The Research Hotspot and Path Evolution of Big Data in Healthcare Industry: Visualized Analysis Based on Cite Space

Chengke Zhu^{1,2}, Xiaofeng Zhao^{3*}

¹ College of Economics and Management, Hefei University, Hefei230601, Anhui, China ² Key Laboratory of Financial Big Data, Hefei University, Hefei230601, Anhui, China ³School of Economics, Guizhou University of Commerce, Guiyang550014, Guizhou, China *Corresponding author.

Abstract

Along with the swift development of big data, the application fields of big data have become more and more broad, as well as the fusing of big data and healthcare. In this paper, CiteSpace, the bibliometrics software, was used for analyzing the research status and hotspots of big data in healthcare fields. Taking big data and healthcare as the keywords, the core database of Web of Science from 2011 to 2020 has been retrieved. According to author/country/organization co-authorship analysis and author/journal/reference co-cited analysis, the research status of big data in healthcare was discussed. Through the keywords co-occurrence network and cluster analysis, research hotspot and path evolution of big data in healthcare were also studied.

Keywords: big data, healthcare, CiteSpace, path evolution, research hotspots

I.Introduction

Due to the swift development of big data, it will be of great value in healthcare. In healthcare fields, big data develops with features of huge volume and multiple styles. At present, how to extract useful information via data mining from giant data occurred in healthcare process is the urgent problem that healthcare big data is facing. Nowadays, big data is applied in the following aspects in the healthcare area.

First, big data serves in the clinical aspect. Electronic Health Records (EHR) and other data information systems benefit medical personnel with more effective diagnostic services. They simplify diagnosis process efficiently, save much time for patients and medical personnel, and increase the accuracy of clinical diagnosis[1]. Personalized healthcare is achieved by big data-driven technology, and then healthcare cost and the admission rate is lowered[2].

Second, big data plays a crucial role in biopharming. Since drug development is based on the unified application of chemistry, biology and a dozen subjects, records of research data and material can be up to thousands of pages for each drug averagely. Big data technology can be applied to analyzing the demands of drugs, and relevant information will be fed back to drug research and development institutions. Predicting and determining the direction of drug research with efficient information obtained through data mining, big data can provide a basis for biopharming. Therefore, big data can help take the effects of drug manufacture and treatment in biopharming[3].

Besides, along with the fusing development of big data and Internet of Things, electronic products about health services are rushing into the market, such as wearable devices[4]. This kind of equipment using big data technology to analyze data collected and give feedback in time. For example, when an intellectual product finds out abnormal data of blood fat by comparing with data, it will alert the users. Then people can make judgement and adjustment according to the information.

Big data has already had certain basis in medical treatment and it is worth to summarize. However, the existing literature demonstrated the impact of big data in health treatment mainly from the qualitative point of view through

literature review and regular statistical methods. Those literatures performed positively in the healthcare fields, but only to a limited extent. That is why it is difficult to fully understand the research status of big data in healthcare field. It is meaningful and essential for future research to give it insight into the research status of big data in healthcare field. In this way, not only the development trend can be understood, but the reference for discovering new problems can be provided as well.

II. Data Resource and Method

2.1 Data Resource

In this paper, the core database of Web of Science (WoS) has been selected in order to reflect research status and hotspots of big data in the healthcare area. Web of Science (WoS) is a database used for scientific document analysis frequently with a wide range of information and positive perception[5, 6]. Retrieving in WoS database in December 30 2020, with keywords of "big data" and "healthcare", limiting document type as "Article", setting up language as "English" and time range "2011 to 2020", totally 1204 eligible articles were obtained after data cleaning and similarity identification.

2.2 Research Method

Mapping knowledge domains, the new method in scientometrics, has developed rapidly in recent years. The evolution and relationships of scientific knowledge can be presented by this method, and development history and frontier in research fields can be observed intuitively. CiteSpace, developed by Professor Chaomei Chen, is a powerful information visualization software. Based on the theory of co-citation analysis, CiteSpace processes visualization analysis on literature in particular fields in order to figure out the key path of the status and evolution withina subject area. Aided by CiteSpace, this paper performed visualization analysis on issues such as numbers of papers published, journals of papers published, high productivity authors, institutions, authors cited, and research hotspots for exploring research hotspots and evolution path of big data in healthcare.

III. Statistical Analysis of Literature

3.1 Time Distribution of Literature

The quantity of articles on a particular subject is a significant indicator for studying the research status. Fig 1 shows the changes in number of papers about big data in healthcare published in the WoS database in the last decade. It can be seen from Fig 1 that the number of papers published was quite small before 2013. However, the quantity began to grow fast since 2014. Overall, the quantity of papers published shows an exponential growth trend.



The curve-fitting equation can be found from Fig 1:

$$y = 6.5403e^{0.4159x}, R^2 = 0.9737$$

In this equation, y presents the numbers of papers published each year and x presents the year. R^2 equals to 0.9737, which represents a high fitting degree, meaning that the number of papers about big data in healthcare will continue to increase in the following time. Therefore, big data has been proved to be a hotspot in the healthcare research field.

3.2 Countries co-authorship analysis

In order to reflect the cooperative network between countries of big data in healthcare, the author set node to country and it came out with co-countries network, shown as Fig 2. In this co-countries network, there are 83 nodes and 530 wires. The country with the largest quantity of articles is USA, and the number is 339. China ranks second, with the number of 200. Followed are England and India, both exceed 100.



Fig 2: Countries co-authorship analysis of literature

3.3 Institution co-authorship analysis

In order to reflect the cooperative network between institutions of big data in healthcare, the author set node to "institution" and it came out with institution co-authorship network, shown as Fig 3. there are 40 institutions which have published more than five papers in theWoS database. Among those institutions, King Saud University ranks first with 31 papers, followed by Huazhong University of Science & Technology, Stanford University and Kyonggi University, and the numbers of papers published are 20, 17 and 17, respectively. When focusing on this network, it can be noticed that network between nodes is quite dense, which means universities work closely. In the case of cooperation between individual universities, Stanford University and Harvard Medical School have closer cooperation with other institutions.



Fig 3: Institution co-authorship analysis of literature

3.4 Author co-authorship analysis

According to Price Law [7], $m = 0.749(n_{max})^{0.5}$, the author analyzed authors who have published a certain number of articles in a big data research field. In this equation, m presents the numbers of authors with a high quantity of papers, and n presents the maximum number of papers of the author. Among the collected literature, the maximum number of papers published is 17, which means m ≈ 3 . There are totally 23 authors publishing more than three papers. Cooperation between authors in the areas of big data in healthcare is shown in Fig 4. It can be noticed that the cooperation between individual authors is not close, neither is the communication. The cooperation between small teams takes up the majority. In Fig 4, there are 335 nodes and 297 links. The team of Kyunghyang Chung and that of M Shamim Hossain are relatively large.



Fig 4: Author co-authorship analysis based on WoS

rank	Freq	Author
1	17	Kyungyong Chung
2	10	M Shamim Hossain
3	7	Min Chen
4	7	Ghulam Muhammad
5	6	Joochang Kim
6	5	Joel J P C Rodrigues
7	5	Jun Yang
8	5	Mohamed Elhoseny
9	5	Angela Brand
10	4	Mohsen Guizani
11	4	Victor Chang
12	4	Shawn N Murphy
13	4	Neeraj Kumar
14	4	Yin Zhang
15	4	Yichuan Wang
16	4	GunasekaranManogaran
17	4	V Vijayakumar
18	4	Matthias Reumann
19	4	Rongxing Lu
20	3	AbdulhameedAlelaiwi
21	3	Hyun Yoo
22	3	Mohammad Mehedi Hassan
23	3	Sandeep K Sood

From Table 1, it can be seen that Kyunghyang Chung ranks first among those authors with articles published as

many as 17 in WoS database. M Shamim Hossain, Min Chen and Ghulam Muhammad come second, with 10, 7 and 7 articles respectively and those three authors belong to the same research team.

Kyunghyang Chung has taken his professorship atKyonggi University since 2017. He mainly studies data mining, healthcare and biomedical. Chung was voted as a highly cited author by Clarivate Analytics in 2017. Under his leadership, Chung's team has published a series of essays about big data used in healthcare. A kind of chatbot for healthcare service is proposed, which used cloud computing technology[8]. This model predicts body activities of users according to data of weather, temperature, humidity, ultraviolet, dust, probability index of skin disease, HRmean and sleeping time, and provides personalized healthcare service.

M Shamim Hossainraised a healthcare system connecting with emotional awareness. This system monitors patients' emotions by capturing speech and image signals of patients. Once the emotional signal turns into "pain", paramedics will carry out a thorough inspection of the patients. The author tested and verified the accuracy of this system standing at 99.87%. Chen et. al.[9]came up with an urban healthcare big data system which can provide healthcare service for citizens in the area of respiratory disease.

IV. Statistical Analysis of Literature

Author Co-citation analysis was raised by henry Small, the American intelligence scientist, in 1973[10]. It analyzes and explores relationships between literatures and reflect discipline structure with highly co-cited literatures. Complex co-cited networks can be simplified to several clusters and visually presented via clustering analysis.

4.1 Journal Co-citation Analysis

Importing data of related literatures into CiteSpace and setting node as "Cited Journal", a mapping knowledge domain of journal co-citation was generated as Fig 5.

According to Fig 5, the range of journals co-cited in the area of big data in healthcare is literally wide. Top comprehensive journals like Nature and Science are on the list, so are those international recognized medical journals such as LANCET, J AM MED INFORM ASSN, JAMA-J MED ASSOC, NEW ENGL J MED and PLOS ONE. There are also journals of computer science like LECT NOTES COMPUT SC and EXPERT SYST APPL.

Among those journals, JAMA-J AM MED ASSOC, PLOS ONE, NEW ENGL J MED, FUTURE GENER COMP SY and IEEE ACCESS are cited over 200 times. LECT NOTES COMPUT SC, J AM MED INFORM ASSN and LANCET come second. Those journals are significant resources of literatures about international research of related subjects.



Fig 5: Journal co-citation analysis

4.2 Literature Co-citation Analysis

By setting node to "Reference", analysis map of co-cited literatures was generated as Fig 6. Raghupathi[11]has the highest frequency of co-citation, reaching 98 times, followed by[1], with 54 times and 43 times.



Fig 6: Literature co-citation analysis

CONVERTER MAGAZINE

Table 2 lists the top 10 literatures co-cited, and they play important roles in research of this area. Raghupathi[11] explored the advantages, frames and methods of big data in healthcare field and analyzed challenges of medical big data research. Murdoch[1] studied on applications of big data in healthcare services and tried to increase the efficiency and quality of medical services by using related data between patients and doctors. Bates[12]believed that big data analysis can improve nursing ability and reduce healthcare costs. The result of Bates' research benefits the reform of healthcare industry in policy implications. Islam[13]analyzed the development of healthcare technology based on IoT, discussed how big data affects in healthcare area and provided suggestions for healthcare research based on IoT in the future. Chen[14] introduced the backgrounds of big data in IoT and healthcare.

rank	Frequency	Centrality	Author	Year	Source
1	98	0.06	Raghupathi W	2014	HEALTH INF SCI SYST
2	54	0.03	Murdoch TB	2013	JAMA-J AM MED ASSOC
3	43	0.11	David W. Bates	2014	HEALTH AFFAIR
4	35	0.06	Islam SMR	2015	IEEE ACCESS
5	30	0.01	Chen HC	2012	MIS QUART
6	28	0.08	Chen M	2014	MOBILE NETW APPL
7	25	0.06	Andreu-Perez J	2015	IEEE J BIOMED HEALTH
8	24	0.02	Manyika J	2011	BIG DATA NEXT FRONTI
9	23	0.01	Gandomi A	2015	INT J INFORM MANAGE
10	23	0.01	Obermeyer Z	2016	NEW ENGL J MED

Table 2	Top 10	literatures	co-cited
---------	---------------	-------------	----------

Then via clustering literatures, 487 nodes, 1715 wires and 69 clusters have been found. After removing small clusters and using LLR algorithm, 7 major clusters remained, and cluster map is shown in Fig 7. More detailed information are listed in Table 3.In data analysis, Silhouette value reflects whether the evaluation of clusters is good or not. Considering the Silhouette values all reach over 0.8, those clusters are rational.

Table 5 Wajor clusters of netrature analysis						
Cluster	Size	Silhouette	mean(Year)	Label (LLR)		
0	96	0.843	2013	healthcare		
1	48	0.865	2014	fog-driven iotehealth		
2	45	0.824	2016	cloud-based framework		
3	42	0.917	2015	sustainable health monitoring		
4	40	0.805	2017	healthcare internet		
5	25	0.99	2016	medicine management		
6	20	0.978	2014	one-size fit		

Table 3 Major clusters of literature analysis



Fig 7: Cluster map of literature

Furthermore, using method of [15], cluster 0, 1 and 2 with citation burst are analyzed. Table 4 to 6 show the top 5 cited references in the clusters listed above.

frequency	Cited references
98	RaghupathiW(2014)Big data analytics in healthcare: promise and potential. HEALTH INF SCI SYST,V2
54	Murdoch TB(2013)The Inevitable Application of Big Data to Health Care. JAMA-J AM MED ASSOC. VOL309, P1351
43	Bates DW(2014)Big Data In Health Care: Using Analytics To Identify And Manage High-Risk And High-Cost Patients. HEALTH AFFAIR.VOL33.P1123
30	Chen HC(2012)Business Intelligence and Analytics: From Big Data to Big Impact. MIS QUART.VOL36.P1165
25	Andreu-Perez J(2015)Big Data for Health. IEEE J BIOMED HEALTH.VOL19.P1193

Table 4 cited references of cluster 0 on healthca

Cluster 0, healthcare, is the largest cluster including 96 articles which is shown in Table 4. It mainly studies the applications of big data in healthcare industry. As we have discussed[1, 11, 12, 16] previously and won't be repeated here. Chen[9] thought Business Intelligence and Analysis (BI&A) is the key research area since issues related to data need to be solved during economic development.

Table 5 cited references of cluster 1 on fog-driven to T health				
frequency	Cited references			
35	Islam SMR(2015)The Internet of Things for Health Care: A Comprehensive Survey. IEEEACCESS.VOL3.P678			

CONVERTER MAGAZINE

Volume 2021, No. 5

20	Gubbi J(2013).Internet of Things (IoT): A vision, architectural elements, and future directions.FUTURE GENER COMP SY.VOL29.P1645
20	Hashem IAT(2015)The rise of "big data" on cloud computing: Review and open research issues.INFORM SYST.VOL47.P98
20	Hossain MS(2016)Cloud-assisted Industrial Internet of Things (IIoT) – Enabled framework for health monitoring. COMPUT NETW.VOL101.P192
13	Alemdar H(2010)Wireless sensor networks for healthcare: A survey.COMPUT NETW.VOL54.P2688

Cluster 1, including 48 articles, is fog-driven IoT health which mainly studies the effect of new technology like IoT and cloud computing in healthcare services shown in Table 5. Gubbi[17] thought IoT has already moved out of the startup stage, and will be the next innovational technology in Internet. Gubbi also put forward the prospect that global IoT will be fulfilled with cloud-centered background, and discussed the key implementation technology and application areas which can push the advancement of research of IoT. Hashem[18] introduced the rise of big data in cloud computing. How to deal with big data has become a challenging task under limited time as big data produced by cloud computing is getting bigger and bigger. To solve this issue, large computing infrastructure is demanded for ensuring data processing and analyzing and the author believes cloud computing is a kind of large computing technology.

frequency	Cited references
19	CatarinucciL(2015)An IoT-Aware Architecture for Smart Healthcare Systems. IEEE INTERNET THINGS.VOL2.P515
13	Rahmani AM(2018)Exploiting smart e-Health gateways at the edge of healthcare Internet-of-Things: A fog computing approach.FUTURE GENER COMP SY.VOL78.P641
12	Farahani B(2018)Towards fog-driven IoT eHealth: Promises and challenges of IoT in medicine and healthcare.FUTURE GENER COMP SY.VOL78.P659
10	Botta A(2016)Integration of Cloud computing and Internet of Things: A survey FUTURE GENER COMP SY.VOL56.P684
9	Abawajy JH(2017)Federated Internet of Things and Cloud Computing Pervasive Patient Health Monitoring System.IEEE COMMUN MAG.VOL55.P48

Table 6 cited references of cluster 2 on cloud-based framework

From Table 6, we can find that cluster 2 is cloud-based framework, including 42 articles. This cluster mainly explores new applications of cloud computing and IoT in healthcare industry based on the previous research. Catarinucci[19] thought IoT has taken a big step and promoted a lot of valuable application. The author raised a kind of biomedical equipment automatically monitoring and tracking hospitals and patients. This equipment can send data to control center for medical staff's visiting. It is a great progress compared with past technologies. Rahmani[20] believed IoT makes it possible to develop more intelligent healthcare solutions. In most medical systems based on IoT, there should be a bridge contact, also known as gateway, between sensor infrastructure and the Internet. On this basis, this paper put forward an intelligent electronic health gateway. Clinically, if intelligent to face various challenges such as energy efficiency and reliability.

4.3 Author Co-citation Analysis

In order to determine the authors cited frequently, node was set to Cited Author. Cited map generated is shown as Fig 8.



Fig 8: Author co-cited analysis

rank	Freq	Centrality	Author				
1	100	0.04	Raghupathi W				
2	83	0.04	Chen M				
3	71	0.14	World Health Organization				
4	69	0.06	Zhang Y				
5	53	0.11	Bates DW				
6	52	0.03	Murdoch TB				
7	41	0.01	Wang Y				
8	41	0.03	Hossain MS				
9	36	0.05	Wang YC				
10	35	0.05	Islam SMR				

Table 7 Top 10 co-cited authors

The top ten co-cited authors are shown in Table 7. Raghupathi, professor in information technology area, has teaching and research interests including analytics, artificial intelligence and deep machine learning algorithm which are applied to finance and corporate governance, culture and ethics, and healthcare fields. Professor Raghupathi has published over 60 articles in his research area. Min Chen is the professor in Huazhong University of Science and Technology, of whom the articles have been cited over 24,800 times by Google Scholar. The number of citations for individual article of Professor Chen was more than 2960 and H-index equals to 78. According to the data of top 1000 scientists in computer science area counted in May, 2020, Min Chen ranked 18 in China and has been voted as highly cited author by Clarivate Analytics from 2018 to 2020. Yin Zhangwas voted as global highly cited author in 2019. In the same year, Zhang was elected the chair of big data technical committee of IEEE Computer Society. Yin Zhang has published over 90 articles in the recent five years. Those articles have been cited nearly 2000 times in SCI database and over 3500 times in Google Scholar, 14 among which were voted as ESI highly cited papers. And Batesmainly focuses on how to use information technology to improve care and care quality, cost effectiveness and results measures in healthcare practice.

V. Research Hotspot and Path Evolution

5.1 Research Hotspot Analysis

Next, for analyzing the research hotspots of literatures, setting node to "Keyword" and knowledge map was

generated as Fig 9. Keyword is extremely significant for understanding information of literatures. By analyzing frequent keywords, the hotspots in the specific area can be figured out. It is noticed that big data, healthcare, Internet of Thing, system and care are the keywords occurring very frequently.



Fig 9: Analysis of keywords co-appearance in literatures

Then clustering those high frequent keywords, Fig 10 appeared. In the database of WoS, clusters of IoT, data protection, stroke, data linkage, epidemiology and deep learning are obtained. Those categories represent different research hotspots of big data in healthcare industry during different periods in WoS database. See Table 8 for details of clusters.



Fig10: Clusters map for keywords

Cluster ID	Size	Silhouette	Year	Label (LLR)	Alternative labels
0	92	0.67	2017	internet of things	cloud computing; security; privacy; blockchain; fog computing; smart cities; edge computing
1	78	0.664	2016	data protection	electronic medical record; analytics; precision medicine; health services research; bioethics; privacy concern
2	71	0.615	2016	stroke	primary care; mortality; chronic disease; cloud computing; personalized medicine; data science; diabetes; osteoporosis; ageing; cancer incidence
3	67	0.717	2015	data linkage	quality; health; patienthood; participation; quality of life; cost analysis
4	47	0.74	2013	epidemiology	healthcare; integrative medicine; validation; socioeconomic factors
5	35	0.688	2018	deep learning	neural networks; feature extraction; computational modeling; support vector machines; CNN; principal component analysis; predictive models; big data analytics

Table 8 Primarilyclusters of keywords

Note: If Silhoutte > 0.5, clustering effect is good.

Labels and alternative labels of clusters 0 and 5 mainly concentrate on technology like IoT, deep learning and cloud computing. Along with the further development of computer science, algorithm about big data has been flourishing, such as cloud computing, fog computing and neural networks supported vector machine which have already been applied to healthcare service area. Obermeyer[21]used algorithm of machine learning based on data and information in healthcare area to investigate how big data technology acts in patients care. Lecun[22] expressed that deep learning can not only be applied to Automatic Speech Recognition (ASR) and object detection, but also be used in drug discovery and genetic research.

Labels and alternative labels of cluster 2 and 4 basically pay attention to common diseases and subjects in medicine field, such as stroke, chronic disease and epidemiology, etc. Nishimura[23] used big data technology to bring new ideas for stroke unit care. The author introduced a new registry for acute stroke unit care. Mooney[24] thought that big data has been used for many subjects including epidemiology. With medical data's growth, combining the data can bring great value to solving issues of epidemiology.

Cluster 3 mainly concentrates on how to connect data effectively after collecting data from patients in order to raise efficiency of healthcare services and quality of life, and keep costs down. Barreto[25] believed that the exploration of big data is happening in business intelligence and healthcare service area and this technology can benefit the whole industry and bring better decisions. Especially in health field, big data is facing big challenge since the quality of data can be fluctuating, and people need to gather and dispose huge amount of data from different databases. In order to generate accurate data collection, the author proposed a data managing approach, which can efficiently connect data, based on spark.

5.2 Path Evolution Analysis

In order to study the evolution path of big data in healthcare industry, this paper unfolded the changes of keywords in the last decade. Setting up the same parameters as before, keywords in WoS database has been collected by CiteSpace. Select the view display type as "Timezone" and Fig 11 was generated as below.



Fig 11: Timezone of keywords clusters

The first stage (2011-2013) basically introduces the background, the relevant technology and the challenge of big data, and studies the applications of big data in healthcare service[14]. In this stage, through connecting data and sharing information effectively of systems, a unified storage pool has been established, laying a solid foundation for the establishment of Electronic Health Record (EHR). Then by integrating the basic IT frame, data management and utilization can be achieved, as well as the purpose of integrated accessing of information. However, because of privacy, ethic and legal issues, the data in EHR can be hardly used completely. This is the most urgent issue to be solved[26].

The second stage (2014-2016) is the explosion of big data applying to healthcare industry. Big data technology has been emerging and cloud computing, machine learning and other technology have begun to be applied to healthcare industry. Cloud computing, a kind of technology which can run large amount of computation, has ascendancy in data disposing and analysis [18], and enjoys huge potential in promoting healthcare service[27, 28]. Machine learning has also developed applications gradually in this stage. Chen et. al.[29]proposed a new algorithm of CNN disease risk prediction using big data of hospitals. This algorithm has been proved to be very effective. Besides, as a branch of mathematics, big data analytics will play a key role in big data analysis area in future. It can help people achieve the goals of modern healthcare service system: high quality, quick response, low cost and efficient nursing service.

The third stage (2017-2020) is the further application of big data technology. As the advent of widespread use of AI, IoT, deep learning and other technologies, healthcare industry benefits from it a lot. Nowadays, AI technology is mainly used in key fields of disease areas including cancer and cardiology areas [30]. Furthermore, with the application of AI, health management service makes breakthroughs, especially in mobile medical equipment focusing on sports, cardiac rhythm and sleeping tests [4]. Health indicators such as blood pressure, electrocardio (ECG) and fat rate can be detected rapidly through AI equipment. The health data collected is uploaded to cloud database to set up individual health file, and personalized health management project is generated through data analysis. Meanwhile, based on the personal habits of users, AI technology can analyze users' data, evaluate users' overall status, and then provide personalized health management projects and assist health management staffs to plan daily healthcare arrangements for users.

VI. Conclusion

The analysis of literatures of big data in healthcare industry has provided a unique viewpoint for applications of big data. Based on all the analysis above, the following conclusions can be gained: First, focus on the quantity of literatures, the numbers of articles published have grown rapidly. This phenomenon proves that the applications of big data in healthcare industry have become more and more. With the development of computer technology, the application of big data started from healthcare data connection, and expanded to healthcare data analysis, medical equipment and other areas; Second, in WoS database, the relevant papers were mainly published by worldwide famous universities or scientific research departments. Those organizations play a huge part in the frontier of big data analysis in healthcare field; Third, according to the analysis to the recent 10 years' literatures generated in this area, several basic research directions are divided: the newest algorithm about big data, machine learning and deep learning including cloud computing, fog computing and neural networks supported vector machine which are now applied to healthcare service area with the help of development of computer technology; attention of the challenges that big data and other technologies may face when applying to healthcare service, like how to protect data privacy efficiently; and how to increase the healthcare service efficiency and living quality while decrease the medical cost after collecting the relevant data of patients.

Acknowledgements

This work is supported in part by the Research Foundation for Talents of Hefei University(18-19RC65), and also in part by Humanities and Social Science Project of Anhui Provincial Education Department (Grant SK2019A0693).

References

- [1] T.B. Murdoch, A.S. Detsky, "The inevitable application of big data to health care," The Journal of the American Medical Association, vol. 309, pp. 1351-1352, 2013.
- [2] N.V. Chawla, D.A. Davis, "Bringing Big Data to Personalized Healthcare: A Patient-Centered Framework," Journal of General Internal Medicine, vol. 28, pp. S660-S665, 2013.
- [3] A. Allen, "The Biopharmaceutical Industry and Big Data," Journal of the American Academy of Child & Adolescent Psychiatry, vol. 55, pp. S308-S309, 2016.
- [4] S.S. Gadde, V.D.R. Kalli, "Applications of Artificial Intelligence in Medical Devices and Healthcare," International Journal of Computer Science Trends and Technology, vol. 8, pp. 182-188, 2020.
- [5] X. Chen, Y. Liu, "Visualization analysis of high-speed railway research based on CiteSpace," Transport Policy, vol. 85, pp. 1-17, 2020.
- [6] C. Birkle, D.A. Pendlebury, J. Schnell, J. Adams, "Web of Science as a data source for research on scientific and scholarly activity," Quantitative Science Studies, vol. 1, pp. 1-14, 2020.
- [7] D.J. De Solla Price, "Little Science Big Science," Columbia: Columbia University Press, 1963.
- [8] K. Chung, R.C. Park, "Chatbot-based heathcare service with a knowledge base for cloud computing," Cluster Computing, vol. 22, pp. 1925-1937, 2019.
- [9] H. Chen, R.H.L. Chiang, V.C. Storey, "Business Intelligence and Analytics: From Big Data to Big Impact," Mis Quarterly, vol. 36, pp. 1165-1188, 2012.
- [10] H. Small, "Co- citation in the scientific literature: A new measure of the relationship between two documents," Journal of the American Society for Information Science, vol. 24, pp. 28-31, 1973.
- [11] W. Raghupathi, V. Raghupathi, "Big data analytics in healthcare: promise and potential," Health Information Science & Systems, vol. 2, pp. 2047-2501, 2014.
- [12] D.W. Bates, S. Saria, L. Ohno-Machado, A. Shah, G. Escobar, "Big data in health care: using analytics to identify and manage high-risk and high-cost patients," Health Aff, vol. 33, pp. 1123-1131, 2014.
- [13] S.M.R. Islam, D. Kwak, M.H. Kabir, M. Hossain, K.S. Kwak, "The Internet of Things for Health Care: A Comprehensive Survey," IEEE Access, vol. 3, pp. 678-708, 2015.
- [14] M. Chen, S. Mao, Y. Liu, "Big Data: A Survey," Mobile Networks & Applications, vol. 19, pp. 171-209, 2014.

- [15] J. Jiang, L. Qu, "Evolution and Emerging Trends of Sustainability in Manufacturing(1999–2019) Based on Literature Visualization Analysis," IEEE Access, vol. 8, pp. 121074-121088, 2020.
- [16] J. Andreu-Perez, C.C.Y. Poon, R.D. Merrifield, S.T.C. Wong, G.Z. Yang, "Big Data for Health," IEEE Journal of Biomedical and Health Informatics, vol. 19, pp. 1193-1208, 2015.
- [17] J. Gubbi, R. Buyya, S. Marusic, M. Palaniswami, "Internet of Things (IoT): A Vision, Architectural Elements, and Future Directions," Future Generation Computer Systems, vol. 29, pp. 1645-1660, 2012.
- [18] I.A.T. Hashem, I. Yaqoob, N.B. Anuar, S. Mokhtar, A. Gani, S.U. Khan, "The rise of "big data" on cloud computing: Review and open research issues," Information Systems, vol. 47, pp. 98-115, 2015.
- [19] L. Catarinucci, D. De Donno, L. Mainetti, L. Palano, L. Patrono, M.L. Stefanizzi, L. Tarricone, "An IoT-Aware Architecture for Smart Healthcare Systems," IEEE Internet of Things Journal, vol. 2, pp. 515-526, 2015.
- [20] A. M. Rahmani, T.N. Gia, B. Negash, A. Anzanpour, I. Azimi, M. Jiang, P. Liljevberg, "Exploiting smart e-Health gateways at the edge of healthcare Internet-of-Things:A fog computing approach," Future Generations Computer Systems, vol. 78, pp. 641-658, 2018.
- [21] Z. Obermeyer, E.J. Emanuel, "Predicting the Future Big Data, Machine Learning, and Clinical Medicine," N Engl J Med, vol. 375, pp. 1216-1219, 2016.
- [22] Y. LeCun, Y. Bengio, G. Hinton, "Deep learning," Nature, vol. 521, pp. 436-444, 2015.
- [23] A. Nishimura, K. Nishimura, A. Kada, K. Iihara, "Status and Future Perspectives of Utilizing Big Data in Neurosurgical and Stroke Research," Neurol Med Chir, vol. 56, pp. 655-663, 2016.
- [24] S.J. Mooney, D.J. Westreich, A.M. El-Sayed, "Epidemiology in the Era of Big Data," Epidemiology, vol. 26, pp. 390-394, 2015.
- [25] M. Barreto, R. Pita, C. Pinto, M. Silva, D. Rasella, "A Spark-based workflow for probabilistic record linkage of healthcare data," in Workshop on Algorithms & Systems for Mapreduce& Beyond, Brussels, pp. 1-10, 2015.
- [26] P.B. Jensen, L.J. Jensen, S.R. Brunak, "Mining electronic health records: towards better research applications and clinical care," Nature Reviews Genetics, vol. 13, pp. 395-405, 2012.
- [27] N. Sultan, "Making use of cloud computing for healthcare provision: Opportunities and challenges," International Journal of Information Management, vol. 34, pp. 177–184, 2014.
- [28] S. Vladimir, C.P. Ricardo, N. Michael, "Cloud Computing Based Systems for Healthcare," The scientific world journal, vol. 2014, pp. 692619, 2014.
- [29] M. Chen, Y. Hao, K. Hwang, L. Wang, L. Wang, "Disease Prediction by Machine Learning Over Big Data From Healthcare Communities," IEEE Access, pp. 8869-8879, 2017.
- [30] F. Jiang, Y. Jiang, H. Zhi, Y. Dong, H. Li, S. Ma, Y. Wang, Q. Dong, H. Shen, Y. Wang, "Artificial intelligence in healthcare: past, present and future," Stroke & Vascular Neurology, vol. 2, pp. 230-243, 2017.