JP (Jurnal Pendidikan) : Teori dan Praktik
Vol. 8, No. 1, April 2023, pp. 29-36
E-ISSN: 2527-6891, DOI: https://doi.org/10.26740/jp.v8n1.p29-36

Bibliometric Analysis of Machine Learning on Development Research for Education in Indonesia

Wigati[1]
[1]Technology of Education, Universitas Negeri Jakarta, Jakarta, Indonesia

Article Info

Article history:

Received December 23, 2021
Revised April 24, 2023
Accepted April 30, 2023

Keywords:

Bibliometric
Publication
Machine learning
Education
Scopus

ABSTRACT

Artificial Intelligence – Machine Learning has great potential to help the development of education in Indonesia. So, it is important to map the research that has been done, whereas it can be a strategy to explore the use of Machine Learning for efficient the development of education. This research was conducted to determine the development of publications through article content, authorship patterns, and author productivity. The method used is descriptive bibliometric analysis, data source from the Scopus indexer, while data collection is carried out with Publish or Perish software and Vos Viewer as bibliometric analysis media. The following research findings: (i) Articles about Machine Learning in Indonesia are spread in various journals. And showing significant improvements with various new themes, the fields that most often appear are those related to computer science or Machine Learning development processes. (ii) In general, in 2009-2021 as many as 446 researchers produced 533 articles. The most research publications in 2020, namely 161 articles (30.1%). And related to the field of education there are only 14 articles (2.6%). (iii) The Keywords related to popular fields/subjects are Indonesian, development, prediction, approach, international conference. While keywords related to education produce higher education, education technology, smart educational robots and engineering education.

Corresponding Author:

Wigati
Technology of Education, Universitas Negeri Jakarta, Jakarta, Indonesia
Email: wigati10@gmail.com

1. INTRODUCTION

Artificial Intelligence is an attempt to imitate human intelligence, so Machine Learning is part of Artificial Intelligence, because Machine Learning is an algorithm to make machines understand data about an object, including its environment, then can analyze it to make decisions for it [1]. In Indonesian academic education such at the Bandung Institute of Technology, Artificial Intelligence research has been going on for almost 30 years. Various areas of Artificial Intelligence research that have been carried out at ITB include Machine Learning, intelligent control/robotics, speech processing, natural language processing, knowledge representation and reasoning, computer vision, expert systems, search methodologies, intelligent instrumentation.

LIPI, the Indonesian Institute of Sciences, as the No. 1 Research Institute in Indonesia, began planning to apply machine learning to its library system in 2004. Thus, the process of accepting Artificial Intelligence has been going on for a long time in Indonesia.

IBM (International Business Machines) in 2020 explained the direction of Machine Learning is as a result of the analysis of existing data, and new data. Machine Learning will continuously train the algorithm until the model is smart enough to be able to make its own decisions in the future. Machine Learning is then a subfield of computer science concerned with building algorithms that, in order to be useful, rely on a collection of examples of several phenomena [2]

Now we have felt the benefits of applying Machine Learning in everyday life that is close to us, such as:

Search Engine usage and results, product recommendations while in the market, social media user recommendations (People You May Know, Face Recognition, Posts on Homepage), Digital Ad Display Content (AdSense), Virtual Personal Assistant (google assistant), Spam Message Filter, Online Fraud Detection (as used by PayPal), Video Surveillance (cctv), Medical Diagnosis, Auto Car and other developments.

Commercial companies have also developed massive use of artificial intelligence. Companies such as Google, Apple, Facebook, Amazon, and Alibaba that have databases certainly play a big role. But manufacturing companies and other sectors are also seeing performance optimization opportunities by using Artificial Intelligence – Machine Learning.

Research Center for International and Strategic Studies (CSIS) in 2019 stated that only 7% of manufacturing companies in Indonesia use artificial intelligence (AI) technology. There are three factors causing the lack of technology adoption in the country. First, the company's ability to see the business potential of adopting AI or other technologies is still low. Second, large capital is needed to implement AI and other advanced technologies. Third, skills and talents are lacking. In its development, Machine Learning has also been applied in global practical technology that is able to help everyday life. Both in the fields of geology, medicine, business, engineering, social and education. Then is the relationship between education and Machine Learning needed, especially in Indonesia?

Education in accordance with Education Law no. 20 of 2003 concerning the National Education System is a conscious and planned effort to create a learning atmosphere and learning process so that students actively develop their potential to have spiritual strength. religion, self-control, personality, intelligence, noble character, and skills needed by himself, society, nation and state.

So according to the country's goals and Machine Learning capabilities, Machine Learning will be able to help the learning process. Among them: automation of school administration, personalization of learning, evaluation, learning content that adapts to student needs. As a result, the learning process can be accessed by all students at all times and education will be easy even for students with special needs (disabled). [3]

Machine Learning is the latest in a long effort to distill human knowledge and reasoning into a form suitable for building machines and automated engineering systems.[4]. Machine Learning and Artificial Intelligence will be very helpful, because later it will be able to personalize the history of a lot of data from a student, data is collected from the beginning of school, so that his behavior, interests, talents will be used for guidance on choosing a suitable career later [5]; [6]). For coverage, Artificial Intelligence can be applied at every level of learning, this is proven in research [7] which found that age, is not a significant factor that affects the success of applying Artificial Intelligence, but the support of environments such as homes and schools. Examples of Machine Learning applications that have been carried out are being able to analyze student data on a large scale to produce predictive models to determine preventive and adaptive interventions [8]

Such as making symptom predictions whether students will drop out [9]. Seeing its potential and usefulness in the present, it is certain that in the future its use will be more practical and widespread. So it is important for Indonesia to be able to utilize this technology to solve educational problems such as inadequate school facilities and infrastructure, lack of teaching materials in terms of diversity and richness of content, and learning solutions for minority groups of people with disabilities [10]

For this reason, it is necessary to conduct high research both in quantity and quality and then published scientifically to encourage discovery and utilization in the field of Machine Learning for Indonesian education. What also needs to be done is to map Artificial Intelligence and Machine Learning research in Indonesia, then see how far progress related to education is. For such purposes used bibliometric method.

If referring to previous research, during the period 2001-2021 there has been no research that specifically maps the progress of Machine Learning research for the field of education in Indonesia. So this research answers the following problems: (1) how the development of the number of international scientific publications in the field of machine learning in Indonesia from 2001 – 2021 in Scopus; (2) what is the productivity level of machine learning researchers; and (3) how to map the progress of international publications for machine learning research based on keywords. In particular, this study also aims to find out: (a) the development of the number of international publications in the field of machine learning in Scopus from 2001 – 2021; (b) core journals in international publications in the field of machine learning; (c) productivity of researchers in the field of machine learning; (d) the development of international research publications in the field of machine learning by subject/field; (e) international publication development map for machine learning research and (f) author (g) education keyword cluster (h) list of education-related articles.

2. METHOD

This study used data from international publications in the field of Machine Learning sourced from the Scopus database (www.scopus.com). Data collection through searching for publications using Publish or Perish version 7 then using Machine Learning keywords in Indonesia with article title categories, ABSTRACT, keywords in the period 2001 – 2021, it turns out that the research was only recorded in 2009, so data collection was also carried out from 2009-2021 which was carried out annually for more precise and thorough results. Development trends & productivity of authors of international publications in the field of machine learning were analyzed using VosViewer software.

3. RESULTS AND DISCUSSION

3.1. The increasing number of international publications in the field of machine learning on Scopus from 2001 – 2021

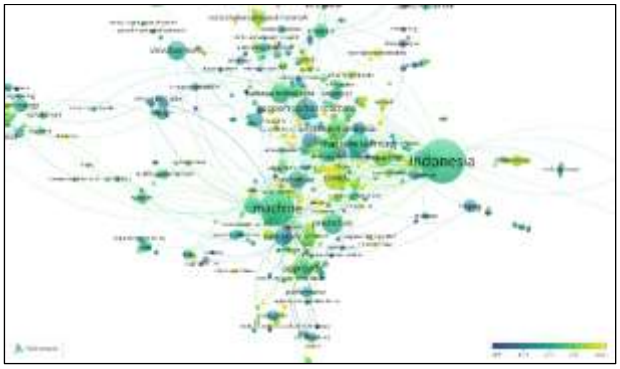


Figure 1. Map overlaying interconnected topics.

From Figure 1 it can be seen that the brighter the item, the newer the topic under study, so covid will be the most talked about theme in 2021. Overall, bright colors predominate, so many new themes are researched.

Table 1. The Number of Articles

Year	The Number	Percentage
2021	119	22.2
2020	161	30.1
2019	116	21.7
2018	62	11.5
2017	35	6.5
2016	20	3.7
2015	3	0.5
2014	6	1.0
2013	7	1.2
2012	5	0.8
2011	2	0.3
2010	2	0.3
2009	1	0.1
Total	533	

From table 1 it can be seen that research is increasingly in demand with a significant increase every year, the most research conducted in 2020 was 30.1% or 161 articles, up 8% from the previous year 2019 which published 116 articles. The temporary decline occurred in 2021 with only 22.2% percent of publications or 119 articles.

Table 2 contains the journals that publish the most articles related to the use of machine learning, (b) core journals in international publications in the field of Machine Learning. There are 5 journals that are the core or largest grouping of all 533 articles.

Table 2. Most Publish Journals on Machine Learning

Journal Name	Article(s)
Journal of Physics: Conference Series	59
IOP Conference Series: Earth and Environmental Science	21
IOP Conference Series: Materials Science and Engineering	22
Procedia Computer Science	13
ACM International Conference Proceeding Series	13

The density visualization figure in the figure below shows (c) the productivity of researchers in the field of machine learning, showing there are 446 authors.

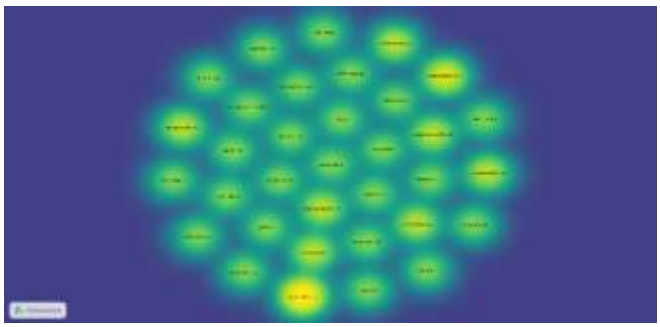


Figure 2. Density Visualization on Researcher Productivity

And 9 of them have written 3 articles seen in the density visualization below.

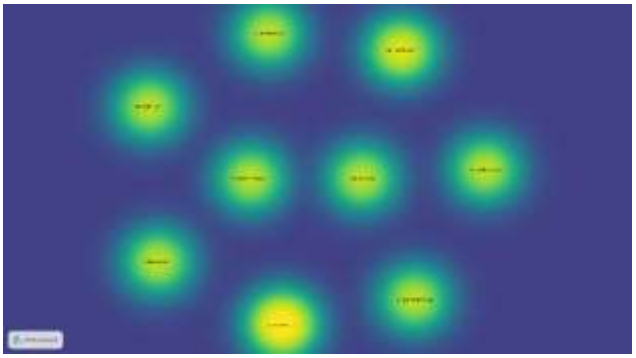
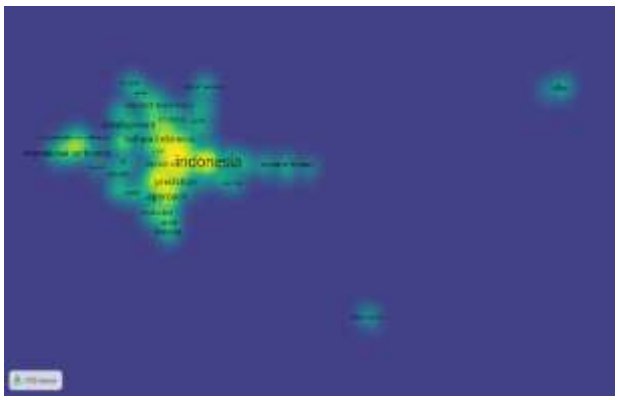
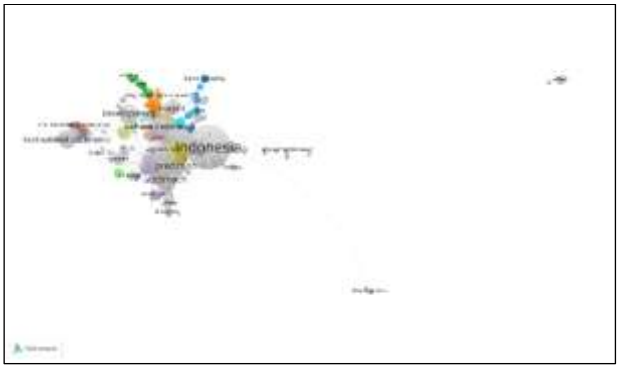


Figure 3. Density Visualization on 3 articles

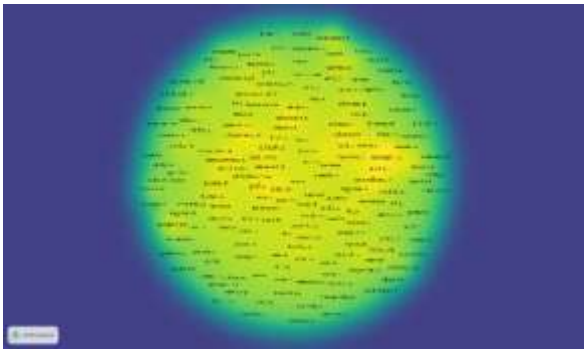
(d) development of international research publications in the field of Machine Learning based on subjects / fields. There are 1158 keywords used and they are divided into 71 linked keyword clusters. The number of keywords shows that there are many variations of themes and subthemes studied by the author. As seen in Figure D.



(e) map of the development of international research publications in the field of machine learning depicted in figure E.



(f) the author is depicted in picture F. From the picture below it can be seen that there are still many variations of themes that have not been studied much. There are even themes that are outliers of the main theme, this shows a wide scope of themes.



(g) Based on the table below, there are 5 clusters of Education keywords with 1 article published each, namely: higher education student cluster, higher education cluster, technology education cluster. Smart Education Robot Cluster, Engineering Education Cluster. (h) List of articles related to the use of Machine Learning for education in Indonesia. There are 14 articles written by 51 authors. And affiliated with 8 universities namely: University of Indonesia (4 articles), Universitas Pendidikan Indonesia (3 articles), Universitas Negeri Solo (2 articles), Institut Sains dan Teknologi Terpadu (1 article), Universitas Pelita Harapan (1 article), University of Johor Malaysia (1 article), Unipdu Jombang (1 article). Bina Nusantara University (1 article). With a range of 2017 – 2021, the most articles published in 2018 were 4 articles, in 2019 and 2021 there were 3 articles each, and in 2017 and 2020 there were 2 articles each. Reflected in Figure H, here:

Table 3. Articles on the use of Machine Learning

Title	Year	Author(s)	Affiliation
Development of Al-Quran sign language classification based on convolutional neural network	2021	Muhamad Zulhairi et al	Johor, malaysia
Development and validation of science, technology, engineering and mathematics (STEM) based instructional material	2017	Ineu Gustiani et al	UPI
AI programming by children using snap! Block programming in a developing country	2018	Ken Kahn et al	Oxford, UPI
Automatic Essay Grading for Bahasa Indonesia with Support Vector Machine and Latent Semantic Analysis	2019	Anak Agung Putri Ratna et al	UI
Detection of Potentially Students Drop out of College in Case of Missing Value Using C4.5	2019	Mutrofin et al	Unipdu, Jombang
Design of Intelligent Robot as A Tool for	2018	Zuhrie, M. S et al.	UNS

Title	Year	Author(s)	Affiliation
Teaching Media Based on Computer Interactive Learning and Computer Assisted Learning to Improve the Skill of University Student			
Learning Management System (LMS) model based on machine learning supports 21st century learning as the implementation of curriculum 2013	2019	B L Putro et al.	UPI
Prediction of number of participants in the online examination system at universitas terbuka	2020	Iwan Susanto, Gede Putra Kusuma	Binus
Data analytics of students' profiles and activities in a full online learning context	2020	Tuti Purwoningsih et al	UI
Predicting Student's Failure in Education Based on Dropout Status	2021	Wildan Adji Nasrullah, et al	ITS
Developing Question Generation System for Bahasa Indonesia Using Indonesian Standard Language Regulation	2021	Maulana Wisnu Prabowo, et al	UI
Mobile sign language recognition for Bahasa Indonesia using convolutional neural network	2018	I. Made Murwantara, Jessica Sean	

From these 14 articles, they can be grouped into 3 main sections about machine learning for education: E-learning, coding, and robotics. In the field of making E-learning there are 3 areas, namely: LMS, assessment and personalization.

There is also research on making LMS prototypes at the beginning of the implementation of the 2013 curriculum which focuses on the concept of machine learning. From the research that has been done, the following results and conclusions are obtained: (1) the use of Machine Learning in the Learning Management System (LMS) Model supports 21st century learning.

The utilization of machine learning in the LMS follows the Online Collaboration, Behavior Tracking, and Learning Analysis Functions; (2) development of an LMS model for the use of ITS 21st century learning architecture (Intelligence Tutoring System). The artificial intelligence domain can be added to every component of the ITS model, both Pedagogic modules, Student Modules, and Interface Modules; (3) The implementation of the 21st century LMS learning model is carried out in SMK in the city of Bandung. [11]. Machine learning in its use for online college exams, namely at the Open University of Indonesia, is also being investigated. Universitas Terbuka (UT) whose students are spread throughout Indonesia, both in big cities and in remote areas. UT as a university has held Final Examinations (UAS) in the form of written examinations since 1984 and has been held in all districts / cities throughout Indonesia.

Over time, there are still students who cannot take UAS because they are constrained by exam dates, so students cannot take exams, even though students do not take UAS, students automatically do not pass the course and will increase their study period. UT strives to provide opportunities for students to be more flexible in conducting end-of-semester exams, so that the Online Examination was developed under the name Online Examination System (SUO). Machine Learning is then used to set SUO implementation quotas to be more effective [12]. As an advanced evaluation, Machine Learning

used to create a database of questions from Indonesian textbooks, such as the Generate Question study. In education, teachers will often ask students to measure each student's level of understanding. Meanwhile, the need for self-evaluation for students requires a system capable of automatically generating questions whenever they want. This study tries to encourage research on question generation, specifically Indonesian. Standard Grammar Indonesian (TBBI) is a set of standard language rules in Indonesian, which is used in this work to convert declarative sentences into question sentences. This method allows for more targeted pattern matching to turn statements into questions. Using this method, the system is able to generate up to 5,000 different questions, with just one book [13]

Research on the use of Machine Learning for essay assessment Indonesian also conducted. To evaluate students' knowledge, they conduct tests that can be done at home in various ways, such as practice, oral, and written in both multiple-choice and essay forms. A few years ago, a process for assessing student evaluations was developed. Written tests such as multiple choice can be graded and evaluated by computer while written essays cannot. Teachers must evaluate written tests in the form of essays manually. It is not effective and

efficient because it will take time and will reduce the objectivity of the essay. Therefore, an essay grading system is needed. The research was used to improve the accuracy of Indonesian vignette assessment automatically by using Support Vector Machine [14]

One of the main characteristics of machine learning is personalization, the ability to adapt to user needs and capabilities. This can be seen from the creation of LMS for teaching materials for deaf friends, where a sign language mobile application was made into the Indonesian. This study was conducted because communication between speech impaired and normal people is an important factor for wider interaction. Hand signals have standards for communicating. However, it is not easy for ordinary people to learn and understand hand sign language. So there is a need to bridge this communication gap by developing mobile apps for android. This application uses the Convolutional Neural Network method, the Mobile Net algorithm on Machine Learning Framework and Tensor Flow. [15]. Even research on Quran teaching materials for the deaf is also carried out. This research needs to be done because sign language is the main form of communication used by deaf people. Most of their activities, such as; speaking, reading, and learning, involving sign language. To read the Qur'an, deaf people use Arabic sign language to recite verses of the Qur'an. For them, assistive technology to assist them in the process of learning and teaching the Qur'an is very important, because traditional methods are very difficult and challenging. One reason is that, traditionally, teachers need to know Arabic Sign Language (ArSL) first to teach them to study the Qur'an. At present, assistive technology is still relatively new and not well developed. In Malaysia and Indonesia, most of the advanced technologies are mobile apps, and web-based devices, both of which require a continuous internet connection and are only suitable for personal use. Previous research on assistive technologies can be classified into two types of devices. First, sensor-based devices, and second are image-based devices [16]. Proper analysis of E-Learning for students can help teachers understand students. What learning contexts and experiences are best suited for E-learning students to improve learning outcomes. However, E-learning teachers often have difficulty in analyzing student data due to the large number of students to be analyzed and limited data. To support research in this area, we conducted a descriptive analysis of a dataset containing data on students from Open and Distance Learning (ODL) who conducted E-learning. The dataset contains data on student demographic profiles and student activities or behaviors during E-learning recorded in the LMS system at Universitas Terbuka Indonesia. [3]

As a result, Machine Learning can also be used to predict problem students and detect their likelihood of dropping out. This study aims to provide projections on the reasons why students drop out of school using machine learning techniques [17]

The reputation of a university can be determined by the number of students who drop out. Therefore, analyzing the problem of missing value data, can reflect the basis why students drop out of school or students who have the potential to drop out of school [18]

Overall, Machine Learning in artificial intelligence has also been studied for its usefulness as a teaching media tool in the form of robotics, because the purpose is used as a demonstration, also attempted, this robot is cheap and applicable. Such as research conducted at the Department of Electrical Engineering, Surabaya State University to develop robots to produce learning devices in the form of modules and kits of smart education robots based on Contextual Learning [19]. With various advances in Machine Learning application applications, it is also necessary to have direct interaction from an early age between students and Machine Learning, the closest way is to teach coding to children, one of the coding available is free software called Snap!, so that even in a developing country like Indonesia, children can be exposed and accustomed to machine learning from the start. As has been researched at SMKN 2 Cimahi, SMK Pelita, SMAN 27 Bandung with funding and collaboration between the Indonesian University of Education and the European Union Horizon 2020 Research and Innovation Coodination &; Action [20]

4. CONCLUSION

Based on the results and international publications in the discussion, it can be concluded that the development of Machine Learning research in 2009 – 2021 with the highest Scopus index occurred in 2020 which reached 161 publications (30.1%). Most of the articles are published by the Journal of Physics Conference. Authors named Rustam Z, Ramadhani F, Kurniawan R, Sinambela M, Yugopuspito P, Suhartono, Liawatimena S, Saragih GS, Toharudin T are the most profile Indonesian writers in the field of Machine Learning with the most fields in Computer Science. The coworks-based machine learning development map is grouped into 5 clusters and co-authors are grouped into 9 clusters. Affiliation with Education can be seen by the presence of authors from the University of Indonesia 4 articles, ITS 1 article, UPH 1 article, Malaysia 1 article, UPI 3 articles, Unipdu Jombang 1 article, UNS 2 articles, Binus 1 article.

A significant increase from year to year shows interest. Machine learning applications are also increasing globally, so these 2 meeting points actually illustrate that artificial intelligence – machine learning

is the future which is an important part of future success and competitiveness. Collaboration between educators and machine learning will improve the quality of Indonesian education.

The authors suggest the need for cooperation between research institutions, industry and academic institutions to increase the quantity of writing and publishing scientific articles on machine learning, especially its use in the field in various aspects of education.

REFERENCES

[1] D. Wankhede, "Artificial Intelligence and its Subsets: Machine Learning and its Algorithms, Deep Learning, and their Future Trends," vol. 9, p. page no.i112-i117, Jun. 2022, doi: 10.6084/m9.jetir.JETIR2205914.

[2] A. Burkov, *The hundred-page machine learning book*, vol. 1. Andriy Burkov Quebec City, QC, Canada, 2019.

[3] I. Saifudin and W. Suharso, "Pembelajaran e-learning, pembelajaran ideal masa kini dan masa depan pada mahasiswa berkebutuhan khusus," *JP (Jurnal Pendidikan): Teori dan Praktik*, vol. 5, no. 2, pp. 30–35, 2020.

[4] M. P. Deisenroth, A. A. Faisal, and C. S. Ong, *Mathematics for machine learning*. Cambridge University Press, 2020.

[5] T. Kabudi, I. Pappas, and D. H. Olsen, "AI-enabled adaptive learning systems: A systematic mapping of the literature," *Computers and Education: Artificial Intelligence*, vol. 2, p. 100017, 2021.

[6] H. Kishan Das Menon and V. Janardhan, "Machine learning approaches in education," *Mater Today Proc*, vol. 43, pp. 3470–3480, 2021, doi: https://doi.org/10.1016/j.matpr.2020.09.566.

[7] E. L.-C. Law and M. Heintz, "Augmented reality applications for K-12 education: A systematic review from the usability and user experience perspective," *Int J Child Comput Interact*, vol. 30, p. 100321, 2021.

[8] K. Zhang and A. B. Aslan, "AI technologies for education: Recent research & future directions," *Computers and Education: Artificial Intelligence*, vol. 2, p. 100025, 2021, doi: https://doi.org/10.1016/j.caeai.2021.100025.

[9] J. Niyogisubizo, L. Liao, E. Nziyumva, E. Murwanashyaka, and P. C. Nshimyumukiza, "Predicting student's dropout in university classes using two-layer ensemble machine learning approach: A novel stacked generalization," *Computers and Education: Artificial Intelligence*, vol. 3, p. 100066, 2022.

[10] Y. A. Purbasari, W. H. Hendriani, and N. H. Yoenanto, "PERKEMBANGAN IMPLEMENTASI PENDIDIKAN INKLUSI," *JP (Jurnal Pendidikan): Teori dan Praktik*, vol. 7, no. 1, pp. 50–58, 2022.

[11] B. L. Putro, R. R. J. Putra, and N. S. Fitriasari, "Learning Management System (LMS) model based on machine learning supports 21st century learning as the implementation of curriculum 2013," in *Journal of Physics: Conference Series*, IOP Publishing, 2019, p. 032032.

[12] Iwan Susanto and Gede Putra Kusuma, "Prediction of Number of Participants in The Online Examination System at Universitas Terbuka," *International Journal of Advanced Trends in Computer Science and Engineering*, vol. 9, no. 5, pp. 7741–7747, Oct. 2020.

[13] M. Wisnu Prabowo, I. Budi, and H. Budi Santoso, "Developing Question Generation System for Bahasa Indonesia Using Indonesian Standard Language Regulation," in *2021 10th International Conference on Software and Computer Applications*, in ICSCA 2021. New York, NY, USA: Association for Computing Machinery, 2021, pp. 258–261. doi: 10.1145/3457784.3457824.

[14] A. A. P. Ratna, H. Khairunissa, A. Kaltsum, I. Ibrahim, and P. D. Purnamasari, "Automatic essay grading for Bahasa Indonesia with support vector machine and latent semantic analysis," in *2019 International Conference on Electrical Engineering and Computer Science (ICECOS)*, IEEE, 2019, pp. 363–367.

[15] P. Yugopuspito, I. M. Murwantara, and J. Sean, "Mobile Sign Language Recognition for Bahasa Indonesia Using Convolutional Neural Network," in *Proceedings of the 16th International Conference on Advances in Mobile Computing and Multimedia*, in MoMM2018. New York, NY, USA: Association for Computing Machinery, 2018, pp. 84–91. doi: 10.1145/3282353.3282356.

[16] M. Z. M. Nizam, S. M. Saad, M. A. Suhaimi, M. A. M. Dzahir, S. Z. A. Rahim, and M. A. M. Dzahir, "Development of Al-Quran sign language classification based on convolutional neural network," in *AIP Conference Proceedings*, AIP Publishing LLC, 2021, p. 020205.

[17] W. A. Nasrullah, J. P. Sugiono, J. Santoso, and A. D. Gunawan, "Predicting Student's Failure in Education Based on Dropout Status," in *2021 3rd East Indonesia Conference on Computer and Information Technology (EIConCIT)*, IEEE, 2021, pp. 183–188.

[18] S. Mutrofin, A. M. Khalimi, E. Kurniawan, R. V. H. Ginardi, C. Fatichah, and Y. A. Sari, "Detection of Potentially Students Drop Out of College in Case of Missing Value Using C4. 5," in *2019 International Conference on Sustainable Engineering and Creative Computing (ICSECC)*, IEEE, 2019, pp. 349–354.

[19] M. S. Zuhrie, I. Basuki, I. G. P. Asto B, and L. Anifah, "Design of Intelligent Robot as A Tool for Teaching Media Based on Computer Interactive Learning and Computer Assisted Learning to Improve the Skill of University Student," *IOP Conf Ser Mater Sci Eng*, vol. 296, no. 1, p. 012049, 2018, doi: 10.1088/1757-899X/296/1/012049.

[20] K. Kahn, R. Megasari, E. Piantari, and E. Junaeti, "AI Programming by Children using Snap," *Block Programming in a Developing Country*, 2018.

BIOGRAPHIES OF AUTHORS

Wigati is a master student in the Department of Educational Technology at Universitas Negeri Jakarta. With a background as an information technology consultant, her research in the field of education currently focuses on using the latest technology to improve the quality of learning & training, such as the use of Artificial Intelligence, Machine Learning, or new learning strategies such as Microlearning, Mobile Collaborative Learning. She can be contacted via email: wigati10@gmail.com

.