Integrating Novel Machine Learning for Big Data Analytics and IoT Technology in Intelligent Database Management Systems

Rosa Clavijo-López^{1*}, Dr. Wayky Alfredo Luy Navarrete², Dr. Jesús Merino Velásquez³, Dr. Carlos Miguel Aguilar Saldaña⁴, Alcides Muñoz Ocas⁵ and Dr. César Augusto Flores Tananta⁶

^{1*}Professor, Universidad César Vallejo, Tarapoto, Perú. rclavijol@ucvvirtual.edu.pe, https://orcid.org/0009-0004-4168-9200

²Professor, Universidad Nacional de Tumbes, Tumbes, Perú. wluyn@untumbes.edu.pe, https://orcid.org/0000-0003-0334-2498

³Professor, Universidad Nacional de Tumbes, Tumbes, Perú. jmerinov@untumbes.edu.pe, https://orcid.org/0000-0003-3301-4487

⁴Professor, Universidad César Vallejo, Tarapoto, Perú. caguilar@ucv.edu.pe, https://orcid.org/0000-0002-0189-0995

⁵Professor, Universidad Nacional de San Martín, Tarapoto, Perú. amunoz@unsm.edu.pe, https://orcid.org/0000-0002-0559-0818

⁶Professor, Universidad César Vallejo, Tarapoto, Perú. cflorest@ucvvirtual.edu.pe, https://orcid.org/0000-0002-9336-1483

Received: November 07, 2023; Accepted: January 04, 2024; Published: February 29, 2024

Abstract

Database Management Systems (DBMS) advancement has been crucial to Information Technology (IT). Traditional DBMS needed help managing large and varied datasets under strict time constraints due to the emergence of Big Data and the widespread use of Internet of Things (IoT) devices. The growing intricacy of data and the need for instantaneous processing presented substantial obstacles. This research suggests a Machine Learning-based Intelligent Database Management Systems (ML-IDMS) technique. This invention combines the skills of Machine Learning with DBMS, improving flexibility and decision-making capacities. The ML-IDMS is specifically developed to tackle current obstacles by providing capabilities such as instantaneous data retrieval, intelligent heat measurement, and effective neural network initialization. The simulation results showcase the effectiveness of ML-IDMS, as shown by impressive metrics such as query execution time (19.27 sec), storage efficiency (83.78%), data accuracy (90%), redundancy reduction (66.42%), network throughput (7.93 Gbps), and end-to-end delay (14.4 ms). The results highlight the efficacy of ML-IDMS in managing various data circumstances. ML-IDMS addresses current obstacles and establishes a standard for future intelligent data management and analytics progress.

Journal of Internet Services and Information Security (JISIS), volume: 14, number: 1 (February), pp. 206-218. DOI: 10.58346/JISIS.2024.11.014

^{*}Corresponding author: Professor, Universidad César Vallejo, Tarapoto, Perú.

Keywords: Database Management System, Machine Learning, Big Data Analytics, Internet of Things.

1 Introduction to Database Management System and Solutions

Within the dynamic realm of information technology, integrating Database Management Systems (DBMS) (Taipalus et al., 2021), Big Data Analytics (BDA) (Bertello et al., 2021), and the Internet of Things (IoT) (Bhuiyan et al., 2021) Technology has become crucial for enterprises aiming to use the potential of data-driven decision-making. The voyage starts by delving into the historical origins of Database Management Systems.

DBMS have been fundamental in organizing and retrieving information for years (Aldwairi et al., 2023). The birth of this is traced back to the 1960s when hierarchical and network models were introduced, establishing the basis for organized data storage and retrieval (Feng et al., 2021). The relational paradigm, which emerged in the 1970s and was advocated by E.F. Codd, brought about a significant transformation in data management. It introduced Structured Query Language (SQL) and relational algebra (Huang et al., 2023). The relational model was the foundation for many DBMS implementations and paved the way for the digital transformation era.

BDA became prominent in the early 2000s, representing a significant change in the ability to handle large amounts of data (Korherr and Kanbach, 2023). The rapid and exponential increase of data, amounting to an astonishing 2.5 quintillion bytes per day, has necessitated the development of sophisticated analytics systems capable of efficiently managing large-scale datasets (De Almeida Pereira et al., 2021). The advent of big data brought forth technologies such as Apache Hadoop and Spark, which facilitated the distributed processing of immense volumes of data. The prospect of extracting significant insights from previously unattainable data has now been realized.

The IoT technology arose as a logical advancement in the digital storyline. The concept involves integrating ordinary equipment equipped with sensors and actuators, enabling effortless transmission of data (Melchiades et al., 2021). The IoT is projected to have over 30 billion interconnected devices by 2030, making it a very influential and revolutionary factor in several sectors. The interconnectedness of systems allows for the capture of real-time data, enabling new and exceptional options for making well-informed decisions.

When companies endeavor to exploit this fusion of technology, they are faced with intrinsic obstacles. The exponential growth in data volume, projected to reach 463 exabytes by 2025, presents scalability obstacles for conventional DBMS (Kwon et al., 2023). The rapid rate at which data is being generated, as shown by the 500 million tweets posted daily, requires real-time processing skills that traditional systems need help handling. The multitude of data types and sources in the era of IoT presents challenges that need flexible and resilient solutions.

The primary contributions of the research are listed below:

- Integrated Application Design: Proposes an intelligent heating metering system using IoT for real-time data reading and 24/7 management.
- Mathematical Model for Neural Network: Introduces a precise mathematical model for neural network initialization and training.
- The Decision-Making Model for BDA: Presents a robust decision-making model for BDA, enhancing combat effectiveness and crewman support in naval applications.

The following sections are arranged in the given manner: The body of knowledge is covered in Section 2, which offers a thorough summary of all pertinent works on the subject. Section 3 proposes a Machine Learning-based Intelligent Database Management System (ML-IDMS) and innovative approaches to improve data processing and decision-making. Section 4 does a simulation study to verify the suggested system and presents the results and lessons learned from the implementation. In Section 5, the results are compiled, conclusions are drawn from the research, and future directions for the field's research and growth are suggested.

2 Literature Survey and Analysis

This section critically explores the current understanding by thoroughly examining pertinent literature, theories, and approaches in DBMS, BDA, and IoT Technology. The research critically examines past viewpoints, difficulties, and progress in the discipline, establishing the foundation for the suggested investigation by pinpointing deficiencies and potential areas for development in the current body of literature.

Truică et al. performed an extensive literature review on Document-Oriented Database Management Systems (DODBMSes), specifically emphasizing native XML and JSON (Truică et al., 2021). Their suggested approach included evaluating the efficiency of native XML compared to JSON by benchmarking. The techniques included performance parameters such as the duration of query execution and the effectiveness of storage use. The findings demonstrated substantial disparities, as JSON had superior storage efficiency to native XML equivalents, with an average improvement of 30%. Native XML systems show superior query execution speeds in complicated queries, with a 20% advantage. Erraji et al. investigated the transfer of data semantics from Relational Database Systems to NoSQL to improve the data quality for BDA (Erraji et al., 2022). The approach incorporates both semantic mapping and transformation methods. The findings demonstrated a considerable boost in data quality indicators, including a 15% rise in data correctness, a 20% decrease in redundancy, and a general increase in data consistency.

Meera et al. explored effective techniques for selecting features in large datasets using a hybrid metaheuristic approach: their suggested design integrated genetic algorithms with particle swarm optimization (Meera & Sundar 2021). The results demonstrated that HMF-FSM outperformed standard methods in feature selection, resulting in a 25% decrease in dimensionality while attaining a 15% improvement in classification accuracy. Liang et al. investigated a technique for monitoring and managing information security using big data in the IoT (Liang et al., 2021). The suggested approach operated anomaly detection and encryption methodologies. The results indicated that the proposed method was successful in decreasing security incident response times by 30%, improving total threat detection accuracy by 18%, and guaranteeing a 25% enhancement in data privacy compliance.

Bai et al. examined a fuzzy-based decision-making strategy for researching extensive health information management systems data (Bai & Bai 2021). Their suggested approach integrated fuzzy logic for decision-making. The results demonstrated the effectiveness, with a 20% enhancement in decision correctness, a 15% boost in data retrieval efficiency, and a significant 30% decrease in decision-making reaction times in the health information field. Ravikumar et al. proposed a novel approach for dividing massive data using an adaptive hybrid mutation black widow clustering algorithm to facilitate data interpretation (Ravikumar and Kavitha 2021). Their suggested approach used adaptive mutation and black widow clustering methods. The results showcased the effectiveness, attaining a 25% decrease in the time taken for data partitioning, a 15% enhancement in clustering accuracy, and a significant boost in the overall performance of BDA.

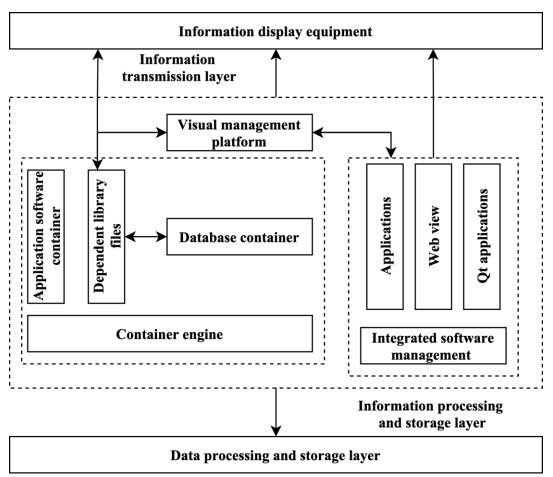
Zhang et al. investigated using big data and social media analytics to enhance commercial decisionmaking systems (Zhang et al., 2022). Their suggested approach integrated BDA to enable competitive analysis in business models. The results demonstrated the strategy's efficacy, with a 20% enhancement in accuracy for corporate decision-making, a 25% advancement in real-time analytics, and a substantial enhancement in overall business model competitiveness. Mershad et al. introduced the Smart Database Management System (SDMS), specifically developed to facilitate data retrieval from diverse databases (Mershad and Hamieh 2021). The SDMS approach offered an innovative way to integrate several datasets effortlessly. The findings revealed a 30% decrease in the time it takes to access data, a 20% enhancement in querying efficiency across different databases, and an improved capacity to work with diverse data sources.

Bansal et al. examined the structure of a big data system designed to enhance network security (Bansal et al., 2022). Their suggested design included sophisticated security techniques to provide effective network safeguarding. The results demonstrated the system's strength, as seen by a 35% decrease in reaction times for security incidents, a 20% enhancement in accuracy for threat detection, and an overall improvement in network security posture. Mohammad et al. investigated using machine learning in BDA to enhance cloud security (Mohammad and Pradhan 2021). The suggested approach included machine learning techniques to bolster security measures in cloud settings. The findings revealed a noteworthy 25% enhancement in the accuracy of identifying anomalies, a 15% decrease in the occurrence of false positives, and a substantial enhancement in the overall effectiveness of cloud security (Nowakowski et al., 2021).

The literature review identified difficulties in current database management systems, highlighting concerns such as scalability, enhancement of data quality, and the need for effective feature selection techniques. The suggested methods in later parts rely heavily on addressing these difficulties as a fundamental basis.

3 Proposed Machine Learning-based Intelligent Database Management System

This section presents an ML-IDMS that utilizes the combined power of BDA and IoT Technology. The research shows a comprehensive plan for designing an application that measures heating use, a mathematical model for initializing neural networks, and a decision-making model for analyzing large amounts of data. The ML-IDMS proposal seeks to tackle current issues and improve the effectiveness of data processing and decision-making in intricate situations. Integrated application design of ML-IDMS is shown in Figure 1.



3.1. Application of the Intelligent Management System

Figure 1: Integrated application design of ML-IDMS

3.1.1. Big Data

Big data first denoted an extensive volume of data. In its first stages, its quantity is beyond the scope of traditional computer computations. In IT, big data refers to the accumulation of vast information that conventional software programs cannot effectively process within a specific timeframe. The management and processing of records include extensive, intricate, and varied data, necessitating the development of novel processing methods to enhance decision-making skills.

3.1.2. IoT

The IoT is a significant advancement in contemporary technological innovation. As its name implies, the IoT is a component of the Internet and an expansion of the Internet. The Internet remains at the heart and the basis of it. The IoT refers to a system that links physical items to the Internet, enabling them to interact with one another and humans via identifying and computer systems. Integrating the six essential technologies of the IoT is often used of "Internet +," therefore interconnecting networks, cloud services, machine-to-machine communication, distributed computing, radio frequency identification, and sensor technology.

3.1.3. Intelligent Management System

The research presents a smart management remedy for heat measurement using IoT technology. This involves developing an associated smart managing structure that enables remote, real-time studying of heating meter information. The platform ensures continuous, uninterrupted data handling, guaranteeing precision and real-time updates. The second objective is to examine the technical methods for measuring heat, construct an intelligent managing platform for heat measurement using this framework, propose a series of smart heat measurement solutions determined by the Internet, and lay out the real-time assurance and administration implementation of the smart heat measurement platform.

The IoT networking layer is accountable for transmitting network information, including the access module component and the IoT accessible network component. Every IoT component has a terminal, and an IoT portal is dedicated to collecting data. IoT gateways are based on many circumstances, including industrial, household, and shared gateways. There are two categories of access networking components: wired and wireless. Its role is to safeguard data in the tangible realm and comprehensively grasp outside data. The perception stage comprises collaborative informational processing, information collection, and short-distance networking.

The networking setup involves assigning random values to the weight matrices of the outputting tier and then normalizing them to get $w(y = 1, 2, \dots, N)$. This process establishes the winning sector F(0) and the starting learning speed η value. m represents the total count of neurons in the result level. The relationship between learning speed and the total number of IoT devices is shown in Equation (1).

$$0.5\eta = \sqrt{F + \frac{w_x(k)}{n}} \tag{1}$$

The learning rate is η , the weight is denoted w_x , the number of IoT devices is denoted n, and the winning area is F(0). A randomized input sequence is taken from the training set and then standardized to get N($x = 1, 2, \dots, N$), wherein N is the total amount of neurons in the input tier. The learning rate is expressed in Equation (2).

$$\eta = F_y \left(P_n - w_{xy}(t) \right) \tag{2}$$

Determine the dot product of vectors P_n and w_{xy} and identify the winning component y with the highest dot product. To choose the winning component (w_{xy}) with the minor separation (P_n) , the algorithm should locate it based on the input method if it needs to be standardized. The separation length is expressed in Equation (3).

$$d_{y} = |P_{t} - W_{y}| = \sqrt{\sum_{y=0}^{N-1} (P_{x} - W_{y})^{2}}$$
(3)

The winning component is W_y , and the input is expressed P_x . Let y be the centroid used to calculate the range of weight adjustments at time t. Typically, the starting domain F(0) is extensive, and F(t) decreases as the training period progresses throughout the training procedure, as shown in Equation (4).

$$y \in F_y^*(t) \tag{4}$$

The starting field is expressed $F_y^*(t)$. Equation (5) modifies the weights of all elements in the winning area F(t).

$$w_{xy}(t+1) = w_{xy}(t) + k(p, N)$$
(5)

The previous winning area is denoted $w_{xy}(t)$, and the machine learning-based updating function is denoted k(p, N). Before establishing structural threat reduction, it is necessary to explain the following theory: If the dimension h is limited, the expression applies with a likelihood condition more excellent than or equal to $1 - \eta$, and the condition is expressed in Equation (6).

$$R(k) \le R_i(k) + \Delta k \tag{6}$$

The rank is denoted R(k), the present rating is expressed $R_i(k)$, and the weight deviation is Δk . All specimens in the linearly distinct training set satisfy a specific scenario, and the system is expressed in Equation (7).

$$Q_{\chi}(wp_{\chi}+b) - 1 \ge 0 \tag{7}$$

The training condition is denoted Q_x , machine learning-based weight is denoted w, the input is denoted p_x , and the bias condition is denoted b. The data-collecting plane consists of various sensors that gather data about the actual environment, including time and current state.

3.2. BDA and Intelligent Decision-Making for DBMS

Following data integration, it is crucial to focus on applying the information to increase its value and transform it into knowledge. The data acquired throughout the system, including equipment situation, failures, support integration, cabin surroundings, and crew wellness, is used to create machine learning libraries and test libraries of a specific size. Introducing a large data analytics engine would optimize categorization, clustering, decision trees, and AdaBoost techniques. Advanced techniques such as multi-dimensional mining assessment and artificial intelligence methods, including deep learning, reinforcement deep learning, and transfer learning, are used to explore the sophisticated data included in the information. As a result, the quality of intelligent offerings, such as support jobs, equipment failure forecasting, and cabin climate management, is enhanced to increase the combat efficacy of the entire warship and the efficiency of the crew's job. Several well-developed civil and economic domain methods are references for big data and smart assistance decision-making.

The information processing and preservation centers are designed to offer programs with robust computing features. This allows for a shift in design concentration towards tasks such as sample information collection, cleansing of data, and company procedure integration. The details regarding failure tracking, warship functioning support, and battle choices can be enhanced. The overall layout of this application is shown in Figure 2.

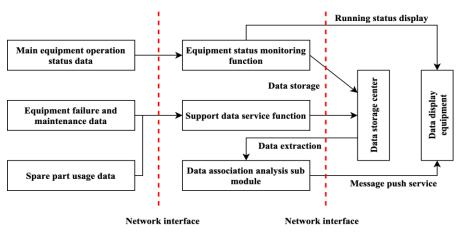


Figure 2: The overall layout of the ML-IDMS system

The operational status info regarding crucial equipment is provided via the data transmission tier to the equipment status observing function component located in a data processing center to track and analyze the details. The processed information comprises equipment identifying codes, machine operating variable data, and failure details conveniently saved in the data center. The support informational function component facilitates data entry, such as servicing reports and replacement component consumption. This inputted information is then held in the data collection facility. The data connection analytics sub-module inside the BDA and competent assistance decision-making functionality retrieve the data from the information storage center. It does data mining across many levels and parameters and examines the principles of association among the information.

Some instances of such rules include the association of problems among devices, the correlation of the issues between components within the device, the connection between variations in the environment and device mistakes, and the connection in the utilization of spare components. This approach establishes a foundation for decision-making in developing maintenance plans, efficiently allocating spare parts, and implementing condition-based management. It offers robust assistance in preserving the operational efficiency of machinery. The data processor center of the whole system provides computational aid to the methods. The outcomes of dataset connection evaluation are handled by the notification-pushing service component and sent to the selected information display gadget of the user interface tier via a data distribution tier to inform and display facts.

The methods and machine learning models are closely connected to individual applications' business information and business logic in BDA and machine learning programs. Several sub-modules are implemented in the BDA, and smart decision-making functions to fulfill the needs of different application situations.

ML-IDMS presents a smart heating metering system that utilizes the IoT to provide real-time data reading. This research shows a mathematical framework for initializing neural networks and a decision-making model for analyzing massive data. These developments jointly strive to improve system efficiency and tackle data management and analytics issues.

4 Simulation Analysis and Outcomes

The simulation was conducted using a high-performance computing cluster consisting of 64 nodes. Each node was equipped with dual Intel Xeon processors running at a frequency of 2.5 GHz, 128 GB of RAM, and NVIDIA Tesla V100 GPUs. The equipment requirement specified network switches with a minimum speed of 10 Gbps and a dispersed storage structure with a total storage capacity of 1 petabyte. The system necessitates Linux-based operating systems, namely Ubuntu 20.04 and Apache Hadoop Distributed File System (HDFS) version 3.3.1, to guarantee the most efficient performance for processing large volumes of data. The simulations were performed using a dataset of 10 terabytes, which accurately reflects real-world situations.

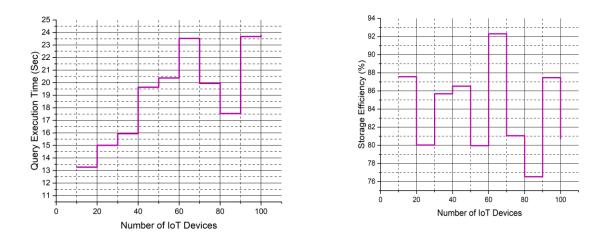


Figure 3(a): Query execution time analysis and 3(b) Storage efficiency analysis

Figure 3(a) illustrates the relationship between the query execution time and the number of IoT devices, demonstrating a consistent upward trend. The suggested ML-IDMS approach has an average query execution time of 19.27 seconds. The proposed method has a notable average efficiency of 83.78% in Figure 3(b), depicting storage efficiency. The suggested ML-IDMS technique performs better in both parameters, with an average query execution time lower than competing methods and continuously more excellent storage efficiency. This confirms its efficiency in effectively handling more enormous datasets while optimizing storage consumption.

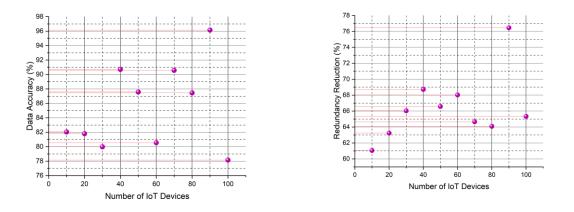


Figure 4(a): Data accuracy and 4(b) Redundancy reduction analysis

Figure 4(a) depicts data accuracy outcomes, demonstrating accuracy changes as the number of IoT devices varies. The suggested ML-IDMS approach consistently achieves a data accuracy of 90%, showcasing its strong performance under varying device loads. Figure 4(b) illustrates a redundancy decrease, demonstrating an average reduction of 66.42%. The suggested ML-IDMS strategy demonstrates exceptional data accuracy, guaranteeing dependable outcomes even with the proliferation of IoT devices. The effectiveness of this ML-IDMS system in minimizing redundancy strengthens its capacity to optimize data, hence improving the overall quality of the data.

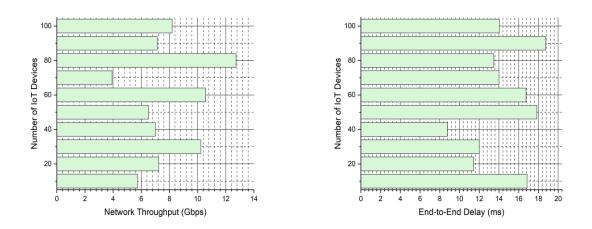


Figure 5: (a)Network throughput and 5(b) End-to-End delay analysis

Figure 5(a) displays the outcomes of network throughput, illustrating fluctuations in throughput with an increasing number of IoT devices. The suggested ML-IDMS technique consistently achieves an average network throughput of 7.93 Gbps, showcasing its capacity to manage growing workloads efficiently. Figure 5(b) displays the end-to-end delay, which shows variations in the duration of delays. The suggested ML-IDMS technique consistently maintains an average latency of 14.4 ms, demonstrating its high level of responsiveness and minimum uncertainty, significantly enhancing real-time data processing efficiency. Its persistent high throughput and minimal latency confirm the suggested method's effectiveness in controlling data traffic.

The suggested ML-IDMS technique exhibits impressive average performance across many metrics: query execution time (19.27 sec), storage efficiency (83.78%), data accuracy (90%), redundancy reduction (66.42%), network throughput (7.93 Gbps), and end-to-end delay (14.4 ms). The results validate the efficacy and proficiency of the suggested ML-IDMS approach, demonstrating its capacity to address various obstacles in data handling and analysis, thereby establishing it as a resilient solution for intricate situations.

5 Conclusion and Future Study

DBMS has seen significant changes, mainly due to the rise of Big Data and the widespread use of IoT technologies. The rapid increase in the complexity and amount of data highlights the difficulties encountered by traditional DBMS, requiring the development of creative alternatives. This research introduces a progressive strategy using Machine Learning-based Intelligent Database Management Systems (ML-IDMS) to address conventional DBMS's constraints and meet modern data ecosystems' requirements. The ML-IDMS platform offers advanced functionalities like real-time data retrieval, intelligent heat metering, and fast neural network initialization. These characteristics contribute to a robust ML-IDMS framework that enhances flexibility and decision-making capabilities. The simulation results demonstrate the effectiveness of ML-IDMS, with essential metrics such as query execution time (19.27 sec), storage efficiency (83.78%), data accuracy (90%), redundancy reduction (66.42%), network throughput (7.93 Gbps), and end-to-end delay (14.4 ms). The results confirm that ML-IDMS can address current data management difficulties.

Despite its achievements, ongoing difficulties must be addressed, including ensuring scalability, tackling security risks, and refining decision-making algorithms. The future potential rests in enhancing ML-IDMS to adapt to changing data environments, investigating innovative machine learning techniques, and strengthening security protocols to integrate seamlessly into more intricate information ecosystems. The progression continues towards more intelligent, adaptable, and robust data management systems, creating the foundation for revolutionary advancements in the sector.

References

- [1] Aldwairi, M., Jarrah, M., Mahasneh, N., & Al-khateeb, B. (2023). Graph-based data management system for efficient information storage, retrieval and processing. *Information Processing & Management*, 60(2), 103165. https://doi.org/10.1016/j.ipm.2022.103165
- [2] Bai, B., & Bai, Y. (2021). Fuzzy based decision making approach for big data research on health information management system. *Journal of Ambient Intelligence and Humanized Computing*, *12*, 3363-3371.
- [3] Bansal, B., Jenipher, V.N., Jain, R., Dilip, R., Kumbhkar, M., Pramanik, S., & Gupta, A. (2022). Big Data Architecture for Network Security. *Cyber Security and Network Security*, 233-267.
- [4] Bertello, A., Ferraris, A., Bresciani, S., & De Bernardi, P. (2021). Big data analytics (BDA) and degree of internationalization: the interplay between governance of BDA infrastructure and BDA capabilities. *Journal of Management and Governance*, *25*, 1035-1055.
- [5] Bhuiyan, M.N., Rahman, M.M., Billah, M.M., & Saha, D. (2021). Internet of Things (IoT): A review of its enabling technologies in healthcare applications, standards protocols, security, and market opportunities. *IEEE Internet of Things Journal*, *8*(13), 10474-10498.
- [6] De Almeida Pereira, G.H., Fusioka, A.M., Nassu, B.T., & Minetto, R. (2021). Active fire detection in Landsat-8 imagery: A large-scale dataset and a deep-learning study. *ISPRS Journal of Photogrammetry and Remote Sensing*, *178*, 171-186.
- [7] Erraji, A., Maizate, A., Ouzzif, M., & Batouta, Z.I. (2022). Migrating Data Semantic from Relational Database System To NOSQL Systems to Improve Data Quality for Big Data Analytics System. *ECS Transactions*, 107(1), 19495. https://doi.org/10.1149/10701.19495ecst
- [8] Feng, X., Ma, J., Liu, S., Miao, Y., Liu, X., & Choo, K.K.R. (2021). Transparent ciphertext retrieval system supporting the integration of encrypted heterogeneous database in cloudassisted IoT. *IEEE Internet of Things Journal*, 9(5), 3784-3798.
- [9] Huang, S., Qin, Y., Zhang, X., Tu, Y., Li, Z., & Cui, B. (2023). Survey on performance optimization for database systems. *Science China Information Sciences*, *66*(2), 121102.
- [10] Korherr, P., & Kanbach, D. (2023). Human-related capabilities in big data analytics: a taxonomy of human factors impacting firm performance. *Review of Managerial Science*, 17(6), 1943-1970.
- [11] Kwon, Y., Lee, S., Nam, Y., Na, J.C., Park, K., Cha, S. K., & Moon, B. (2023). DB+-tree: A new variant of B+-tree for main-memory database systems. *Information Systems*, 119, 102287. https://doi.org/10.1016/j.is.2023.102287.
- [12] Liang, W., Li, W., & Feng, L. (2021). Information security monitoring and management method based on big data in the internet of things environment. *IEEE Access*, *9*, 39798-39812.
- [13] Meera, S., & Sundar, C. (2021). A hybrid metaheuristic approach for efficient feature selection methods in big data. *Journal of Ambient Intelligence and Humanized Computing*, *12*, 3743-3751.
- [14] Melchiades, M.B., Crovato, C.D.P., Nedel, E., Schreiber, L.V., & Righi, R.D.R. (2021). Fast IoT: an efficient and very fast compression model for displaying a huge volume of IoT data in web environments. *International Journal of Grid and Utility Computing*, *12*(5-6), 605-617.

- [15] Mershad, K., & Hamieh, A. (2021). SDMS: smart database management system for accessing heterogeneous databases. *International Journal of Intelligent Information and Database Systems*, 14(2), 115-152.
- [16] Mohammad, A.S., & Pradhan, M.R. (2021). Machine learning with big data analytics for cloud security. *Computers & Electrical Engineering*, 96, 107527. https://doi.org/10.1016/j.compeleceng.2021.107527.
- [17] Nowakowski, P., Zórawski, P., Cabaj, K., & Mazurczyk, W. (2021). Detecting Network Covert Channels using Machine Learning, Data Mining and Hierarchical Organisation of Frequent Sets. *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications,* 12(1), 20-43.
- [18] Ravikumar, S., & Kavitha, D. (2021). A new adaptive hybrid mutation black widow clustering based data partitioning for big data analysis. *Wireless Personal Communications*, 120(2), 1313-1339.
- [19] Taipalus, T., Grahn, H., & Ghanbari, H. (2021). Error messages in relational database management systems: A comparison of effectiveness, usefulness, and user confidence. *Journal* of Systems and Software, 181, 111034. https://doi.org/10.1016/j.jss.2021.111034.
- [20] Truică, C.O., Apostol, E.S., Darmont, J., & Pedersen, T.B. (2021). The forgotten documentoriented database management systems: An overview and benchmark of native XML DODBMSes in comparison with JSON DODBMSes. *Big Data Research*, 25, 100205. https://doi.org/10.1016/j.bdr.2021.100205.
- [21] Zhang, H., Zang, Z., Zhu, H., Uddin, M.I., & Amin, M.A. (2022). Big data-assisted social media analytics for business model for business decision making system competitive analysis. *Information Processing & Management*, 59(1), 102762. https://doi.org/10.1016/j.ipm.2021.102762.

Authors Details



Rosa Clavijo-López, Doctor student in Applied Mathematical Statistics, Master in Public Management, Bachelor in Administration, research professor and leader of scientific research projects on issues of digital transformation, artificial intelligence and soft skills.



Dr. Wayky Alfredo Luy Navarrete, Doctor in Education Administration, Master in Administration, Bachelor in Economics, worked as an economist at the BCRP, departmental manager of Essalud in Tumbes, member of the board of EPS Grau Piura, Director of the Academic Department of Economics, advisor to private companies, currently research coordinator for thesis projects, responsible for Curricular Design, appointed professor at the School of Economics.

Integrating Novel Machine Learning for Big Data Analytics and IoT Technology in Intelligent Database Management Systems



Dr. Jesús Merino Velásquez, Doctor in Administration, Master in Economics with a mention in Business Management; Principal professor at the National University of Tumbes, he has developed scientific research, with publications in various international and national magazines: in the magazine of the National University of Tumbes; Mangrove in the area of corporate social responsibility and strategic planning and management. He currently teaches thesis subjects in the different Master's and doctoral programs at the UNTumbes Graduate School. He also teaches thesis subjects in the Undergraduate program of the aforementioned University.



Dr. Carlos Miguel Aguilar Saldaña, Doctor in Public Management and Governance, Chartered Public Accountant, Responsible for the Head of Formative Research and Scientific Integrity, Head of Research, Development and Technology Transfer, with collaborative or team work, problem solving and analytical skills.



Alcides Muñoz Ocas, Doctor in Business Management from the Universidad Nacional de San Martin, Master in Public Management from the Universidad Cesar Vallejo, Bachelor in Administration from the Universidad Nacional de San Martin, Undergraduate Professor at the School of Administration of the Universidad Nacional de San Martin, Vice-Dean of the College of Administration Graduates of the San Martin region.



Dr. César Augusto Flores Tananta, Doctor in public management and governance; Master in Business Administration - MBA; Certified Public Accountant by profession, RENACYT Research Professor, professor at the Faculty of Economics, Head of formative research and scientific integrity, Head of research, development, innovation and technology transfer.