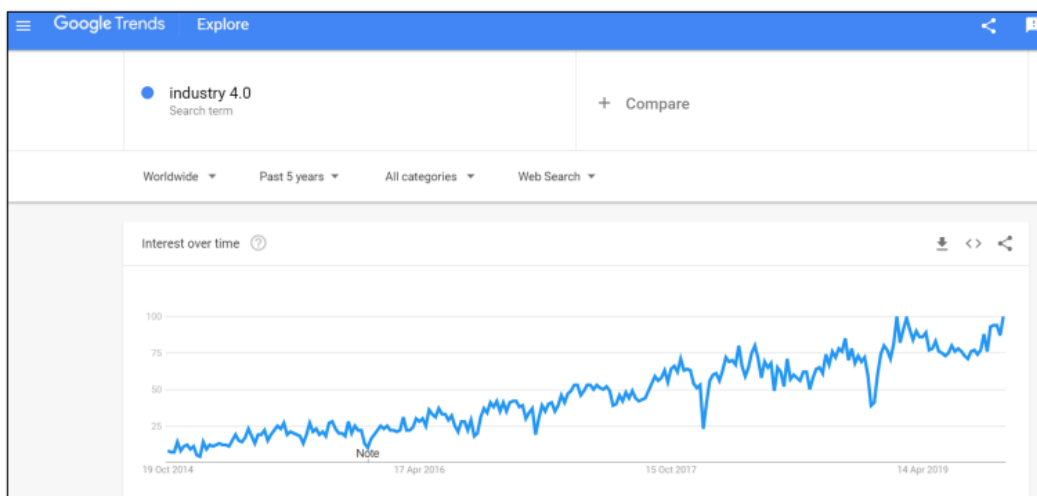
This converts into significantly more prominent number of definitions, terms and plans characterizing the Industry 4.0. Since in the space of Industry 4.0 advancements assume a critical part, it was chosen to recognize them based on plans accessible on the Internet. For the hunt "Industry 4.0 advancements" countless different plans introducing innovations in the space of Industry 4.0 was shown.

**Examine the plans concerning advances in Industry 4.0**

Table 1.
Analysis of schemes concerning technologies in Industry 4.0

ID.	Technology	Fig. 1	Fig. 2	Fig. 3	Fig. 4	Fig. 5	Fig. 6
1	Cloud Computing	x	x	x	x		x
2	Augmented Reality	x			x	x	x
3	Internet of Things	x	x	x	x		
4	Cybersecurity		x		x	x	
5	Additive manufacturing			x	x	x	
6	Big Data Analytics	x	x			x	
7	3D Printing	x	x				x
8	Big Data			x	x		x
9	Advanced Robotics	x	x				
10	Cyber Physical System			x			x
11	Autonomous Robots				x	x	
12	Simulation				x	x	
13	System integration				x		
14	Horizontal and vertical system integration					x	
15	The industrial Internet of Things					x	
16	The cloud					x	
17	System Security						X
18	Humanoid Robots						X
19	Location Detection	x					
20	Machine learning			x			
21	Cobot			x			
22	Smart Sensor	x					

The investigation of definitions and plans is one of the habits of distinguishing proof of advancements in the space of Industry 4.0. Notwithstanding the way that it is a tedious interaction, with an enormous number of examined plans a rundown of the most often named innovations can be acquired. A blemish of this technique is the way that plans are much of the time duplicated by many creators and they don't consider the importance and improvement of a given innovation in the field of Industry 4.0. It was expected, that the more changes occur in a given innovation and its utilization is the area of examination inside Industry 4.0, the more articles there are. Therefore, it was concluded that a complementation of this investigation can be an investigation of logical texts and check how much of the time specific words and expressions related with advances show up in them.
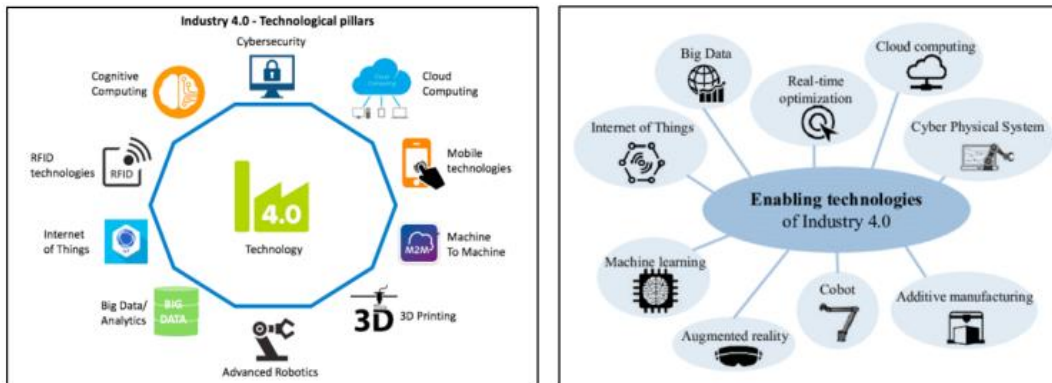
Five characteristic stages
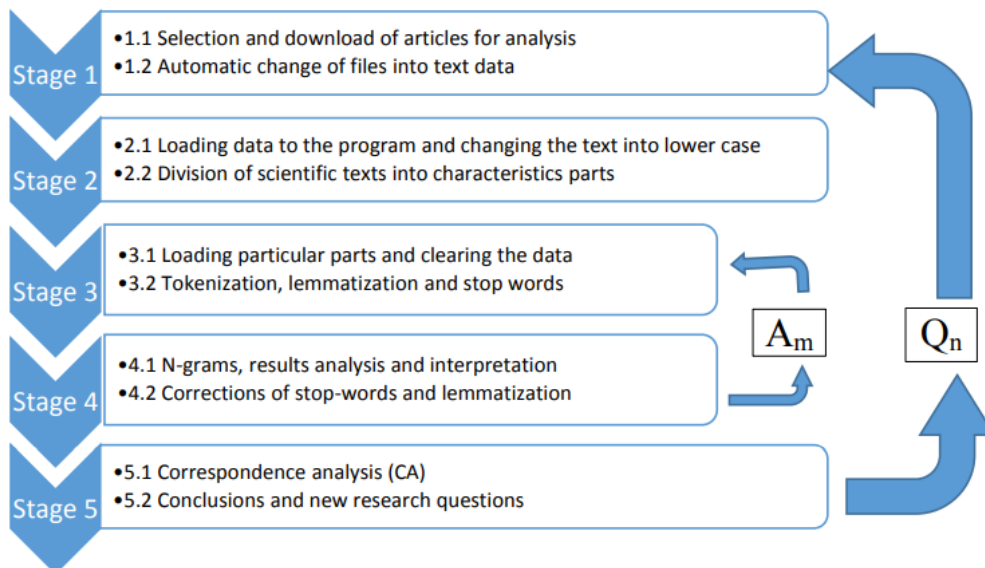
- ➢ The stage one is utilized for choosing and downloading of articles for examining and furthermore changing the records consequently into text information.

- ➢ The stage two is tied in with stacking the entire gathered information in text documents into that scientific apparatus. for that reason. Rstudio was utilized.

- ➢ At stage three, POS-tag was overlooked deliberately (Nakagawa, 2007). It shapes a piece of NLP and 2 worries denoting the words with a fitting piece of the discourse (for instance: thing, 3 descriptor, action word and so on.). It was concluded that it isn't required with the end goal of this exploration

- ➢ The stage four was tied in with making recurrence measurements of the pre-owned single, twofold or triple 5 words in the text with the utilization of n-grams.

➤ The stage five was tied in with dissecting the got data and making the new 9 information on the broke down reports with the utilization of chosen strategy for information examination.

OVERVIEW OF ARTICLE

Understanding the subject of Industry 4.0 with the assistance of text mining is a fascinating method for enhancing your insight around here. It requires significantly a ton of investment. As a matter of some importance, it comes from being required to compose the necessary scripts and change them while examining unstructured text. In view of just 40 articles, we had the option to determine the methods that are happening around here. You can likewise attempt to gathering and look at them. This influences researchers as well as experts and individuals who pursue various choices in light of information sources from the social and financial climate, notwithstanding that because of the examination of logical text utilizing text mining, you can pose new inquiries about your requirements in view of the outcomes acquired with them. consequently, it is vital to foster a basic and helpful instrument and strategy 4304 for investigating unstructured text records. In this survey we have depicted suspicion of the text mining strategy used to decipher records connected with industry 4.0 was made, because of her a rundown of important articles about industry 4.0 the pdf document was switched over completely to a txt record the information was tidies up and the text was parted into unmistakable parts .the most ordinarily utilized set words and barely any words. In light of correspondence examination, chose advances and related articles were recognized and gathered. Text mining and steps while dissecting logical articles utilizing step related constraints ought to likewise be thought of. They are fundamentally connected with plans ,tables and drawings in pdf records that are skipped yet contain a great deal of helpful data. Results from definition examination and text mining have permitted to recognize four significant advances today they are distributed computing ,IOT cps(cyber actual framework) and enormous information .text mining examination is accessible to specialists in all disciplines inside industry4.0 and will be fascinating supplement to explore strategies, for example, research ,meetings ,perceptions and writing investigation, later on, we want to refine the suppositions of this strategy and investigation more logical texts on industry 4.0.

CONCLUSION

Word	Freq.
industry	1917
system	1601
data	1413
manufacture	1245
production	1033
process	1009
technology	929
product	814
industrial	734
information	697
time	653
smart	634
base	618
machine	556
model	534
company	515
network	491

Using the assumptions and steps offered inside the bankruptcy above a search containing 40 files on industry 4.0 became done. The first iteration produced a list of the most frequently used words and phrases related to industry 4.0 a total of 8,782 words were determined, of which 3,846 (about 44%) were used only once in all documents the most common words used were industry, systems, data, manufacturing, production, processes, technology, and so on.

REFERENCES

1. [1] R. Sagayam, A survey of text mining: Retrieval, extraction and in- dexing techniques, International Journal of Computational Engineering Research, vol. 2, no. 5, 2012.

2. [2] N. Padhy, D. Mishra, R. Panigrahi et al., "The survey of data mining applications and feature scope," arXiv preprint arXiv:1211.5723, 2012.

3. [3] W. Fan, L. Wallace, S. Rich, and Z. Zhang, "Tapping the power of text mining," Communications of the ACM, vol. 49, no. 9, pp. 76–82, 2006.

4. [4] S. M. Weiss, N. Indurkhya, T. Zhang, and F. Damerau, Text mining: predictive methods for analyzing unstructured information. Springer Science and Business Media, 2010.

5. [5] S.-H. Liao, P.-H. Chu, and P.-Y. Hsiao, "Data mining techniques and applications–a decade review from 2000 to 2011," Expert Systems with Applications, vol. 39, no. 12, pp. 11 303–11 311, 2012.

6. [6] W. He, "Examining students online interaction in a live video streaming environment using data mining and text mining," Computers in Human Behavior, vol. 29, no. 1, pp. 90–102, 2013.

7. [7] G. King, P. Lam, and M. Roberts, "Computer-assisted keyword and document set discovery from unstructured text," Copy at http://j. mp/1qdVqhx Download Citation BibTex Tagged XML Download Paper, vol. 456, 2014.

8. [8] N. Zhong, Y. Li, and S.-T. Wu, "Effective pattern discovery for text mining," IEEE transactions on knowledge and data engineering, vol. 24, no. 1, pp. 30–44, 2012.

9. [9] A. Henriksson, H. Moen, M. Skeppstedt, V. Daudaravicˇius, and M. Duneld, "Synonym extraction and abbreviation expansion with ensembles of semantic spaces," Journal of biomedical semantics, vol. 5, no. 1, p. 1, 2014.

10. [10] B. Laxman and D. Sujatha, "Improved method for pattern discovery in text mining," International Journal of Research in Engineering and Technology, vol. 2, no. 1, pp. 2321–2328, 2013.

11. [11] C. P. Chen and C.-Y. Zhang, "Data-intensive applications, challenges, techniques and technologies: A survey on big data," Information Sci- ences, vol. 275, pp. 314–347, 2014.

12. [12] R. Rajendra and V. Saransh, "A Novel Modified Apriori Approach for Web Document Clustering," International Journal of Computer Applications, pp. 159– 171, 2013.

13. [13] K. Sumathy and M. Chidambaram, "Text mining: Concepts, applica- tions, tools and issues-an overview," International Journal of Computer Applications, vol. 80, no. 4, 2013.

14. [14] P. J. Joby and J. Korra, "Accessing accurate documents by min- ing auxiliary document information," in Advances in Computing and Communication Engineering (ICACCE), 2015 Second International Conference on. IEEE, 2015, pp. 634–638.

15. [15] Z. Wen, T. Yoshida, and X. Tang, "A study with multi-word feature with text classification," in Proceedings of the 51st Annual Meeting of the ISSS-2007, Tokyo, Japan, vol. 51, 2007, p. 45.

16. [16] V. Gupta and G. S. Lehal, "A survey of text mining techniques and applications," Journal of emerging technologies in web intelligence, vol. 1, no. 1, pp. 60–76, 2009.

17. [17] R. Agrawal and M. Batra, "A detailed study on text mining techniques," International Journal of Soft Computing and Engineering (IJSCE) ISSN, pp. 2231– 2307, 2013.

18. [18] D. S. Dang and P. H. Ahmad, "A review of text mining techniques associated with various application areas," International Journal of Science and Research (IJSR), vol. 4, no. 2, pp. 2461–2466, 2015.

19. [19] R. Steinberger, "A survey of methods to ease the development of highly multilingual text mining applications," Language Resources and Evaluation, vol. 46, no. 2, pp. 155–176, 2012.

20. [20] A. M. Cohen and W. R. Hersh, "A survey of current work in biomedical text mining," Briefings in bioinformatics, vol. 6, no. 1, pp. 57–71, 2005.

21. [21] E. A. Calvillo, A. Padilla, J. Mun̅oz, J. Ponce, and J. T. Fernan- dez, "Searching research papers using clustering and text mining," in Electronics, Communications and Computing (CONIELECOMP), 2013 International Conference on. IEEE, 2013, pp. 78–81.

22. [22] B. L. Narayana and S. P. Kumar, "A new clustering technique on text in sentence for text mining," IJSEAT, vol. 3, no. 3, pp. 69–71, 2015.

23. [23] B. A. Mukhedkar, D. Sakhare, and R. Kumar, "Pragmatic analysis based document summarization," International Journal of Computer Science and Information Security, vol. 14, no. 4, p. 145, 2016.

24. [24] R. Al-Hashemi, "Text summarization extraction system (tses) using extracted keywords." Int. Arab J. e-Technol., vol. 1, no. 4, pp. 164– 168, 2010.

25. [25] I. H. Witten, K. J. Don, M. Dewsnip, and V. Tablan, "Text mining in a digital library," International Journal on Digital Libraries, vol. 4, no. 1, pp. 56–59, 2004.

26. [26] S. Ayesha, T. Mustafa, A. R. Sattar, and M. I. Khan, "Data mining model for higher education system," Europen Journal of Scientific Research, vol. 43, no. 1, pp. 24–29, 2010.

27. [27] A. Henriksson, J. Zhao, H. Dalianis, and H. Bostro ̈m, "Ensembles of randomized trees using diverse distributed representations of clinical events," BMC Medical Informatics and Decision Making, vol. 16, no. 2, p. 69, 2016. 28. [28] I. Alonso and D. Contreras, "Evaluation of semantic similarity metrics applied to the automatic retrieval of medical documents: An umls approach," Expert Systems with Applications, vol. 44, pp. 386–399, 2016.

29. [29] C. Ding and H. Peng, "Minimum redundancy feature selection from microarray gene expression data," Journal of bioinformatics and com- putational biology, vol. 3, no. 02, pp. 185–205, 2005.

30. [30] Y. Zhao, "Analysing twitter data with text mining and social network analysis," in Proceedings of the 11th Australasian Data Mining and Analytics Conference (AusDM 2013), 2013, p. 23.

31. [31] F. Fatima, Z. W. Islam, F. Zafar, and S. Ayesha, "Impact and usage of internet in education in pakistan," European Journal of Scientific Research, vol. 47, no. 2, pp. 256–264, 2010.

32. [32] R. Sharda and M. Henry, "Information extraction from interviews to obtain tacit knowledge: A text mining application," AMCIS 2009 Proceedings, p. 283, 2009.

33. [33] H. Solanki, "Comparative study of data mining tools and analysis with unified data mining theory," International Journal of Computer Applications, vol. 75, no. 16, 2013.

34. [34] A. Kumaran, R. Makin, V. Pattisapu, and S. E. Sharif, "Automatic extraction of synonymy information:-extended abstract," OTT06, vol. 1, p. 55, 2007.

35. [35] A. Kaklauskas, M. Seniut, D. Amaratunga, I. Lill, A. Safonov, N. Vatin, J. Cerkauskas, I. Jackute, A. Kuzminske, and L. Peciure, "Text analytics for android project," Procedia Economics and Finance, vol. 18, pp. 610–617, 2014.

36. [36] N. Samsudin, M. Puteh, A. R. Hamdan, and M. Z. A. Nazri, "Immune based feature selection for opinion mining," in Proceedings of the World Congress on Engineering, vol. 3, 2013, pp. 3–5.