2020

Exploring Data Mining Research in West Africa: A Bibliometric Analysis

Oladosu O. Oladimeji

Olayanju O. Oladimeji

Follow this and additional works at: <https://aquila.usm.edu/slisconnecting>

Part of the [Archival Science Commons](#), [Collection Development and Management Commons](#), [Information Literacy Commons](#), [Scholarly Communication Commons](#), and the [Scholarly Publishing Commons](#)

Recommended Citation

Oladosu O. Oladimeji and Olayanju O. Oladimeji (2020) "Exploring Data Mining Research in West Africa: A Bibliometric Analysis," *SLIS Connecting*: Vol. 9 : Iss. 2 , Article 5.

DOI: 10.18785/slis.0902.05

Available at: <https://aquila.usm.edu/slisconnecting/vol9/iss2/5>

This Article is brought to you for free and open access by The Aquila Digital Community. It has been accepted for inclusion in SLIS Connecting by an authorized editor of The Aquila Digital Community. For more information, please contact Joshua.Cromwell@usm.edu.

Exploring Data Mining Research in West Africa: A Bibliometric Analysis

By Oladosu O. Oladimeji^a and Olayanju O. Oladimeji^b

^aUniversity of Ibadan, Ibadan, Nigeria

^bBowen University, Iwo, Nigeria

Introduction

This present century is being referred to as the information age, and it is believed that information is key in decision making nowadays. With the advent of technologies such as computers, tremendous amounts of information are being collected and stored.¹ Unfortunately, these massive collections of information (data) stored in various storage mechanisms are becoming overwhelming, even after the creation of structured databases and database management systems (DBMS). Now, far more information can be handled, ranging from transactions, text, pictures, and intelligence. Therefore, information retrieval is not enough for making right decisions. With this, there is a need for help to make the right decision, including summarizing data, extraction of relevant information stored, and discovering patterns.

Data mining is the new powerful technology that meets these needs. It is the automated analysis of large or complex data sets to discover significant, interesting patterns or trends that would otherwise go unrecognized.^{2,3} Data mining nowadays is a crucial subject in either academic research or practical application.⁴ A number of developed countries are already reaping the benefits of data mining. Developing countries also have the potential to be influenced by data mining, especially the areas of commerce, health-care, and government, to enhance daily activities in which West African countries are not left out. West Africa, according to the United Nations, includes 16 countries: Benin, Burkina Faso, Cape Verde, The Gambia, Ghana, Guinea, Guinea-Bissau, Ivory Coast, Liberia, Mali, Mauritania, Niger, Nigeria, Senegal, Sierra Leone, Togo, as well as Saint Helena, Ascension and Tristan da Cunha (United Kingdom Overseas Territory).⁵ As information technology is developing in West Africa and a large amount of data generated, there is a need to apply data mining to make better decisions.⁶ Data mining has been applied virtually in most West African sectors, especially in business and commerce, where there is competition.

Thus, this research paper's main objective is to quantify and analyze the research output of data mining in West Africa. The second section of this paper discusses the methodology used in this research, while the third section showcases the results, followed by the discussion of the results, and finally, the conclusion is drawn.

Materials and Methods

The data analyzed to pinpoint the bibliometric analysis in this study were downloaded from *Scopus* database, which is the largest abstract and citation database of peer-reviewed literature, comprising journals, workshops, and conference proceedings.^{7,8,9} The choice was made based on the number of studies done in retrospective times that analyzed the coverage difference between other databases, including *Google Scholar*, *Web of Science*, *CSA Illumina*, and *Scopus* database.^{10,11,12} In addition, Norris analyzed differences in these databases and concluded, "Scopus offers the best coverage from amongst these databases."¹³

In this research, we have investigated various aspects of research growth in the field of data mining in West Africa. The factors that are being considered include (1) growth in publications as measured by the Relative Growth Rate (RGR) and Doubling Time (Dt), (2) productivity and quantification of research output for different institutions of the region (West Africa).

Data Mining as a field of study has a lot of terms related to it, which includes: Cluster Analysis, ID3, Neural Network, Pattern Extraction, Decision Tree, Neural Network, etc. for this research, the generic term "Data Mining" or "Knowledge Discovery" was adopted. Therefore, the research data were retrieved based on the carefully-crafted string "Data* Mining*" OR "Knowledge* Discovery,*" with the years 2020 and 2019 excluded. The search country was limited to West African Countries that were indexed in *Scopus* database at the time of this research: Benin, Burkina Faso, Cape Verde, Gambia, Ghana, Mali, Nigeria, Senegal, Sierra Leone. The initial output data that were retrieved were cleaned by removing irrelevant and duplicate records. Also, journals with zero impact factors were removed. Finally, 146 publications were considered for further analysis.

Results

Based on data retrieved, it is important to note the first document on data mining to be indexed in *Scopus* from West Africa was from Nigeria in the year 2000.¹⁴

Preferred Journals and Areas of Research

This research, as stated above, covers 17 years, which ranges from 2000 – 2018. The total number of West

African data mining publications retrieved from *Scopus* was 146, of which 81 were research articles, 47 were in conference proceedings, 12 were book series, and 6 were books. A list of the top 13 journals preferred by West African researchers is presented in Table 1.

The number of research publications per subject area of data mining is listed in Table 2. Data mining research in West Africa is dominated by the sectors related to computer science (29.10 %), followed by engineering (15.67 %), social sciences (8.21 %), and medicine (6.34 %). The classification of subjects is based on the one provided by *Scopus*.

Table 1. Top 14 Sources Publishing Data Mining Research in West Africa

S/No	Journal	Number of Publications	Impact Factor
1	<i>Data in Brief</i>	8	0.85
2	<i>Advances in intelligent Systems and Computing</i>	4	0.37
3	<i>CEUR Workshop Proceedings</i>	3	0.34
4	<i>European Journal of Scientific Research</i>	3	0.21
5	<i>information Management and Computer Security</i>	3	5.00
6	<i>Lecture Notes in Computer Science including Subseries Lecture Notes in Artificial intelligence and Lecture Notes in Bioinformatics</i>	3	0.60
7	<i>Lecture Notes in Engineering and Computer Science</i>	3	0.45
8	<i>Arpn Journal of Engineering and Applied Sciences</i>	2	0.22
9	<i>Environmental Monitoring and Assessment</i>	2	1.12
10	<i>Expert Systems With Applications</i>	2	3.25
11	<i>Journal of Computing and information Technology</i>	2	0.73
12	<i>Journal of Engineering Science and Technology</i>	2	0.36
13	<i>Lecture Notes of the Institute for Computer Sciences Social Informatics and Telecommunications Engineering</i>	2	0.31
14	<i>Proceedings of the IEEE international Conference on Computing Networking and Informatics (ICCNI 2017)</i>	2	N/A

Table 2. Subject Area Classification of Data Mining Publications in West Africa

S.NO	Subject Area	No. of Publications	% of Total
1	Computer Science	78	29.10
2	Engineering	42	15.67
3	Social Sciences	22	8.21
4	Medicine	17	6.34
5	Mathematics	16	5.97
6	Biochemistry, Genetics and Molecular Biology	12	4.48
7	Earth and Planetary Sciences	12	4.48
8	Multidisciplinary	11	4.10
9	Agricultural and Biological Sciences	10	3.73
10	Environmental Science	9	3.36
13	Decision Sciences	8	2.99
14	Business, Management and Accounting	6	2.24
15	Chemical Engineering	4	1.49
16	Energy	4	1.49
17	Materials Science	4	1.49
18	Immunology and Microbiology	3	1.12
19	Health Professions	2	0.75
20	Physics and Astronomy	2	0.75
21	Psychology	2	0.75

22	Chemistry	1	0.37
23	Economics, Econometrics and Finance	1	0.37
24	Nursing	1	0.37
25	Pharmacology, Toxicology and Pharmaceutics	1	0.37

Growth of Publications

The growth in the number of publications is measured by means of two related parameters viz. RGR and D_t ¹⁵. RGR in classical growth analysis is defined as:

$$RGR = (\ln S_2 - \ln S_1) / (t_2 - t_1)$$

Here S_2 and S_1 are the cumulative publications in two years t_2 and t_1 . In the present analysis $t_2 - t_1$ is taken as one year. RGR can then be expressed as $RGR = \ln (S_2 / S_1)$. The D_t is the time required for publications to double in number for a given RGR. D_t is expressed as:

$$D_t = (t_2 - t_1) \ln 2 / (\ln S_2 - \ln S_1) \text{ or } D_t = \ln 2 / RGR$$

A constant value for RGR in each year subsequently is an indication that the growth rate is exponential. The D_t is a characteristic time for this exponential growth. In Table 3, the annual number of publications in the discipline of data mining is represented and their cumulative number, RGR

and D_t , for the period 2000–2018. Starting from only one publication in 2000 to 38 in 2018, there has been a significant increase in the number of publications output. The average annual growth rate of publications (Figure 1) during the last 18 years (2000–2018) is an impressive 28 percent.

Quality of Publications

The number of citations that a paper receives is a good measure of its quality and importance. In columns six and seven of Table 3, we give the number of citations that a paper has received for that year and the average citations per paper (ACCP), respectively. The citations already reported in a certain year are not included in the subsequent year. Figure 2 shows the growth trend in the number of citations.

Table 3. West Africa Data Mining Research Output, Relative Growth Rate, Doubling Time, Citation per Paper

Years	Number of Publication	Cumulative	RGR	D_t	Citation	Citation per paper
2000	1				0	0.00
2001	1	2	0.69	1	15	15.00
2002	0	2	0.00	N/A	0	N/A
2003	0	2	0.00	N/A	0	N/A
2004	2	4	0.69	1	318	159.00
2005	2	6	0.41	1.71	14	7.00
2006	3	9	0.41	1.71	142	47.33
2007	0	9	0.00	N/A	0	N/A
2008	2	11	0.20	3.45	13	6.50
2009	5	16	0.37	1.85	27	5.40
2010	5	21	0.27	2.55	23	4.60
2011	10	31	0.39	1.78	106	10.60
2012	6	37	0.18	3.92	35	5.83
2013	7	44	0.17	4.00	62	8.86
2014	9	53	0.19	3.72	91	10.11
2015	9	62	0.16	4.42	34	3.78
2016	18	80	0.25	2.72	153	8.50
2017	28	108	0.30	2.31	112	4.00
2018	38	146	0.30	2.30	524	13.79

Figure 1. Data Mining Publication Pattern, 2000-2018

Figure 2. Citations of Data Mining Papers, 2000–2018

Table 4. Top 10 West African institutions' Data Mining Publications and Citations

	Institution	TP	TC	ACPP
1	Covenant University	18	101	5.61
2	Federal University of Agriculture, Abeokuta	11	51	4.64
3	Federal University of Technology, Akure	10	45	4.50
4	University of Ibadan	10	57	5.70
5	University of Lagos	8	20	2.50
6	Redeemer's University	6	19	3.17
7	University of Ghana	6	482	80.33
8	international institute of Tropical Agriculture IITA, Ibadan	5	357	119.0
9	Ahmadu Bello University	4	4	1.00
10	University of Abuja	4	35	8.75

Figure 3. West African Countries and Number of Data Mining Publications

Table 4 presents the total publications (TP), total citations (TC), and ACPP for the top 10 West African institutions with publications in data mining for the period 2000–2018. Covenant University tops the table. Covenant University is the leading private university in Nigeria¹⁶. Covenant University is closely followed by the Federal University of Agriculture, Abeokuta, while the Federal University of Technology, Akure, comes after.

National Contribution to Data Mining

Figure 3 shows the West African countries indexed Scopus in the research field of Data Mining and their number of publications.

Discussion and Conclusion

The trend of data mining research in West Africa is dominated by Nigerian institutions being the first country in the region to key into data mining research, which was followed by Ghana 9 years later,¹⁷ then followed by other countries. Other West African countries yet to key into this research area are encouraged to do so through knowledge exchange from countries already indexed in this research, especially from Nigeria, in order to boost data mining research in West Africa. Although Covenant University has the highest number of publications, in terms of quality of publication based on citations, the University of Ghana and IITA has the highest number of citations.

In tracing the research in the field of data mining in West Africa from 2000, the number of research publications has shown a rising trend, though some years have zero publications. From just one publication in 2000, the number increased to 146 for 2018, though more work needs to be done. The trend of data mining research in West Africa is dominated by Nigerian institutions.

References

1. Ramangeri, B.M (2010). Data mining techniques and application, *India Journal of Computer Science and Engineering* 1(4), 301-305.
2. Jiawei, H. & Kamber, M. (2011). Data mining: Concepts and techniques. Burlington, MA: Morgan Kaufman.
3. Feng Chen, et al. (2015). Data mining for the internet of things: Literature review and challenges. *international Journal of Distributed Sensor Networks*, 11(8). Retrieved from <https://journals.sagepub.com/doi/10.1155/2015/431047>
4. Tseng S.F., Won Y.L., & Yang J.M. (2016). A bibliometric analysis on data mining and big data. *international Journal of Electronic Business*, 13(1), 38-69.

5. Paul R.M. & Catherine, A.P. (2001). Monetary union in West Africa (ECOWAS): Is it desirable and how it could be achieved? *international Monetary Fund*. Retrieved from <https://www.imf.org/external/pubs/nft/op/204/>
6. Odachi G.N. & Nwosu I.N. (2015). Data mining: A strategy for economic and educational development in West Africa. *World Applied Sciences Journal*, 33(12), 1846-1851.
7. Azubuike A.E. & Igbekele E. (2019). Cloud computing research in Nigeria: A bibliometric and content analysis. *Asian Journal of Scientific Research*, 12, 41-53.
8. Bajwa, R.S., Yaldram, K. (2013). Bibliometric analysis of biotechnology research in Pakistan. *Scientometrics*, 95, 529–540.
9. Elsevier, (2016). SCOPUS.
10. Archambault, E., et al. (2009). Comparing bibliometric statistics obtained from the Web of Science and Scopus. *Journal of the American Society for information Science and Technology*, 60(7), 1320–1326.
11. Bajwa, R. S. & Yaldram, K. (2012). Research output in nanoscience and nanotechnology: Pakistan scenario. *Journal of Nanoparticle Research*, 14(2), 1–6.
12. Lopez-Illescas, C., de Moya-Anegon, F., & Moed, H.F. (2009). Comparing bibliometric country-by-country rankings derived from the Web of Science and Scopus: The effect of poorly cited journals in oncology. *Journal of information Science*, 35(2), 244–256.
13. Norris, M. & Oppenheim, C. (2007). Comparing alternatives to the Web of Science for coverage of the social sciences' literature. *Journal of informetrics*, 1(2), 161–169.
14. Olaleye J.B., Sangodina J.O. (2000). Geostatistical analysis by vector space methods. *international Archives of the Photogrammetry, Remote Sensing and Spatial information Sciences - ISPRS Archives*, 33, 769-778. Retrieved from https://www.isprs.org/proceedings/xxxiii/congress/part4/769_XXXIII-part4.pdf
15. Mahapatra, M. (1985). on the validity of the theory of exponential growth of scientific literature. in *15th IASLIC Conference Proceedings* (pp. 61–70). Bangalore: IASLIC.
16. Webometrics (2018). Webometrics.info/Africa/Nigeria.
17. Kankam S. & Boon E.K. (2009). Energy delivery and utilization for rural development: Lessons from Northern Ghana. *Energy for Sustainable Development* 13(3), 212-218.