International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

An In-Depth Review to Explore Cost Optimization Strategies for Healthcare Domain in Cloud Computing

Mayur Patel[1], Tejas Thakkar[2], Ami Trivedi[3], Palak Patel[4]

Submitted: 26/11/2023 **Revised:** 06/01/2024 **Accepted:** 16/01/2024

Abstract: Cloud Computing (CC) is a technological innovation that enables the provision of computing ability and data storage in a dynamic and flexible manner via pay-as-you-go services using the Internet. This technology has significantly advanced the field of Information Technology (IT). Over the recent years, the progression of cloud computing is increased to the emergence of novel technologies, including fog computing, edge computing, and cloud federation. However, the advent of the Internet of Things (IoT) has introduced various challenges associated with these innovative technologies. Hence, this manuscript delves into an examination of each of these evolving cloud-oriented technologies, encompassing their architectures, prospects, and challenges. The objective of this study is to assess the issue of cost optimization in healthcare (HC) by conducting a thorough survey of existing approaches in cloud computing. The paper aims to present a comprehensive classification of the aspects and parameters related to cost optimization in HC. Additionally, it offers a categorization of cost-based metrics, distinguishing between monetary and temporal cost parameters across various scheduling stages. The intention is to provide valuable insights for researchers and practitioners, aiding them in the selection of the most suitable cost optimization approach based on identified aspects and parameters. Furthermore, the paper outlines potential avenues for future research in this ongoing and evolving research domain.

***Keyword**: healthcare, categorization, Internet of Things, innovative, Information Technology*

1. Introduction:

Cloud computing has transformed the aspiration of scalable computational resources into a reality and is now poised to be integrated into various usage models. The IT sector acknowledges cloud computing as an emerging technology with wide-ranging applications across all disciplines. Cloud computing's primary functions include hosting and providing various software and services via the Internet [1][2]. It is an essential component of many companies' infrastructure. After more than ten years of development, cloud computing has not only proven to be incredibly successful but has also significantly changed how the economy, society, industries, and scientific community operate. The rapid advancements in mobile Internet and big data based technology have further solidified cloud computing's role, with a majority of online services and data services being established on its foundation. Consequently, cloud computing has found extensive applications in various business, marketing, as well as in the medical and research domains [3]. Cloud data centres represent extensive computer infrastructures designed to efficiently meet the needs of the IT industry and support multiple applications seamlessly. Nevertheless, various significant issues are linked to energy conservation across different domains of knowledge, exerting a profound influence on the economy. Two primary impacts stand out: eco-friendly and performance oriented. The environmental effects are both eco-friendly and inevitable, but the quality of work can be enhanced by developing diverse scenarios to address these concerns[4].

[1]*Computer Science, The CVM University, mayur.patel@cvmu.edu.in*
[2]*Computer Science, The CVM University, tejas.thakkar@cvmu.edu.in*
[3]*Computer Science, The CVM University, ami.trivedi@cvmu.edu.in*
[4]*Computer Science, The CVM University, palakben.patel@cvmu.edu.in*

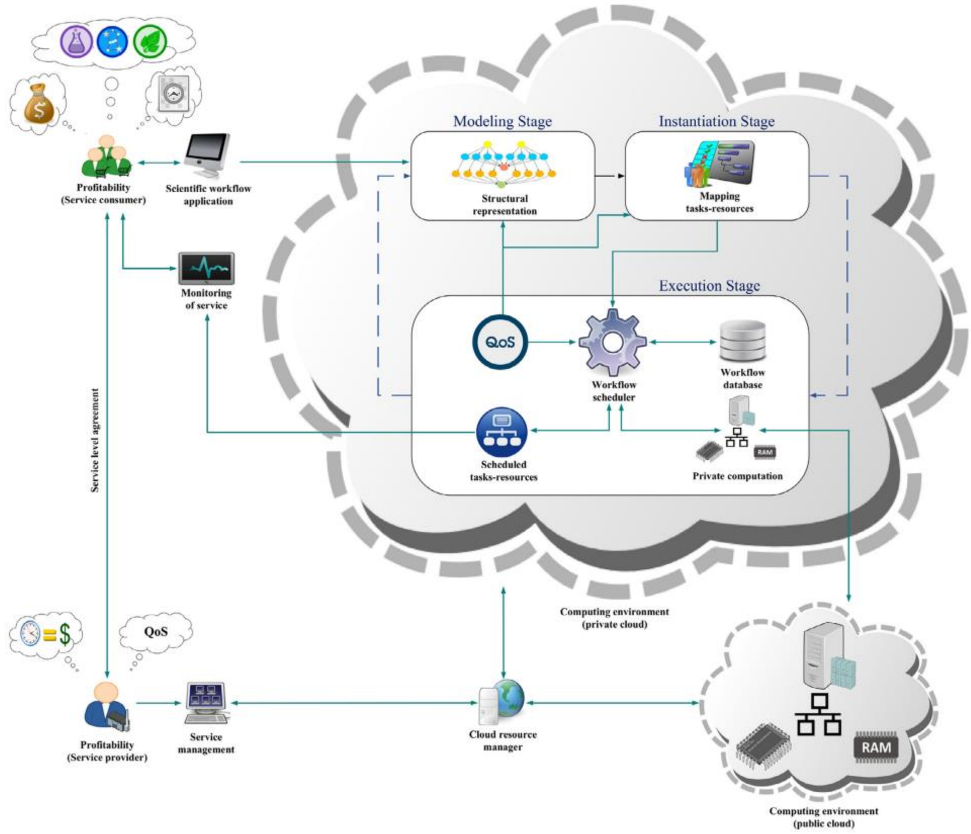

Fig 1: Cloud computing scalable computational model

Researchers now have more chances because of the move in the life sciences toward data-driven science, which makes it possible to generate large amounts of omics data quickly and affordably. Nevertheless, there are difficulties in the gathering, storing, sharing, evaluating, and understanding of this data [5]. Advances in a variety of Information Technology (IT) sectors are necessary due to the significant heterogeneity of data in terms of nature and source, raising problems regarding processing, computing power, privacy, security, storage, and sharing. As such, it is imperative to develop software tools and algorithms specifically designed for the analysis of various forms of omics data, including gene expressions, protein sequences, and single nucleotide polymorphisms (SNPs). These instruments are essential for learning about gene expression, control, and the DNA mutations causing hereditary illnesses. In addition, there is a need for the development of graphical user interfaces that can efficiently display data collected from several sources.

Healthcare organizations generate an extensive array of data and information. With the advancements in High-Throughput (HT) enabled technologies, there are possibilities of exponential surge in generated data, including gene terminologies, sequence alignments, and protein classifications. This renders traditional computational methods inadequate for managing the vast and heterogeneous datasets, transforming omics sciences into the realm of Big Data science. In the health and medical fields, the management of Big Data necessitates robust infrastructures to enhance data storage and handling [21].

Given the critical importance of data sharing and security in healthcare, seamless and secure access to extensive datasets is essential for scientific data investigation and result dissemination. The CC solutions tailored for healthcare sectors play a pivotal role in optimizing data investigation, distribution, access, and storage. Cloud based services, with their scalability capabilities, prove to be effective as they can adapt to the increasing volume of data. Consequently, CC services emerge as a cost-effective resolution for the efficient storage, access, analysis, sharing, and protection of healthcare data and information [22].

2. Cloud Computing in Health Care Domain:

The integration of cloud computing in the healthcare sector has witnessed substantial growth, particularly highlighted during the pandemic. This transformation has touched every facet of medical institutions, offering benefits such as heightened privacy, reduced costs, and improved patient care through remote accessibility and collaboration. Cloud computing has introduced novel avenues for enhancing the functionality of IT systems within healthcare[6]. The healthcare cloud computing market is projected to reach $71,730.64 million by 2027, indicating a remarkable acceptance of new technologies within the industry. This signals a willingness among healthcare organizations to invest significantly in technological advancements. Cloud computing in healthcare revolves around the use of remote servers accessible via the internet for storing, managing, and processing medical information securely. The adoption of cloud-based solutions, especially with the Electronic Medical Records (EMR) Mandate, has become prevalent among medical companies aiming to safeguard patient records. The shift to cloud solutions has proven instrumental in managing the massive amount of digital data generated by healthcare providers annually. Key benefits of incorporating cloud computing in healthcare include affordability in storing healthcare data, the growth of telemedicine, improved patient experiences through real-time access to medical information, enhanced collaboration among healthcare professionals, and convenient interoperability, enabling seamless data integration across the healthcare system. Different types of cloud computing in healthcare, categorized by deployment and distribution models, include S-a-a-S (Software as a Service), I-a-a-S (Infrastructure as a Service), P-a-a-S (Platform as a Service), and distribution models such as community, private, hybrid, and public. However, despite these advantages, challenges exist in the adoption of healthcare cloud computing. These challenges encompass the necessity for broader technology integration beyond cloud solutions, a shortage of expert developers in healthcare software, security concerns associated with storing sensitive healthcare data, and difficulties in transitioning from legacy frameworks to cloud technologies, requiring comprehensive task management process transformations and training [7] [23].

3. Advantages of Cloud in Healthcare Sector:

Key benefits of incorporating cloud computing in healthcare include affordability in storing healthcare data, the growth of telemedicine, improved patient experiences through real-time access to medical information, enhanced collaboration among healthcare professionals, and convenient interoperability, enabling seamless data integration across the healthcare system. Different types of cloud computing in healthcare, categorized by deployment and distribution models, include SaaS (Software as a Service), IaaS (Infrastructure as a Service), PaaS (Platform as a Service), and deployment models such as community, private, hybrid, and public. However, despite these advantages, challenges exist in the adoption of healthcare cloud computing. These challenges encompass the necessity for broader technology integration beyond cloud solutions, a shortage of expert developers in healthcare software, security concerns associated with storing sensitive healthcare data, and difficulties in transitioning from legacy frameworks to cloud technologies, requiring comprehensive task management process transformations and training [8] [24].

Cost Efficiency:

The CC offers on-demand resources, reducing the need for costly in-house resources. Cloud vendors deliver agility and ample storage dimensions at a reasonable cost. Technical tasks, upgrades, and maintenance are handled by providers, leading to significant operational cost savings for healthcare organizations [25].

Unparalleled Scalability:

Cloud systems provide unmatched scalability, allowing healthcare institutions to adjust data storage and network requirements based on their specific needs. This flexibility is crucial in dynamic healthcare environments, enabling organizations to scale up during peak periods and scale down during calmer periods, optimizing expenses. This agility contrasts with the inflexibility of on-premise systems, where acquiring new equipment involves significant investments [26].

Enhanced Patient Experience:

Cloud computing in healthcare elevates medical services, enhancing patient satisfaction. Adopting the cloud allows healthcare organizations to offer personalization, transparency, and convenience. With telemedicine powered by the cloud, patients can consult healthcare specialists remotely, receiving medical services at home. Cloud-driven healthcare apps, patient portals, and wearables further improve and personalize medical care, empowering patients with increased access to their health data [27].

Data-Driven Decision Making:

Cloud computing enables healthcare organizations to leverage big data analytics, AI, and machine learning. This technology processes large amounts of raw data quickly, providing valuable perceptions for data-driven decision-making. It enhances the eminence of medical services and helps organizations meet their business purposes by streamlining processes, allocating resources efficiently, and managing facilities effectively [28].

Improved Interoperability and Collaboration:

Cloud computing creates a connected environment for healthcare organizations, centralizing and organizing data. This facilitates easy access for medical staff, allowing timely medicines, treatment strategies, and efficient data-linked tasks. Real-time updates enable seamless tracking of medical records across specialists, reducing cross-department communication and fostering collaboration among healthcare professionals, even across geographical limits.

4. Risks of Cloud Computing in Medicine:

Security Concerns:

Security is a primary concern in cloud adoption, particularly for sensitive healthcare data. Data breaches can be financially costly and damage a healthcare organization's reputation. Reliable cloud providers implement multi-level privacy measures, meticulous user access monitoring, data encoding, two-step authentication, and blockchain to protect data from outside threats.

Acquiescence with Data-Regulatory Ethics:

Healthcare organizations essential comply with data-regulatory standards such as HIPAA, HITECH, and GDPR. Selecting a reputable cloud provider with robust security measures helps ensure agreement with these critical ethics and allows for adaptability to changes in guidelines.

System Downtimes:

Despite the reliability of cloud services, concerns persist about potential data and system control loss during provider downtimes. Implementing contingency plans, reducing dependence on a single provider, or adopting hybrid cloud solutions are strategies to mitigate risks and enhance resilience in medical cloud computing.

5. Related Work Analysis:

The integration of cloud computing in the healthcare sector has witnessed substantial growth, particularly highlighted during the pandemic. This transformation has touched every facet of medical institutions, offering benefits such as heightened privacy, reduced costs, and improved patient care through remote accessibility and collaboration. The adoption of cloud-based solutions, especially with the Electronic Medical Records (EMR) Mandate, has become prevalent among medical companies aiming to safeguard patient records. The shift to cloud solutions has proven instrumental in managing the massive amount of digital data generated by healthcare providers annually.

In [9], authors have Involving both Fog and Cloud computing, the focus is on price and value of service in terms of reply time. The approach employs Cost Optimisation based on Task Deadline (COTD) to efficiently optimize energy usage by reducing task loads. However, it faces challenges in understanding the complex cloud environment, resulting in higher energy consumption.

In [10], authors have focused on concentrating on Cloud computing, the issue addressed is Load Balancing (LB). The proposed technique combines Harries Hawks Optimization (HHO), and Ant Colony Optimization (ACO) to achieve less reversal time, and LB time. The approach is specific to Infrastructure as a Service (IaaS) and requires implementation in a real-world scenario.

In [11], authors have focused onDealing with various computing resources in the Cloud, the challenge is optimizing both cost and makespan for high-performance applications. Despite being a complex problem for such applications, the technique aims to reduce time and economic costs, albeit at a higher cost to meet deadlines.

In [12], authors have focused on Fog computing, the objective is cost optimization for elastic computing resources on demand. The approach employs a gradient-based optimizer (GBO) to enhance constraint satisfiability and cost optimization. However, it faces drawbacks when compared to traditional workflow scheduling algorithms, leading to increased energy consumption.

In [13], authors have focused on Cloud computing, the emphasis is on developing a load balancer algorithm. The Lion optimizer is introduced to balance loads, demonstrating unsettled performance in terms of extreme turnaround time. Nevertheless, it incurs higher costs for meeting deadlines of larger workflows.

In [14], the field of fog computing, a method intended to mitigate load balancing, processing, bandwidth, and storage issues, has caught the authors' attention. Performance measures including bandwidth, distance, and least cost path are optimized through the use of the salp swarm algorithm (SSA) and the Arithmetic Optimization Algorithm (AOA). It's crucial to remember that this method can only be used with hybrid clouds.

In [15], authors have focused on Cloud computing, the goal is effective resource management for dynamic workloads. Deep neural network (DNN) enabled schedulers are utilized to bring about enhancements in implementation costs, energy utilization, and service level contract defilements, offering a temporary saving of operational costs. The technique allows for modeling dependencies among tasks.

In [16], authors have focused on addressing scheduling, resource allocation, and load balancing in the Cloud, the technique involves Ant Colony Optimisation (ACO). It is only relevant to reserved instances and attempts to improve service composition, energy usage, and replication; nonetheless, difficulties still exist in cloud settings.

In [17], To enhance the performance of Internet of Things applications, the authors have focused their research on workload modelling, resource provisioning, and workload scheduling in the Cloud domain. The strategy depends on computing and storage resources that are available on demand, therefore it might not be appropriate for every user.

In [18], authors have focused to precisely and proficiently predict cloud jobs, the technique employs an efficient supervised learning-based DNN. This results in a significant reduction in mean square errors, utilizing the Gated Recurrent Unit (GRU) for precise prediction. However, it may not be suitable for handling complete workflow structures.

In [19], the authors focused on CC, with a particular focus on issues with host or server overload and underload. The suggested method makes use of a Q-learning algorithm in combination with the Artificial Bee Colony Algorithm (ABC). The main goals are to improve throughput and average while concurrently reducing makespan, cost, and imbalance degree.. Nevertheless, system performance optimization across all test datasets is challenging, especially in multi-cloud, fog cloud, or edge cloud environments.

In [20], authors have focused on Addressing the overload problem, the technique involves a copy allocation algorithm. It effectively reduces data transmission time and storage overhead but falls short in focusing on resource management for osmotic computing. Similar to other techniques, it allows for modeling dependencies among tasks.

Table I: Comparative analysis previous work done

References		**Problem Statement**	**Technique**	**Advantage**	**Disadvantage**	**Findings**
[9]	Fog and Cloud	considerations of cost and quality of service are evaluated with a focus on response time.	Cloud Cost Optimization based on Task Deadline (COTD)	Efficient energy optimization techniques by reducing task loads.	abortive to understand cloud in complex situation	More Energy consumption
[10]	Cloud	Load Balancing (LB)	HHO and ACO are crossbred in the proposed method.	Low turnaround and LB time	Infrastructure as a Service (IaaS),	Need an implementation on real scenario
[11]	Cloud	Computing resources with various combinations of configurations and prices.	Achieving optimization for both cost and time in the operation of high-performance computing applications within a cloud computing environment remains a challenging and intricate problem.	remains a complex problem for high-performance computing applications	Reduce the lifetime and economic costs	It Takes More Cost to meet the deadlines
[12]	Fog	Cost optimization on elastic computing resources on demand.	gradient-based optimizer (GBO)	Improved interms of constraint satisfiability and cost optimization	Traditional workflow scheduling algorithms	More Energy consumption

[13]	Cloud	load balancing technique for CC	The Lion Optimizer is designed to distribute and balance loads effectively.	Enhanced performance is achieved in relation to the maximum turnaround time.	Task scheduling probability	It Takes More Cost to meet the deadlines of larger workflows.
[14]	Fog computing	Fog computing diminishes the requirements for load balancing, processing, bandwidth, and data storage.	Arithmetic Optimization Algorithm, and the Salp swarm algorithm	performance measures as distance, bandwidth and least cost path,	task scheduling probability	Applicable only for hybrid clouds.
[15]	Cloud	effective resource management for dynamic workloads	DNN based schedulers	improvement s in execution costs, energy utilization and SLA violations	temporarily save operational costs.	Dependencies among tasks can be modeled
[16]	Cloud	scheduling, resource allocation, load balancing	Ant Colony Optimisation (ACO)	Improved ser vice composition, energy consumption, and replication	Several challenges exist in cloud environments such as Security Concerns:	Applicable for only reserved instances
[17]	Cloud	workload modeling, resource provisioning, workload scheduling		improve the performance of IoT applications	on-demand storage and computation facilities	Not suitable for all users
[18]	Cloud	accurately and efficiently predict cloud workloads.	supervised learning based DNN	.reduce the mean square errors significantly	Gated Recurrent Unit (GRU) to attain precise forecast.	Not suitable for complete workflow structure
[19]	Cloud	underloaded, which may affect the processing time or may result in a system crash.	Artificial Bee Colony Algorithm using a Q-based learning technique	reducing makespan, reducing cost, reducing degree of imbalance, increasing throughput and average	The optimization of system performance may not be achievable across every test dataset.	Scheduling tasks in a multi-cloud, fog cloud, or edge cloud environment poses significant challenges.

				resource utilization.		
[20]		tackling the overload issue due to its inadequate capacity.	replica allocation technique	effectively decrease the data broadcast time and storage overhead cost	Does not focused on supply supervision for the CC	Dependencies among tasks can be modeled

6. Conclusion:

In conclusion, the research emphasizes the importance of fluctuations in critical factors such power demand, electricity costs, and renewable power generation, with a particular emphasis on their impact on a data centre's total power consumption cost. To tackle these issues, an intelligent power management system is proposed that makes use of a strong energy cost optimization approaches. The proposed algorithm orchestrates various elements such as data centre load, battery storage, power generators, renewable energy sources, and electrical energy prices from both actual-time power markets. This real-time coordination aims to mitigate the expected energy consumption cost by making informed decisions based on the dynamic nature of the parameters involved. To address uncertainties associated with data centre load and renewable energy, forecasting algorithms are employed to compute these parameters. The paper proposed a model for calculating prices in service level agreements, ensuring that clients engaging in on-demand cloud services are billed in a manner that reflects the operational cost of the data centre. Simulation results conducted on real data centre load, climate conditions, and electrical energy prices validate the effectiveness of the proposed methodology. The outcomes demonstrate that the smart power management system, with its robust optimization algorithm, proves to be a powerful and efficient tool for minimizing the operational cost of a data centre. This research contributes to the ongoing efforts in advancing sustainable and cost-effective practices in the management of data centre resources.

References

[1] Channi, H. K., Sandhu, R., Faiz, M., & Islam, S. M. (2023, August). Multi-Criteria Decision-Making Approach for Laptop Selection: A Case Study. In 2023 3rd Asian Conference on Innovation in Technology (ASIANCON) (pp. 1-5). IEEE.

[2] Faiz, M., & Daniel, A. K. (2023). A hybrid WSN based two-stage model for data collection and forecasting water consumption in metropolitan areas. International Journal of Nanotechnology, 20(5-10), 851-879.

[3] Narayan, V., Faiz, M., Mall, P. K., & Srivastava, S. (2023). A Comprehensive Review of Various Approach for Medical Image Segmentation and Disease Prediction. Wireless Personal Communications, 132(3), 1819-1848.

[4] Saxena, A., Chauhan, R., Chauhan, D., Sharma, S., Sharma, D., & Narayan, V. (2022). Comparative Analysis Of AI Regression And Classification Models For Predicting House Damages İn Nepal: Proposed Architectures And Techniques. Journal of Pharmaceutical Negative Results, 6203-6215.

[5] Narayan, V., Awasthi, S., Fatima, N., Faiz, M., Bordoloi, D., Sandhu, R., & Srivastava, S. (2023, May). Severity of Lumpy Disease detection based on Deep Learning Technique. In 2023 International Conference on Disruptive Technologies (ICDT) (pp. 507-512). IEEE.

[6] Mall, P. K., Narayan, V., Srivastava, S., Sabarwal, M., Kumar, V., Awasthi, S., & Tyagi, L. (2023). Rank Based Two Stage Semi-Supervised Deep Learning Model for X-Ray Images Classification.

[7] Chaturvedi, P., Daniel, A. K., & Narayan, V. (2021). Coverage Prediction for Target Coverage in WSN Using Machine Learning Approaches.

[8] Narayan, V., Daniel, A. K., & Chaturvedi, P. (2023). E-FEERP: Enhanced Fuzzy based Energy Efficient Routing Protocol for Wireless Sensor Network. *Wireless Personal Communications*, 1-28.

[9] Danthuluri, S., & Chitnis, S. (2023). Energy and cost optimization mechanism for workflow scheduling in the cloud. *Materials Today: Proceedings*, *80*, 3069-3074.

[10] Kamanga, C. T., Bugingo, E., Badibanga, S. N., & Mukendi, E. M. (2023). A multi-criteria decision making heuristic for workflow scheduling in cloud computing environment. *The Journal of Supercomputing*, *79*(1), 243-264.

[11] Godhrawala, H., & Sridaran, R. (2023). A dynamic Stackelberg game based multi-objective approach for effective resource allocation in cloud computing. *International Journal of Information Technology*, *15*(2), 803-818.

[12] Malathi, K., & Priyadarsini, K. (2023). Hybrid lion–GA optimization algorithm-based task scheduling approach in cloud computing. *Applied Nanoscience*, *13*(3), 2601-2610.

[13] Mohamed, A. A., Abdellatif, A. D., Alburaikan, A., Khalifa, H. A. E. W., Elaziz, M. A., Abualigah, L., & AbdelMouty, A. M. (2023). A novel hybrid arithmetic optimization algorithm and salp swarm algorithm for data placement in cloud computing. *Soft Computing*, *27*(9), 5769-5780.

[14] Tuli, S., Casale, G., & Jennings, N. R. (2023). Learning to Dynamically Select Cost Optimal Schedulers in Cloud Computing Environments. *ACM SIGMETRICS Performance Evaluation Review*, *50*(4), 29-31.

[15] Asghari, S., & Jafari Navimipour, N. (2023). The role of an ant colony optimisation algorithm in solving the major issues of the cloud computing. *Journal of Experimental & Theoretical Artificial Intelligence*, *35*(6), 755-790.

[16] Jeyaraj, R., Balasubramaniam, A., MA, A. K., Guizani, N., & Paul, A. (2023). Resource management in cloud and cloud-influenced technologies for internet of things applications. *ACM Computing Surveys*, *55*(12), 1-37.

[17] Xu, M., Song, C., Wu, H., Gill, S. S., Ye, K., & Xu, C. (2022). esDNN: deep neural network based multivariate workload prediction in cloud computing environments. *ACM Transactions on Internet Technology (TOIT)*, *22*(3), 1-24.

[18] Kruekaew, B., & Kimpan, W. (2022). Multi-objective task scheduling optimization for load balancing in cloud computing environment using hybrid artificial bee colony algorithm with reinforcement learning. *IEEE Access*, *10*, 17803-17818.

[19] Li, C., Bai, J., Chen, Y., & Luo, Y. (2020). Resource and replica management strategy for optimizing financial cost and user experience in edge cloud computing system. Information Sciences, 516, 33-55.

[20] Faiz, M., & Daniel, A. K. (2022). A Multi-Criteria Dual Membership Cloud Selection Model based on Fuzzy Logic for QoS. International Journal of Computing and Digital Systems, 12(1), 453-467.

[21] Narayan, Vipul, et al. "7 Extracting business methodology: using artificial intelligence-based method." Semantic Intelligent Computing and Applications 16 (2023): 123

[22] Narayan, Vipul, et al. "A Comprehensive Review of Various Approach for Medical Image Segmentation and Disease Prediction." Wireless Personal Communications 132.3 (2023): 1819-1848.

[23] Mall, Pawan Kumar, et al. "Rank Based Two Stage Semi-Supervised Deep Learning Model for X-Ray Images Classification: AN APPROACH TOWARD TAGGING UNLABELED MEDICAL DATASET." Journal of Scientific & Industrial Research (JSIR) 82.08 (2023): 818-830.

[24] Narayan, Vipul, et al. "Severity of Lumpy Disease detection based on Deep Learning Technique." 2023 International Conference on Disruptive Technologies (ICDT). IEEE, 2023.

[25] Saxena, Aditya, et al. "Comparative Analysis Of AI Regression And Classification Models For Predicting House Damages İn Nepal: Proposed Architectures And Techniques." Journal of Pharmaceutical Negative Results (2022): 6203-6215.

[26] Kumar, Vaibhav, et al. "A Machine Learning Approach For Predicting Onset And Progression"“Towards Early Detection Of Chronic Diseases “." Journal of Pharmaceutical Negative Results (2022): 6195-6202.

[27] Chaturvedi, Pooja, Ajai Kumar Daniel, and Vipul Narayan. "Coverage Prediction for Target Coverage in WSN Using Machine Learning Approaches." (2021).

[28] Chaturvedi, Pooja, A. K. Daniel, and Vipul Narayan. "A Novel Heuristic for Maximizing Lifetime of Target Coverage in Wireless Sensor Networks." Advanced Wireless Communication and Sensor Networks. Chapman and Hall/CRC 227-242.