Artificial Intelligence in MOOCs:

A Bibliometric Perspective

Ishteyaaq Ahmad[1]

Sonal Sharma[2]

Gulsun Kurubacak[3]

Ajay Kumar[4]

Rajiv Kumar[5]

Sikha Ahmad[6]

[1, 2, 5] *Uttaranchal Institute of Management, Uttaranchal University, Dehradun, India, ishteyaaq@gmail.com[1], sonal_horizon@rediffmail.com[2], rajiv.gill1@gmail.com[5]*

[3]*Anadolu University College of Open Education, Turkey,gkurubac@anadolu.edu.tr*

[4]*Bharati Vidyapeeth (Deemed to be University) Institute of Management and Research,New Delhi, ajay.kumar@bharatividyapeeth.edu*

[6]*Doon University, Dehradun, India, sikhaahmad@gmail.com*

Abstract.

Artificial intelligence has come a long way in recent years, and it currently represents an emerging technology that will change the way people live. Despite the fact that many instructors are unaware of its scope and, more significantly, what it implies, the use of AI in education is growing rapidly, with the aim of improving the quality of teaching and learning. There are many different approaches to this, but they all involve the use of data analytics and machine learning to make improvements in education. The purpose of this research was to look at the scientific output on artificial intelligence in MOOCs that were indexed in Scopus databases between 2012 and 2021. For this a customised approach for bibliometric investigations was applied on the SCOPUS databases. A total of 476 articles were included in the sample. The findings revealed that the issue has a lot of interest, but that the literature on it is still in its early stages.

Keywords. Intelligent systems; artificial intelligence; machine learning; MOOCs; bibliometric study

1. INTRODUCTION

From time immemorial human beings have been learning and adapting to the new environment. This movement has continued for millions of years in all dimensions of human life but picked up incredible momentum in the modern era. The advent of information technology amplified this learning and adapting process by leaps and bounds [1], [2], [3], [4]. The world in which we are living is changing at such a fast and unpredictable rate and it is difficult to visualize how different things were just a few years back. Artificial Intelligence (AI) has gone a great way in recent years, and it is now regarded as a disruptive technology that will affect people's lives in the not-too-distant future. Various studies have already examined the primary applications of artificial intelligence in education, as well as the fundamental design difficulties and methodologies used in the system's adaptation to the user [5]. Traditionally, intelligent educational systems (IESs) were considered in the context of their functional components. Domain knowledge, a representation of the learner's current state, and instructional knowledge are the three mandatory subsystems [6]. There are many different ways in which AI can be used in education. One way is through the use of data analytics to improve teaching and learning practices. Here systems can involve machine learning, where algorithms can be trained on large sets of data to learn from them and then make predictions or recommendations.

Massive open online courses (MOOCs), which have been popular for a few years, are one of the fields of education where a large number of learners are enrolled. These courses can be taken at any time, on any device and are a cheaper and more flexible way to learn. The best thing about MOOCs is that people can choose from a wide range of courses and study at their own pace, allowing them to complete the course faster or take it slower. As artificial intelligence is also growing rapidly in education where its primary goal is to improve the quality of teaching and learning by making it more accessible, personalized, adaptive, and interactive. Because of enormous MOOCs enrolments, AI is likely to play a significant role in analysing the large data set of MOOCs. Data analytics may be used by AI to improve teaching and learning techniques. The machine learning algorithms can be trained on large datasets of MOOCs to learn from them and then make predictions or recommendations about how to teach better or learn something new.

A search of the scientific literature on artificial intelligence in MOOCs can be useful in this scenario to see how many research papers are available. As a result, this bibliometric analysis looks at the scholarly output on artificial intelligence in MOOCs that was indexed in Scopus between 2012 and 2021, and would address the following research questions:

RQ1: What is the status of paper publication in the last ten years (2012 – 2021)?

RQ2: What are the major organisations and countries that produce the most artificial intelligence research within MOOCs?

RQ3: What are the most-cited publications and popular keywords for research in artificial intelligence output in MOOCs?

2.   METHODOLOGY

In this research paper, we employed a bibliometric study approach that adhered to the PRISMA pronouncement's requirements. The metadata of scientific work on the Scopus database in the last ten years (2012–2021) was analysed for this. For this purpose, the search was broadened to include the popular subsets of artificial intelligence viz., "machine learning" and "intelligent systems". The Scopus database was searched using the keywords ("Artificial Intelligence" OR "Machine Learning" OR "Intelligent System") AND "MOOC". The results were limited to publication years (2012–2021). The data analysis was carried out on 476 research papers that were indexed in Scopus. Software, such as Microsoft Excel, VOSviewer version 1.6.18 were used for investigations.

3.   RESULTS AND DISCUSSIONS

3.1.   *Status of paper publication in the last ten years*

Taking into account the "year of production" variable, the findings revealed that the largest production peaks in the Scopus database in the year 2020 with 91 publications, while the years 2021 and 2019 each had 89 and 85 publications, respectively. Figure 3.1 depicts the year-by-year release of materials in MOOCs on the theme of artificial intelligence from 2012 to 2021. The findings indicate that there is a growing interest in the subject, since the number of publications has increased in the last ten years.
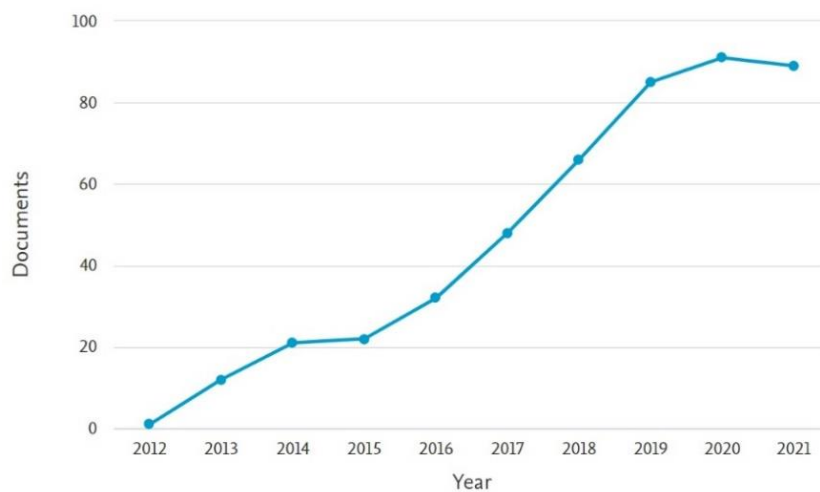

Figure 3.1. Year of production of published documents

In terms of "document type" conference papers accounted for 71.7 percent of overall output in the Scopus database, followed by journal articles, which accounted for 20.3 percent of total production. Table 1 depicts the document type of research publication in MOOCs on the theme of artificial intelligence from 2012 to 2021.

Comparing the documents by source for the counts of at least 10, the series "Lecture Notes in Computer Science" has the highest number of publications in the last ten years with 72

documents, followed by "ACM International Conference Proceeding Series" with 31 documents. The complete result is shown in Figure 3.2.

Table1. Documents by type

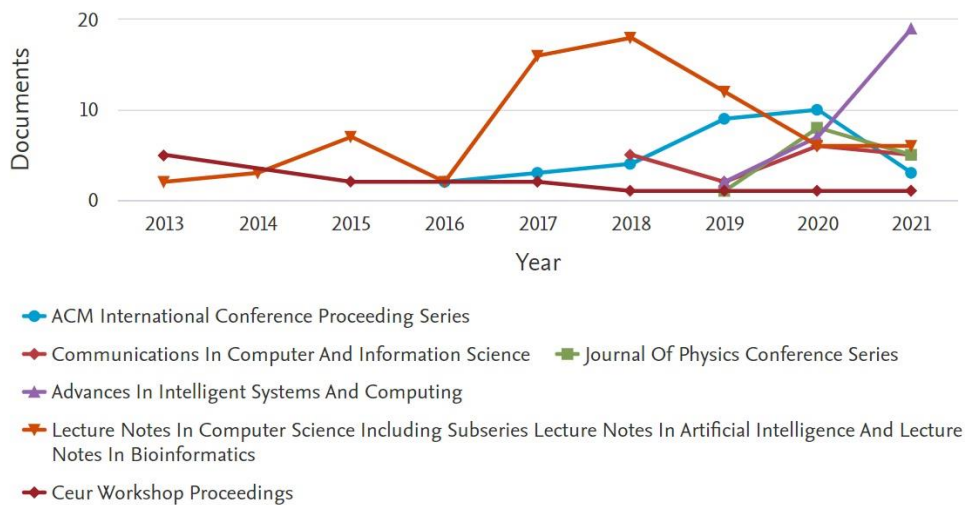| Document type	Documents	Percentage
Conference Paper	335	71.7
Journal Article	95	20.3
Conference Review	25	5.4
Review	5	1.1
Book Chapter	4	0.9
Book	1	0.2
Erratum	1	0.2
Short Survey	1	0.2

Figure 3.2. Published documents by source

Analysing the documents by subject area, the research found that 44.3% of the documents were related to computer science, 14.2% were related to mathematics, and 13.7% were related to social sciences. The complete subject analysis is shown in Table 2.

3.2. *Major organisations and countries that produce the most research*

According to the Scopus database, the top institution with research publications on artificial intelligence in MOOCs, in the previous ten years, from 2012 to 2021, was Carnegie Mellon University (USA), followed by Universidad Autónoma de Madrid

(Spain), and Norges Teknisk-Naturvitenskapelige Universitet (Norway). Table 3 summarises the top 10 results for this world-wide search.

Table 2. Documents by Subject Area

Subject area	Documents	Percentage
Computer Science	391	44.3
Mathematics	125	14.2
Social Sciences	121	13.7
Engineering	115	13.0
Decision Sciences	37	4.2
Physics and Astronomy	21	2.4
Business, Management and Accounting	15	1.7
Materials Science	9	1.0
Psychology	9	1.0
Medicine	8	0.9
Others	32	3.6

Table 3. Documents by Affiliation

Organizations	n	%
Carnegie Mellon University (USA)	11	2.31
Universidad Autónoma de Madrid (Spain)	8	1.68
Norges Teknisk-Naturvitenskapelige Universitet (Norway)	7	1.52
Universidad de Valladolid (Spain)	7	1.52
Universidad Carlos III de Madrid (Spain)	6	1.26
Huazhong Normal University (China)	6	1.26
Ecole Mohammadia d'Ingenieurs, Mohammed V University (Morocco)	6	1.26
Stanford University (USA)	6	1.26
Durham University (England)	6	1.26
Massachusetts Institute of Technology (USA)	6	1.26

In terms of "countries", according to the Scopus database, China has most publications on artificial intelligence research on MOOCs, with 110, followed by the United States with 83, India with 37, and Spain with 34 research articles. Figure 3.3 depicts the top ten countries of the entire search outcome.
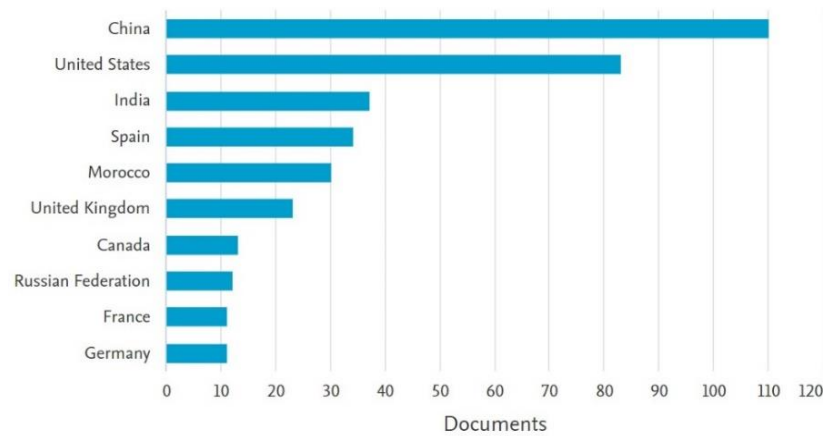
Figure 3.3. Published documents by country

3.3. *Most-cited publications and popular keywords for research in artificial intelligence output in MOOCs*

Finally, data on the variable "the most-cited publications" was gathered. This was picked as a criterion for highlighting the most significant research. Table 4 summarises the top five most referenced publications in Scopus database.

Table 4. The most cited articles in Scopus on "artificial intelligence in MOOCs"

Title	Authors	Journal	Year	Citations
Rebooting MOOCresearch[7]	Reich, J.	Science	2015	187
MOOCs: So many learners, so much potential [8]	Kay, J., Reimann, P., Diebold, E., Kummerfeld, B.	IEEE Intelligent Systems,	2013	157
Temporal Models for Predicting Student Dropout in Massive Open Online Courses [9]	Fei, M., Yeung, D.-Y.	Proceedings - 15th IEEE International Conference on Data Mining Workshop, ICDMW	2015	135
Characteristics of massive open online courses (MOOCS): A research review, 2009-2012 [10]	Kennedy, J.	Journal of Interactive Online Learning,	2014	112
Combination of machine learning algorithms for recommendation of courses in E-Learning System based on historical data [11]	Aher, S.B., Lobo, L.M.R.J.	Knowledge-Based Systems,	2013	107

The Scopus database was analysed for keywords using VOSviewer version 1.6.18. The minimum threshold of 25 appearances was required for a word to be included in the research. Only 24 keywords out of 2792 terms satisfied this criterion. Figure 3.4 depicts a co-occurrence network map, generated by VOSviewer using 24 keywords relying on a relevance score with three final clusters depicted in different colours (red, green, and blue). With phrases like "forecasting", "machine learning", "learning systems", "learning algorithms", "decision trees", "learning analytics" and "data mining" the red cluster appears to focus on the predictive aspects of MOOC. The green cluster appears to be focused on many elements that may influence MOOC acceptance, with phrases like "online education", "distance education", "online learning", "big data", "intelligent systems", "students", "education computing" and "curricula" included. Finally, with the keywords like "education", "computer aided instruction", "engineering education", "education", "artificial intelligence" and "teaching" the blue cluster appears to be focused on the implementation of the MOOCs.
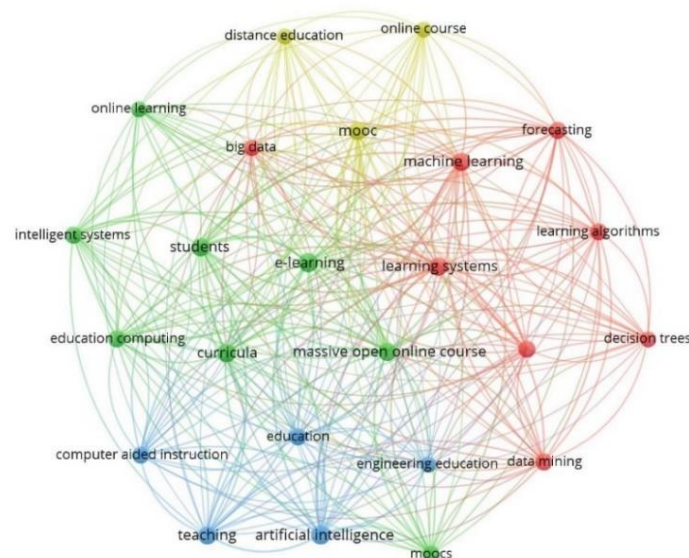

Figure 3.4. Keyword co-occurrence network map in the published documents

4.   DISCUSSION AND CONCLUSIONS

Scopus database reveals that the number of documents published on use of artificial intelligence in MOOCs has increased in recent years. The majority of the publications were proceedings papers, signifying a high level of interest in the subject. Although, there exists a significant gap between conference papers (71.7%) and journal articles (20.3%). The literature suggests that MOOCs integrated with artificial intelligence are now a possibility. The study looked at a wide range of published works on artificial intelligence in MOOCs, demonstrating the researchers' heterogeneity on the subject. Several authors have discussed the MOOC dropout prediction studies [9], [12], [13][14]. While significant

scholars laid emphasis on numerous other topics some of these were viz., clickstream analysis of learners' [15], [16], satisfaction of the learners [17], [18], Time-based metrics related to assessments and student interactions [19], use of K-means algorithm MOOC dataset [20], machine learning algorithms to filter and classify the MOOC learners [21], emotional tendencies of learners [22], learning behaviours in MOOC [23], intelligent analysis [24], Virtual and Augmented Reality [25] etc.

MOOCs (Massive Open Online Courses) are slowly conquering the education business. 2017 Higher Education Edition of the Horizon Report predicted that by 2022, Artificial intelligence will be applied in higher education [26]. Moreover, the technologies like IOT, blockchain and AR / VR can be incorporated in the future. Although pandemic has disrupted the last two years, AI is set to become a widespread feature in MOOCs within few years. While artificial intelligence and MOOCs are both real, we acknowledge that artificial intelligence research in MOOCs is still in its early phases.

5.   REFERENCES

[1]   S. Al-Qatawneh, N. Alsalhi, A. Al Rawashdeh, T. Ismail, and K. Aljarrah, "To E-textbook or not to E-textbook? A quantitative analysis of the extent of the use of E-textbooks at Ajman University from students' perspectives," *Educ. Inf. Technol.*, vol. 24, no. 5, pp. 2997–3019, 2019, doi: 10.1007/s10639-019-09912-4.

[2]   I. Ahmad, S. Jasola, and A. Anupriya, "Supplementing higher education with MOOCs: A case study," in *2017 International Conference on Emerging Trends in Computing and Communication Technologies, ICETCCT 2017*, 2018, vol. 2018-Janua, doi: 10.1109/ICETCCT.2017.8280346.

[3]   I. Ahmad, S. Jasola, S. Ahmad, and S. Sharma, "Technology Enabled Learning in Higher Education: A Case Study," *Xplore Xavier's Res. J.*, vol. 11, no. 1, pp. 68–73, 2020, [Online]. Available: https://9dec1e70-3d6c-406f-960b-920f18cdab25.filesusr.com/ugd/03b37e_0a4fe5f7f9764c23a239c95f05ae24fc.pdf.

[4]   I. Ahmad, S. Ahmad, A. K. Pant, and R. Kumar, "Intelligent Decision Support System for Tourism," *Int. J. Latest Res. Sci. Technol.*, vol. 8, no. 5, pp. 1–4, 2019, [Online]. Available: https://www.mnkjournals.com/journal/ijlrst/pdf/Volume_8_5_2019/10978.pdf.

[5]   C. Soledad, "Intelligent Systems in Education: A review of current research lines," *RELIEVE*, vol. 10, pp. 3–22, 2004.

[6]   G. Cumming and J. Self, "Intelligent educational systems: identifying and decoupling the conversational levels," *Instr. Sci. 1990 191*, vol. 19, no. 1, pp. 11–27, Jan. 1990, doi: 10.1007/BF00377983.

[7]   J. Reich, "Rebooting MOOC Research," *Science (80-.).*, vol. 347, no. 6217, pp. 34–35, Jan. 2015, doi: 10.1126/SCIENCE.1261627.

[8]   J. Kay, P. Reimann, E. Diebold, B. Kummerfeld, and B. (2013). Kay, J., Reimann, P., Diebold, E., & Kummerfeld, "MOOCs: So Many Learners, So Much Potential ... IEEE Intelligent Systems," *IEEE Intell. Syst.*, vol. 28, no. 3, pp. 70–77, 2013.

[9]   M. Fei and D. Y. Yeung, "Temporal Models for Predicting Student Dropout in Massive Open Online Courses," *Proc. - 15th IEEE Int. Conf. Data Min. Work. ICDMW 2015*, pp. 256–263, 2016, doi: 10.1109/ICDMW.2015.174.

[10]   J. Kennedy, "Characteristics of massive open online courses (MOOCS): A research review, 2009-2012," *J. Interact. Online Learn.*, vol. 13, no. 1, pp. 1–16, 2014, [Online]. Available: http://www.ncolr.org/jiol/issues/pdf/13.1.1.pdf.

[11]   S. B. Aher and L. M. R. J. Lobo, "Combination of machine learning algorithms for recommendation of courses in E-Learning System based on historical data," *Knowledge-Based Syst.*, vol. 51, pp. 1–14, 2013, doi: 10.1016/j.knosys.2013.04.015.

[12]   J. Liang, C. Li, and L. Zheng, "Machine learning application in MOOCs: Dropout prediction," *ICCSE 2016 - 11th Int. Conf. Comput. Sci. Educ.*, no. Iccse, pp. 52–57, 2016, doi: 10.1109/ICCSE.2016.7581554.

[13]   D. Sun, Y. Mao, J. Du, P. Xu, Q. Zheng, and H. Sun, "Deep learning for dropout prediction in MOOCs," *Proc. - 2019 8th Int. Conf. Educ. Innov. through Technol. EITT 2019*, pp. 87–90, 2019, doi: 10.1109/EITT.2019.00025.

[14]   F. Dalipi, A. S. Imran, and Z. Kastrati, "MOOC dropout prediction using machine learning techniques: Review and research challenges," in *IEEE Global Engineering Education Conference, EDUCON*, May 2018, vol. 2018-April, pp. 1007–1014, doi: 10.1109/EDUCON.2018.8363340.

[15]   A. Alamri, Z. Sun, A. I. Cristea, G. Senthilnathan, L. Shi, and C. Stewart, "Is MOOC learning different for dropouts? A visually-driven, multi-granularity explanatory ML approach," 2020, doi: 10.1007/978-3-030-49663-0_42.

[16]   M. M. Al-Rifaie, M. Yee-King, and M. D'Inverno, "Boolean prediction of final grades based on weekly and cumulative activities," *2017 Intell. Syst. Conf. IntelliSys 2017*, vol. 2018-Janua, no. September, pp. 462–469, 2018, doi: 10.1109/IntelliSys.2017.8324334.

[17]   B. Hmedna, A. El Mezouary, O. Baz, and D. Mammass, "A machine learning approach to identify and track learning styles in MOOCs," *Int. Conf. Multimed. Comput. Syst. -Proceedings*, vol. 0, pp. 212–216, 2017, doi: 10.1109/ICMCS.2016.7905606.

[18]   K. F. Hew, X. Hu, C. Qiao, and Y. Tang, "What predicts student satisfaction with MOOCs: A gradient boosting trees supervised machine learning and sentiment analysis approach," *Comput. Educ.*, vol. 145, p. 103724, 2020, doi: 10.1016/j.compedu.2019.103724.

[19]   N. I. Jha, I. Ghergulescu, and A.-N. Moldovan, "OULAD MOOC Dropout and Result Prediction using Ensemble, Deep Learning and Regression Techniques," doi: 10.5220/0007767901540164.

[20]   S. Shrestha and M. Pokharel, "Machine Learning algorithm in educational data," Nov. 2019, doi: 10.1109/AITB48515.2019.8947443.

[21]   Y. Mourdi, M. Sadgal, W. B. Fathi, and H. El Kabtane, "A machine learning based approach to enhance MOOC users' classification," *Turkish Online J. Distance Educ.*, vol. 21, no. 2, pp. 54–68, 2020, doi: 10.17718/TOJDE.727976.

[22]   L. Wang, G. Hu, and T. Zhou, "Semantic analysis of learners' emotional tendencies on online MOOC education," *Sustain.*, vol. 10, no. 6, 2018, doi: 10.3390/su10061921.

[23]   W. Feng, J. Tang, and T. X. Liu, "Understanding dropouts in MOOCs," *33rd AAAI Conf. Artif. Intell. AAAI 2019, 31st Innov. Appl. Artif. Intell. Conf. IAAI 2019 9th AAAI Symp. Educ. Adv. Artif. Intell. EAAI 2019*, pp. 517–524, 2019, doi: 10.1609/aaai.v33i01.3301517.

[24]   C. Li and H. Zhou, "Enhancing the efficiency of massive online learning by integrating intelligent analysis into MOOCs with an Application to Education of Sustainability," *Sustain.*, vol. 10, no. 2, 2018, doi: 10.3390/su10020468.

[25]   F. M. Nuraliev, U. E. Giyosov, and Y. Okada, "Enhancing Teaching Approach with 3D Primitives in Virtual and Augmented Reality," in *World Conference

Intelligent System for Industrial Automation, Springer, Cham., 2021, pp. 155–163.

[26]   S. A. Becker *et al.*, *NMC Horizon Report: 2017 Higher Education Edition*. 2017.

Biographies

Ishteyaaq Ahmad is an academician in Uttaranchal Institute of Management, Uttaranchal University, Dehradun. His areas of interest include Online education, Software Engineering and Artificial Intelligence. He has to his credit a MOOC developed for SWAYAM platform and two e-Learning courses developed for Commonwealth of Learning (COL). He is also a recipient of Young Scientist and Professional research grant funded by Department of Science & Technology, Government of India.

Sonal Sharma is Professor & head-Department of Computer Applications, Uttaranchal University with an experience of more 21 years. She possesses strong leadership and administrative skills. Her research interest is in Data Warehousing, Data Mining, Machine Learning, Big Data Analytics and Data Sciences. She has presented/published 45 research papers in national and international conferences/seminars/journals, out of which 11 are Scopus Indexed. She has presided over as a speaker in IEEE conference at Chengdu, China.

Gulsun Kurubacak is a senior professor in Distance Education at the College of Open Education of Anadolu University.Dr. Kurubacak undertook graduate studies at Anadolu University, Turkey (MA. Educational Technology) and the University of Cincinnati, USA (Ed.D. Curriculum & Instruction), and also has worked a post-doctoral fellow at the College of Education at New Mexico State University, USA. She has abundance experience with R&D as well as Accreditation and Quality in open and distance learning at the College of Open Education.

Ajay Kumar working as Assistant Professor in Bharati Vidyapeeth (Deemed to be university) Institute of Management &Research New Delhi. He has more than 10 years of experience in teaching and his area of interest is Network Security and Machine Learning. He is Pursuing PhD. from Mewar University.

Rajiv Kumar is Associate Professor in Uttaranchal Institute of Management, Uttaranchal University, Dehradun. He is B.Tech, M.Tech in Computer Science and Engineering, PhD in computer science and Engineering. Dr. Rajiv Kumar has published several research papers in national and international journals of repute. His research interest includes Genetic Algorithm, Memetic Algorithm.

Sikha Ahmad is an academician in School of Social Sciences, Doon University, Dehradun. She is gold medalist in Economics from Gauhati University. She specializes in Mathematical Economics and Econometrics. Her research interests include Online learning, Behavioural Economics, Neuroeconomicsand Tourism. She is also the Course Coordinator of a popular MOOC on SWAYAM.