Journal of Information Systems and Informatics
Vol. 6, No. 3, September 2024 e-ISSN: 2656-4882 p-ISSN: 2656-5935
DOI: 10.51519/journalisi.v6i3.782 Published By DRPM-UBD

Bibliometric Analysis of Data Analytics Techniques in Cloud Computing Resources Allocation

Sello Prince Sekwatlakwatla[1], Vusumuzi Malele[2]

[1,2]North-West University, Mafikeng Campus, South Africa
Email: sek.prince@gmail.com[1], vusi.malele@nwu.ac.za[2]

Abstract

Cloud computing provides on-demand computing services over the Internet, allowing for quicker innovation, more flexible resources, and economies of scale while reducing the need for physical data centers and servers. With this benefit, most organizations are adopting this technology, and some organizations are also operating fully on cloud computing. This causes traffic to increase, and most of these organizations are struggling with resource allocation, resulting in complaints from users regarding inactive system performance, timeouts in applications, and higher bandwidth use during peak hours. In this regard, this study investigates data analytics techniques and tools for the allocation of resources in cloud computing. The study indexed journal articles from the Scopus Database and Web of Science (WOS) between 2010 and 2024. This article brings new insights into the analysis of data analytics techniques in Africa and collaborations with other developing countries. The findings present tools and approaches that may be used to allocate cloud computing resources and give recommendations.

Keywords: Resource allocation, Data Analytics Techniques and Traffic prediction

1. INTRODUCTION

Most organizations are adopting cloud technology due to its benefits, one of the benefits of cloud computing is that it allows automation in the cloud to eliminate human error, increase resource efficiency and improve system adaptability by eliminating manual labor [1-2]. Organizations are adopting it more frequently in order to increase efficiency since it allows rote operations to be completed in real time, however allocation of resources remains a challenge most of the time, even if cloud automation is available [3], due to Support varies from vendor to vendor and can be challenging at times business rely on the vendor's development teams and SLAs (service level agreements) to fix any bugs or outages, Although service and uptime are still the provider's responsibility some businesses may find it preferable to have less control. Artificial intelligence (AI) driven diabetes prediction is driven by the recognition of the dynamic nature of risk factors, such as lifestyle choices, environmental exposures, and genetic predispositions this model can dynamically update risk assessments, providing real-time and personalized health monitoring however there is no information about the required medical treatment to solve the situation [4].

It is said in [5] that block chain and edge computing may be combined to overcome each other's shortcomings in resource allocation and traffic protection. The potential advantages of block chain technology for cloud computing are explored in [6], where cloud services are used to categorize works pertaining to block chain and cloud computing resource allocation. However a cloud-edge hybrid architecture and a double block chain-based transaction model for flexible trust management are presented in [7], which also discusses the works pertaining to block chain-based trust management in cloud computing and the limitations of the model overfitting of resource allocation and wasting the cost of the organizations [8]. In this regard, cloud service providers can use the data analytics technique to enhance resource allocation in cloud computing. Support vector machines are effective in predicting traffic casualties and supercritical water heat transfer processes, with accuracy rates of 0.5894 before and 0.8247 after normalization. However, they struggle with long-running time and complex attribute classification [9].

The push-and-pull search architecture proposed for resource allocation in cloud computing for IoT-enabled compute offloading in collaborative edge and cloud computing shows superiority over state-of-the-art methods in handling computation difficulties [10]. In order to satisfy various quality of service (QoS) requirements for the benefit of the guest, machine learning techniques are increasingly being used in cloud-integrated computing paradigms; however, there is limited information about accurate prediction of traffic flow [11]. Artificial neural networks (ANNs) and other soft computing techniques are being used to optimize several QoS parameters in order to address the dynamic scheduling of cloud computing resources. On the other hand, ANN and linear regression (LR) are also being used in adaptive resource provisioning in order to satisfy future resource demands in cloud computing [12-13]. Data from restricted devices and networks is necessary for traffic categorization; however concept drift can alter the properties of the data making it difficult for models trained on certain time periods and scene sets to be applied to open networks [14].

The SARIMA-LSTM model for continuous anomalous traffic detection is proposed in order to allocate the resources. The three algorithms such as random forest, Bayesian network and support vector machine are combined and the model incorporates genetic algorithms and fuzzy logic for traffic anomaly detection. Using a virtual computer, processing models for resource allocation are provided. The difficulties with this approach big data is not well defined, nor is it evident how it employs machines [15]. The author discusses the resource allocation problem in automotive systems using cloud computing they compare public and private cloud standards, focusing on decentralized binding schemes and centralized binding schemes [16]. They demonstrate how task deadlines and workloads affect optimal bidding policies through numerical examples [17].

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

Cloud resource management entails making intricate decisions and regulations that take energy efficiency, resource usage, and application scalability into account. Most IT organizations are unable to address issues like erratic workloads and a lack of global state information while effectively allocating resources to satisfy the needs of a sizable user base. The purpose of this study is to analyze and investigate the data analytics techniques in cloud computing to identify the trend of the resources that are allocated effectively and efficiently and provide recommendations of the techniques that are available and that can be used to allocate the resources effectively. The article also highlighted the importance of the allocation and management of resources in cloud computing using multiple techniques. The remaining portion of the article is organized into four sections: section two discusses the methodology, section three deals with the findings and discussion, and section four provides a quick summary of the conclusion.

2. METHODS

To accomplish the goal of the study, procedures are suggested, as shown in Figure 1.

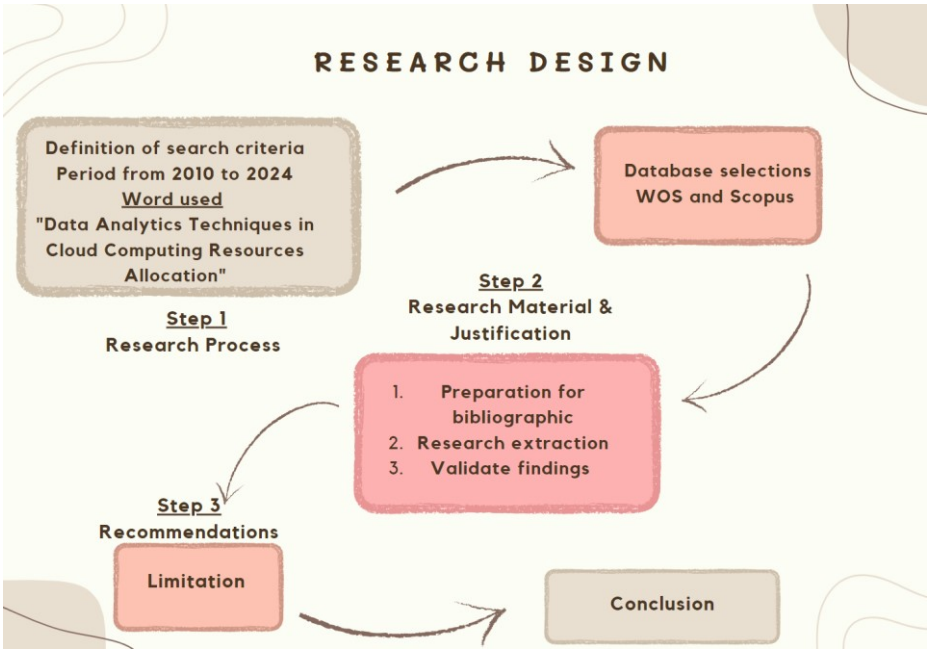

Figure 1. Research Design

2.1 Research Process

The investigation employed various databases, such as Web of Service and Scopus, to explore the field of data analytics techniques in cloud computing resource

allocation. The search was conducted using the specific criteria of "Data Analytics Techniques in Cloud Computing."? The keywords that were identified in the search criteria were traffic flow, data analytics, and resource allocation. The databases were accessed on June 22, 2024. The data collected spanned from January 1, 2010 to June 15, 2024 and encompassed conferences, journals, early-access articles, and magazines that were deemed relevant to the research, Research questions (RQs) were used as the guide for this study. The following research question was utilized: (i) RQ 1: What are data analytics techniques that can be used to enhance cloud computing resource allocation? (ii) RQ 2: What are the cloud computing resource concerns when allocating resources in cloud computing?

2.3 Research Material and Justification

The process consists of three primary steps: preparing the data, analyzing the data, and producing the result. The dataset was also renamed to BibTeX in order to improve bibliographic organization. Moreover, useful data was retrieved, and graphs were downloaded to support decision-making.

2.4 Recommandations

Based on the results of the bibliometric study, the methodologies are identified, and their limitations are discussed in this section.

3. RESULTS AND DISCUSSION

3.1 Data Analytics Techniques and Cloud Computing Resource

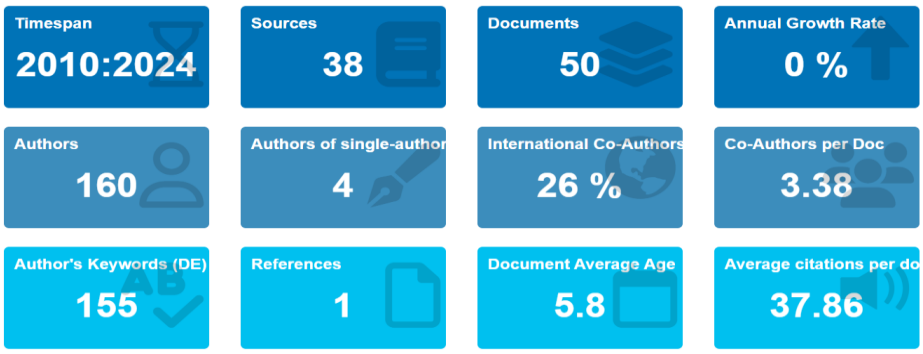

Figure 2. Results summary

The databases were accessed on June 22, 2024. The data collected spanned from January 1, 2010 to June 15, 2024. There are 26% international co-authors and 37.86 average articles per document, and the author's keywords are 155. Over 160

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

authors are cited with over 50 documents and 38 sources; however, there is a 0% annual growth rate, as shown in Figure 2. Big data and data analytics have become important forces behind business intelligence, allowing companies to glean insightful information from enormous databases. The introduction of cloud computing has completely changed how large data processing and data analytics are carried out; in this regard, there is a need to increase the annual growth rate of the data analytics research.

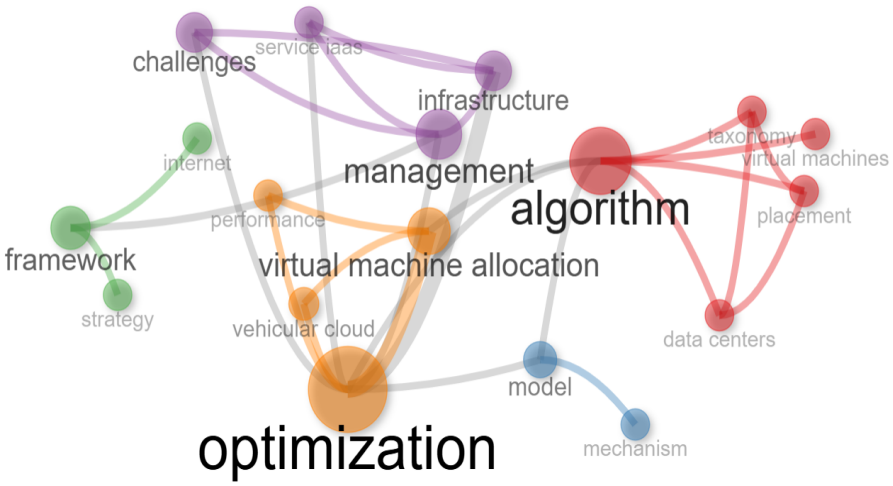

Figure 3. Techniques to improve resource allocation

Resource allocation plays a critical role as computer networks handle complex optimization issues. The goal of resource allocation, as it pertains to this subject, is to maximize throughput, or total computing efficiency in this regards the results of the analysis of current data show that data centers, virtual machine allocation, vehicular cloud for the automotive industry, Infrastructure as a Service (service IAAS), strategy ,framework and mechanism taxonolomy virtual machine, are on the top list of the results to allocate the resources effectively and efficiently , therefore the combination of these techniques may be used to enhance the analysis of data analytics techniques in cloud computing resource allocation, as shown in Figure 3.

A qualitative data analysis technique called thematic analysis involves going over a collection of data and looking for patterns in the meaning of the data in order to uncover themes. In this regard, Figure 4 indicates that data centers, algorithm models and virtual machine allocation support the strategy to enhance the allocation of resources in cloud computing. On the other hand, optimization management challenges are also highlighted as a strategy for optimizing resources, as shown in Figure 4.

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

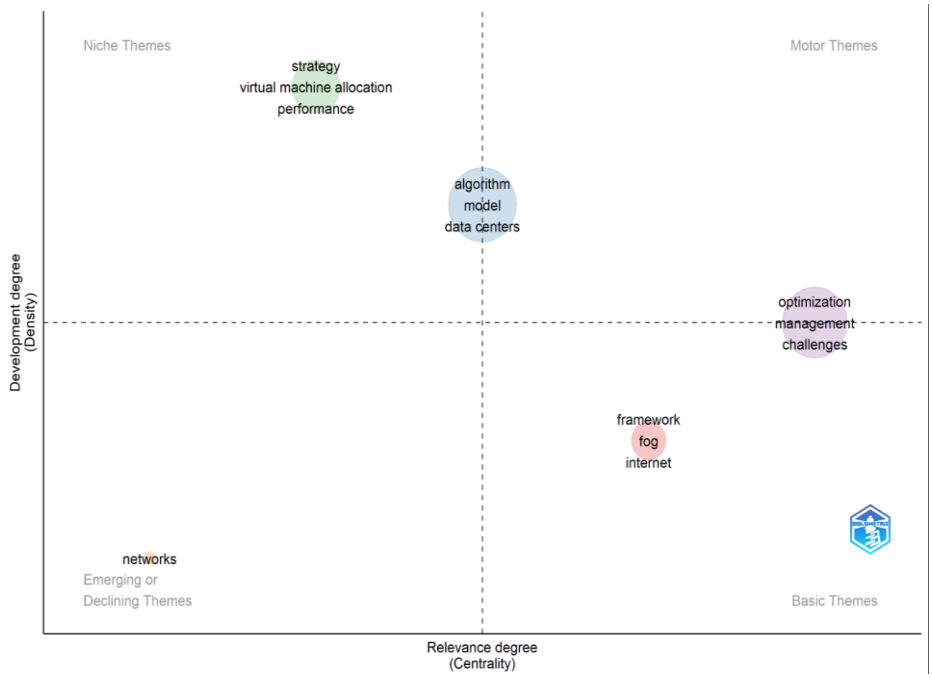

Figure 4. Techniques to improve resource allocation

Most searched words are essential components. Researcher comprehension of the publication's primary themes and subjects to facilitate and also help locate pertinent papers in this regard Most of the searched words results for to assist with the alignment of the study in order to identify trend are resource allocation, computation of fading, energy utilization, edge cloud computing, and also block-chain electronic data interchange, as shown in Figure 5.

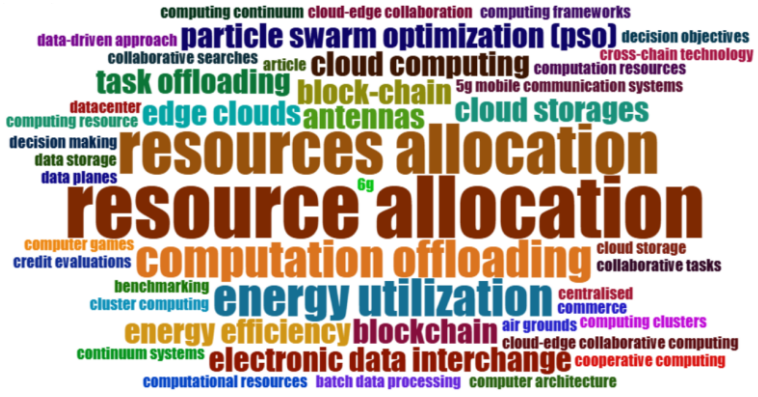

Figure 5. Summary of the most searched words

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

From the resource allocation in cloud computing, there are antennas, block chains, and computation offloading techniques that need to be investigated further as techniques to enhance data analytics techniques in cloud computing resource allocation, in this regard's ensemble method can be used as a tool to enhance and allocate the resources effectively and efficiently, as shown in Figure 6.

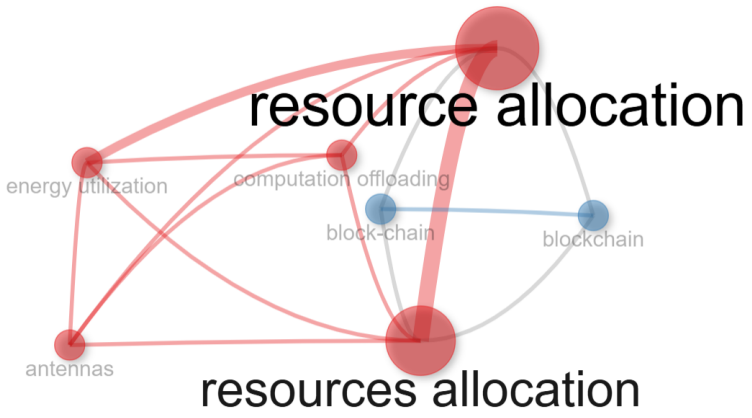

Figure 6 Co-network relationship of resources allocation

The top ten most researched countries in the world for the analysis of data analytics techniques in cloud computing are China with over 3250, followed by India with over 2500 documents, the United States with 15 000 documents, Canada with a record of over 600 and the United Kingdom with about 500 documents, followed by Australia with 400 documents, followed by Italy and Iran with 260 documents, and the last country that is doing this research is France with 260, as shown in Figure 7.

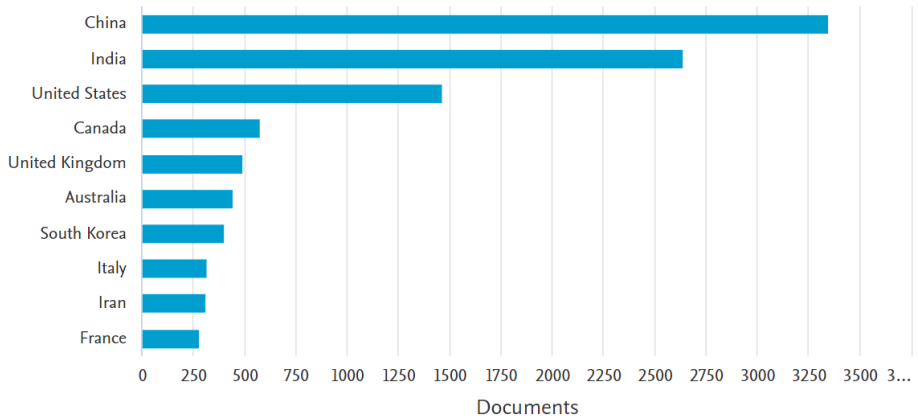

Figure 7 The top ten most researched countries in the world

The top ten most researched countries in Africa for the analysis of data analytics techniques in cloud computing are Egypt with over 115, followed by Tunisia with over 82 documents, Algeria with 78 documents, Morocco with a record of over 67 and South Africa with about 25 documents, followed by Ethiopia with 19 documents, Senegal with nine documents and Ghana with seven documents. The last countries are Angola and Kenya with two documents, as shown in Figure 8.

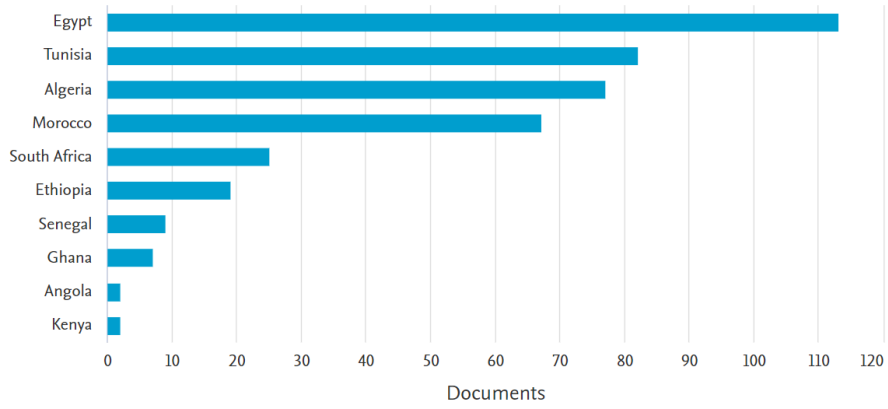

Figure 8 the top ten most researched countries in the Africa

3.2 Discussion

The bibliometric analysis conducted in this study identifies various techniques that can be employed to enhance resource allocation in cloud computing. Resource allocation is a crucial factor in optimizing the performance of computer networks, particularly as they increasingly address complex optimization challenges. The tools and techniques highlighted in this study, such as virtual machine allocation and Infrastructure as a Service (IaaS), are not only well-supported by existing literature but also hold significant potential for improving cloud computing resource management. These methods are vital as cloud computing continues to evolve, requiring ever more efficient strategies to manage growing data and processing demands.

In answering RQ1: What are data analytics techniques that can be used to enhance cloud computing resource allocation?**,** this study identifies several key techniques, including virtual machine allocation, vehicular cloud computing for the automotive industry, and IaaS. Additionally, the application of ensemble methods—where multiple models are combined to improve accuracy and reliability—emerges as a powerful approach to optimizing these techniques. Ensemble methods leverage the strengths of various models, offering a more comprehensive solution for resource allocation challenges in cloud computing.

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

These findings suggest that a combination of these techniques can lead to more effective and efficient use of cloud resources.

RQ2: What are the cloud computing resource concerns when allocating resources in cloud computing? is also addressed through this analysis. One of the primary concerns is the need to handle complex optimization issues efficiently while maintaining high throughput and minimizing costs. The identified techniques, such as those highlighted in Figure 3, focus on addressing these concerns by providing structured frameworks and mechanisms for resource allocation. The use of thematic analysis in this study further supports the notion that optimizing resource allocation is not only about technical efficiency but also about strategic alignment with broader business goals.

The study also highlights that countries like China and India are leading the way in researching cloud computing resource allocation techniques. Their significant output in this field demonstrates a strong commitment to advancing technology and optimizing resource use in cloud computing. This level of engagement from developing countries serves as a benchmark for other regions, particularly Africa, where research output in this area remains relatively low. By increasing focus and resources on cloud computing research, African nations could significantly enhance their technological capabilities, contributing to a more balanced global advancement in this critical area.

The tools identified in this research, including virtual machine allocation, vehicular clouds, and IaaS, are essential for improving resource allocation in cloud computing. These techniques, when combined with ensemble methods, can lead to more sophisticated and effective data analytics tools, optimizing resource allocation more comprehensively. As indicated in Figure 3, integrating these approaches can significantly enhance the efficiency of cloud computing systems, making them more adaptable to the varying demands of different industries and regions.

In conclusion, this study provides valuable insights into the techniques available for enhancing resource allocation in cloud computing, particularly through the use of data analytics and ensemble methods. The findings suggest that while some regions, like China and India, are making significant progress, there is a critical need for increased research and development in other areas, such as Africa. By addressing the identified resource concerns and leveraging the techniques outlined in this study, the global community can work towards more efficient and equitable resource allocation in cloud computing, ultimately supporting the broader adoption and advancement of cloud technologies worldwide.

4. CONCLUSION

This study used a bibliometric analysis of data analytics techniques in cloud computing in order to identify techniques that can be used to enhance resource allocation in cloud computing. The study indexed journal articles from the Scopus Database and Web of Science (WOS) between 2010 and 2024. This article brings new insights into the analysis of data analytics techniques in Africa and collaborations with other developing countries. The results of the analysis of current data show that data centers, virtual machine allocation, vehicular cloud for the automotive industry, Infrastructure as a Service (IAAS), strategy, framework, framework and mechanism taxonolomy virtual machine can be used as ensemble method to enhance cloud computing resource allocation. China and India are leading with research on data analytics techniques in cloud computing; on the other hand, Egypt and Tunisia are leading from the Africa point of view; however, Africa needs to increase research in this area. The limitation of the study is that it only focuses on two databases. For future research, the study will include more databases.

REFERENCES

[1] P. Sharma and V. Jadhao, "Molecular Dynamics Simulations on Cloud Computing and Machine Learning Platforms," *Proc. 2021 IEEE 14th Int. Conf. Cloud Comput.* (CLOUD), Chicago, IL, USA, pp. 751-753, 2021.

[2] G. Peng, H. Wu, H. Wu, and K. Wolter, "Constrained Multiobjective Optimization for IoT-Enabled Computation Offloading in Collaborative Edge and Cloud Computing," in *IEEE Internet Things J.*, vol. 8, no. 17, pp. 13723-13736, 2021.

[3] P. Kumar, A. Tharad, U. Mukhammadjonov, and S. Rawat, "Analysis on Resource Allocation for Parallel Processing and Scheduling in Cloud Computing," *Proc. 2021 5th Int. Conf. Inf. Syst. Comput. Networks* (ISCON), Mathura, India, pp. 1-6, 2021.

[4] J. Kamiri and G. Mariga, "Research Methods in Machine Learning: A Content Analysis," *Int. J. Comput. Inf. Technol.*, vol. 10, no. 2, 2021.

[5] I. Obeidat and A. Mughaid, "Implementing Factors of Information Security in Governmental Organizations of Jordan," *Proc. Thirteenth Int. Conf. Digital Soc. eGovt.*, 2019.

[6] T. Ahmad and N. Zhou, "Ensemble Methods for Probabilistic Solar Power Forecasting: A Comparative Study," *Proc. 2023 IEEE Power Energy Soc. Gen. Meeting* (PESGM), Orlando, FL, USA, pp. 1-5, 2023.

[7] T. Qin, G. Cheng, Y. Wei, and Z. Yao, "Hier-SFL: Client-Edge-Cloud Collaborative Traffic Classification Framework Based on Hierarchical Federated Split Learning," *J. Future Gener. Comput. Syst.*, vol. 149, 2023. doi: 10.1016/j.future.2023.07.001

[8] Z. Zhao et al., "Research on Edge Cloud Architecture of Intelligent Power Plant," *Proc. 2022 IEEE 2nd Int. Conf. Digital Twins Parallel Intell.* (DTPI), Boston, MA, USA, pp. 1-3, 2022.

[9] N. Mohan and V. Jain, "Performance Analysis of Support Vector Machine in Diabetes Prediction," *Proc. 2020 4th Int. Conf. Electron. Commun. Aerosp. Technol.* (ICECA), Coimbatore, India, pp. 1-3, 2020.

[10] N. H. Haay and M. N. N. Sitokdana, "Analysis of Information Technology Governance on Communication and Information Service of Papua Province Using COBIT 2019," *J. Inf. Syst. Informatics*, vol. 4, no. 2, pp. 349–360, 2022.

[11] K. Ntanzi, A. K. Lumbwe, S. C. Mukwakungu, and N. Sukdeo, "The Relationship Between the Implementation of Quality Management Practices and Service Quality in the South African Financial Service Industry," *Proc. 2020 IEEE Int. Conf. Ind. Eng. Eng. Manage.* (IEEM), Singapore, pp. 1235-1240, 2020.

[12] M. Hao et al., "Artificial Neural Network-Based Approach to Modeling Energy Bands of GaN-Based Heterojunction Materials," *Proc. 2023 Int. Conf. High Perform. Big Data Intell. Syst.* (HDIS), Macau, China, pp. 71-76, 2023.

[13] S. Arlis and S. Defit, "Machine Learning Algorithms for Predicting the Spread of Covid-19 in Indonesia," *TEM J.*, vol. 10, no. 2, pp. 970–974, 2021.

[14] Y. Sun and J. Zhu, "Forecasting Chinese Electric Vehicle Volume Using Keyword Search Data Based on SARIMA-LSTM Model," *Proc. 2023 8th Int. Conf. Inf. Syst. Eng.* (ICISE), Dalian, China, pp. 467-472, 2023.

[15] S. K. Panigrahi, "Forecasting Twitter's Favorite Cryptocurrencies Based on a Comparative Assessment of SARIMA, LSTM, and Fb Prophet Machine Learning Approach," *Proc. 2023 Int. Conf. Sustain. Islamic Bus. Finance* (SIBF), Bahrain, pp. 42-45, 2023.

[16] P. Kumar, A. Tharad, U. Mukhammadjonov, and S. Rawat, "Analysis on Resource Allocation for Parallel Processing and Scheduling in Cloud Computing," *Proc. 2021 5th Int. Conf. Inf. Syst. Comput. Networks* (ISCON), Mathura, India, pp. 1-6, 2021.

[17] X. Mou and H. Chen, "Network Information Resource Allocation under Big Data Technology," *Proc. 2023 IEEE 2nd Int. Conf. Electr. Eng. Big Data Algorithms* (EEBDA), Changchun, China, pp. 15-19, 2023.