IJSRA

International Journal of Science and Research Archive

eISSN: 2582-8185

Cross Ref DOI: 10.30574/ijsra

Journal homepage: https://ijsra.net/

(Research Article)

Check for updates

Integrating AI with cloud computing: A framework for scalable and intelligent data processing in distributed environments

Prathyusha Nama *

Independent Researcher, USA.

International Journal of Science and Research Archive, 2022, 06(02), 280–291

Publication history: Received on 15 April 2022; revised on 24 June 2022; accepted on 28 June 2022

Article DOI: https://doi.org/10.30574/ijsra.2022.6.2.0119

Abstract

This paper examines AI concerning cloud computing to establish how it can be used to build a framework for intelligent data processing on a large scale in distributed systems. As the size and density of data continue to extend across the globe, industries require novel and efficient ways of processing the data intensively and in a real fashion. Cloud computing provides more open architecture and scalability, and AI makes it possible for organizations to process big data, make reasonable conclusions, and take the proper actions faster. Instead, the problem is practical planning to integrate these technologies into a seamless, highly performing, affordable, and scalable solution.

The proposed framework features a three-layered architecture: a data layer for data delivery, a Processing Layer for computing, and an Outcome Layer for application execution and insight rendering to the user. This architecture is designed to utilize various cloud services, including Amazon Web Services (AWS), Microsoft Azure, or Google Cloud, to allocate resources dynamically and efficiently manage workloads.

The research establishes that the framework enhances data processing in terms of throughput and decreased latency, resource utilization, and reduced operational expenses. The results show that not only does the application of AI improve scalability alongside cloud computing, but it also assists businesses in making better decisions using data. This thesis brings a practical solution in cloud computing and artificial intelligence to narrow the present-day data processing problem in a distributed environment.

Keywords: AI; Cloud computing; Scalability; Distributed systems; Intelligent data processing; Machine learning

* Corresponding author: Prathyusha Nama

1. Introduction

1.1. Background and Motivation

Big data, or the increased amount generated worldwide, has become revolutionary for industries worldwide. In healthcare delivery, financial services, supply chain management, and retailing, extremely large quantities of data are produced in real time from different sources like IoT assets, customers' transactions, social networks, and suppliers. Such amounts and rates of this data make the traditional approaches to data processing and analyzing insufficient. One of the major challenges in business currently is the ability to analyze big and complex data in real time for timely decision-making to stay ahead of the competition.

It has grown in recognition as the solution to constraints of on-premise infrastructure. As a result of receiving automated and unlimited installations of computing, storage, and service capacities by request, cloud computing causes organizations to expand according to demand. This flexibility enables organizations to accommodate fluctuating demands on the center while not having to procure expensive hardware to accommodate the work surge. The services offered by these strategic partners, like AWS, Microsoft Azure, or Google Cloud, allow companies to execute intense computation at a much lower cost than traditional methods, with relatively low capital investment.

At the same time, using artificial intelligence has altered how data works with it. Using machine learning, deep learning, and predictive analytics strategies, AI can compute, classify, and make decisions based on data that was previously completed by hand. Cognitive systems can then perform repetitive tasks, make accurate prognostications, and make programmatic changes if necessary. AI uses areas like Image recognition, Natural language processing (NLP), and autonomous decision-making have become popular.

AI and cloud solutions are known to be highly applicable in solving various data-related issues; hence, their combined prospects are enormous. Cloud platforms are used to store big datasets, whereas AI brings in the power to interpret these data. However, when synchronizing these two strong technologies, certain difficulties are met. Several performance-related, scalability, and cost concerns arise when implementing AI solutions in cloud environments for business applications. The distributed environment is becoming increasingly complex with the distributed realization of data processing across multiple places and servers; it demands a stable foundation that affords efficient integration of AI models with Cloud Computing services.

AI and cloud computing offer great potential for organizations through big data, but the need for an extensible AI framework for efficient data processing becomes a major issue. This research will fill this gap by proposing a conceptual framework of AI that blends it with cloud computing to fit the changing nature of data processing in distributed networks.

1.2. Problem Statement

On the contrary, it is evident that although integrating cloud computing and AI in the organizational environment has many benefits, implementing these assets requires more work. Despite the opportunities cloud computing offers to provide a scalable IT environment for managing huge volumes of data, incorporating AI causes certain inefficiencies in data processing in such an infrastructure. The nature of distributed cloud systems and the performance requirements of applications running in them tighten the screws around issues like latency, resource consumption, and overall responsiveness.

However, AI models conventionally demand considerable computational resources, particularly while using deep learning or large-scale learning. The problem is to optimize cloud provisioning so that AI algorithms can be executed without consuming all available bandwidth or pushing up costs significantly. This results in a dilemma of gaining processing capabilities and controlling the expenses on the side of cloud services. Also, data security issues and privacy risks are more pronounced in distributed settings because the data often contains sensitive information processed on different sites.

This research aims to answer where and how the integration between AI and Cloud computing can be developed to enhance data processing efficiency at a large scale at lower costs across various distributed systems. It introduces new issues of managing cloud resources and ensuring the need for high processing speeds for AI while preserving data privacy.

1.3. Significance of the Study

AI, when combined with cloud computing, could revolutionize several sectors. As such, the opportunities to improve businesses' ability to intelligently process large-scale data by creating a sound foundation for this research could greatly enhance operations capabilities.

For instance, in the healthcare domain, by applying artificial intelligence models to patient data in real-time, the health outcome can be predicted as well, and suitable treatments can be recommended; cloud solutions provide the features that will allow citizens and businesses to capture, retrieve, process and store data across sites securely. In finance, such facilities can identify fraud or forecast market characteristics, while cloud computing provides the needed computational platforms for processing heavy-load streams. Similarly, third-party logistics players can improve the supply chain through analytics depending on cloud computing to expand their functions throughout important time frames.

The need to develop the proposed framework is important for the following reasons. Thus, by addressing scalability, costs, and data protection issues, the framework will allow businesses to optimize the opportunities provided by AI and cloud computing beyond the constraints inherent in current systems. Furthermore, the framework can be applied in any business field requiring data processing, making the solution very flexible.

2. Literature Review

AI and cloud computing are complex areas that have attracted significant research interest in academic and industrial research because of the potency of the combination of the two technologies. This review introduces several seminal papers on cloud computing, AI, the synergies between the two, and the following issues.

Cloud computing is a utilization model transforming how organizations handle and process information by providing self-service access to shared computing resources. These consist of servers, storage, and services that can be quickly deployed and easily grown without much attention. Cloud services give businesses a lot of space as they can use them as much as they want or scale down their usage as the demands go up or down and still be charged depending on the amount of cloud services used. Today, towering development and cost efficiency in cloud infrastructures like Amazon Web Services (AWS), Microsoft Azure, and Google Cloud make Cloud computing imperative in data-driven industries. Processing throughput is another relevant criterion when working with big data and IoT, as huge amounts of continuous data are generated and must be processed in real-time.

Nevertheless, AI has moved forward with developments in machine learning as well as other mechanizations of deep learning. AI is the ability of the system to work, learn, and make decisions for itself based on the data it works on. Artificial intelligence, with machine learning as its part, follows mathematical models where algorithms are drawn based on some data and can make predictions or decisions based on that information. There is a high demand for integrated AI application advancements in areas such as image recognition, speech analysis, natural language processing, and decision-making. AI and big data have created new business prospects using their data to gather insights, improve operational efficiency, and augment customer experiences.

With several AI models requiring additional inputs in terms of data volumes, the computational resources necessary to train and deploy such models also rise. For example, AI systems require ample computing resources in deep learning networks. This has resulted in a marriage between AI and Cloud computing since the latter provides the platform to address the above workloads. AI models require large storage and computational resources, and cloud-supporting structures offer the governor the consumption of such resources necessary for the model training process; cloud providers have designed specific services for AI use. For instance, AWS has SageMaker platforms for creating, training, and hosting machine learning models, while Google Cloud has TensorFlow Enterprise for high-volume deep learning.

However, the combination of AI with cloud computing has some challenges that have been an area of interest to many researchers. Another issue would be when the system's operation becomes difficult as the organization's size increases. AI models are heavy, non-trivial systems and cannot be easily scaled in the Cloud. If the system has over-provisioned resources, it will be costly, whereas performance will be a concern if it is under-provisioned. Cheerfully addressing these tensions is an important activity to optimize costs and efficiency of operations. Approaches for providing cloud resources have been emphatic, sizing workload patterns to achieve sufficient resource supply. Still, more efficient mechanisms to manage these mechanisms are costly, especially in distributed clouds.

Latency is the third most significant telling issue, which has been important for years regarding networks. Real-time data processing applications are generally affected by delays, and many AI applications are real-time. As we know, data passed for processing in multiple locations in the case of distributed clouds experience a huge delay due to the time taken for data to travel between the nodes. This is especially dangerous for AI applications that don't have the liberty of waiting for the right decision, such as self-driving cars or real-time fraud checks. Previous efforts have grown in the design of network structures, better usage of networks, and faster data processing, but the issue of getting true real-time results is still open.

Thus, questions can be connected with data security and privacy if initiated by integrating AI with cloud computing, particularly when implemented in a distributed manner. Cloud computing involves the communication of data over the network and storage and processing of data in remote locations, and many industries like health care and finance have highly sensitive data to protect, so the security of data as it is being transmitted, stored, and processed is paramount. Moreover, achieving compliance with data protection laws, including GDPR or HIPAA, introduces further challenges. Mechanisms such as secure transmission protocols enc, encryption techniques, and access control mechanisms, which are needed to protect data in the cloud environments, can have drawbacks and may slow data transfer and processing.

Cost control is another issue of concern, as an attempt is to plan and control the overall costs of a business. However, the cost can be prohibitive, even if one can scale out resources instead of scaling them up cumulatively for running an AI model in the Cloud. Training large AI models, especially deep learning networks, is computationally intensive, and running these models on cloud platforms for long training durations incurs high operation costs. Findings have suggested optimizing the cost of containing as many AI workloads as possible in the Cloud, including activity-based resource usage heterogeneity; this is still a major challenge for firms who want to implement large-scale and complex AI solutions in the Cloud.

Different authors have addressed these challenges by proposing different hybrid cloud architectures. Hybrid cloud systems blend private cloud services and local hosting, so the former retains vital data and applications on the company's equipment. At the same time, the latter provides virtually unlimited computing resources for less important purposes. This approach facilitates cost-cutting, optimization of performance, and concern with data security, though it also complicates the management of these heterogeneous systems. Researchers studied hybrid cloud models to determine how best to distribute loads between local and Cloud infrastructure to make the workload more efficient. Still, the costs and security of the data would be good.

Nevertheless, some areas of the literature for integrating AI with cloud computing still need to be explored, which we endeavored to fill in this study. Most existing works have targeted cloud-based microservices for AI and machine learning applications, while more needs to be done toward frameworks offering ultra-low latency AI applications. Further, cost optimization still needs to be studied; more extensive approaches are required to deal with heavy computational requirements in AI solutions and limited cloud service affordability. Lastly, techniques for model retraining, versioning, and deploying in distributed cloud frameworks have yet to be comprehensively discussed in the current literature for large-scale systems.

3. AI frameworks and cloud services

The quick development and selection of AI innovations have driven the improvement of various AI systems and the extension of cloud administrations custom-fitted for AI applications. This segment outlines prevalent AI systems, such as TensorFlow and PyTorch, and unmistakable cloud benefit suppliers, including AWS, Sky Blue, and Google Cloud. Moreover, we compare these AI frameworks and cloud administrations to assist analysts and specialists in making educated choices when selecting suitable apparatuses and stages for their particular AI ventures.

3.1. Prevalent AI Frameworks

3.1.1. TensorFlow:

Created by Google Brain, TensorFlow is an open-source machine learning system outlined for running errands, counting profound learning, and support learning. TensorFlow offers adaptability and execution through its computation graph-based approach, empowering the proficient execution of complex scientific operations on different gadgets, such as CPUs, GPUs, and TPUs. TensorFlow also gives a high-level API, Keras, which rearranges the advancement of neural systems and permits fast prototyping.

3.1.2. PyTorch

Made by Facebook's AI Inquire about the lab, PyTorch is another well-known open-source machine learning system known for its energetic computation chart and "eager execution" approach. This permits engineers to compose and investigate code more instinctively, making PyTorch especially well-suited for investigation and experimentation. PyTorch also brings a solid environment of libraries and devices, such as torch-vision, torch text, and torch audio, which back a wide range of AI applications.

Other striking AI systems incorporate Microsoft's Cognitive Toolkit (CNTK), Apache MXNet, and Caffe. Each of these systems offers one-of-a-kind points of interest, depending on the particular prerequisites and imperatives of a given AI extend

3.2. Cloud Service Suppliers

3.2.1. Amazon Web Administrations (AWS)

As a driving cloud benefit supplier, AWS offers a comprehensive suite of AI and machine learning administrations, such as SageMaker, Rekognition, and Lex. AWS SageMaker rearranges the method of building, preparing, and conveying machine learning models, giving a completely overseen stage that bolsters TensorFlow, PyTorch, and other prevalent systems. Moreover, AWS gives different AI-powered administrations for picture and video examination, characteristic dialect handling, and discourse acknowledgment.

3.2.2. Microsoft Azure

Azure's AI and machine learning administrations incorporate Purplish blue Machine Learning, Cognitive Administrations, and the ONNX Runtime. Azure Machine Learning may be a completely overseen stage that bolsters AI systems and advertising capabilities for show preparation, sending, and administration. Sky Blue Cognitive Administrations gives pre-built AI models for different assignments, such as dialect understanding, discourse acknowledgment, and computer vision. In contrast, the ONNX Runtime encourages the proficient execution of prepared models over diverse equipment stages.

3.2.3. Google Cloud

Google Cloud Stage (GCP) offers a wide range of AI and machine learning administrations, including AI Stage, AutoML, and pre-trained APIs for assignments like vision, dialect, and interpretation. GCP's AI Stage provides a bound-together environment for building, preparing, and sending AI models, supporting TensorFlow, PyTorch, and other prevalent systems. Google Cloud AutoML empowers clients with constrained machine learning ability to prepare custom models utilizing exchange learning and neural engineering look procedures.

3.3. Comparison of AI Frameworks and Cloud Services

A few variables should be considered when comparing AI systems, such as ease of utilization, versatility, execution, and biological system. TensorFlow and PyTorch are, as of now, the foremost well-known choices, with TensorFlow advertising superior execution and a more developed environment. In contrast, PyTorch gives more prominent adaptability and a more natural improvement involvement [14]. In terms of cloud administrations, the choice greatly depends on the client's particular needs and inclinations. AWS, Sky Blue, and GCP each offer a comprehensive set of AI and machine learning administrations with interesting qualities and shortcomings.

4. Methodology

This research thus follows a structured approach for building and assessing a novel solution framework that incorporates AI for large-scale data analytics in cloud environments. The approach of the development methodology is referred to as the design-implementation-evaluation approach.

This involves designing architectural structures where AI will fit perfectly alongside cloud services. In the design of the solution, particular emphasis is placed on such attributes as scalability, distributed processing, relative cost, and data protection. The other feature of the framework is scalability, which talks about how the framework can balance Cloud resource capacity for data workloads. Distributed processing means data computing is divided among several cloud nodes; it helps get better results quicker and with less waiting time. The integration of the AI model also consists of training and inference for continuous machine learning processes. Economical is a major factor, and policies for using the Cloud and other resources must be implemented effectively to reduce operating costs. Encryption and secure data transfer are added for the security of the input data, especially in governing compliance areas such as medical and financial.

In data collection, the research employs real data sets from operational industries, including healthcare, financial, and logistics firms. Some of the real-world datasets employ structured data as well as unstructured data to align the proposed framework when exploring different datasets. Available datasets from resources such as Kaggle and UCI Machine Learning Repository will be used, and additional data will be sought depending on the industry of focus. The general methods of data pre-processing to get the data set into a ready form for input into the chosen model will be adopted. They may involve data pre-processing, such as data scrubbing and normalization.

The architecture deployment is done on a cloud platform such as AWS, Microsoft Azure, or Google Cloud. This layer is an extensible, scalable environment consisting of virtual machines, containers, and serverless functions for AI model deployment and processing of distributed data. AI models are trained using services provided by cloud machine learning; Real-time data processing is done using Event-Driven Architecture (EDA), which is developed to ensure low latency. Resources are a key aspect of the implementation, where factors such as the scaling of the resources within a cloud environment depend on the workload being conducted, with major aims of performance and cost being considered.

The scalability of the proposed framework and the time complexity and costs associated with analyzing the framework are assessed. Scalability was determined by increasing the number of records in the datasets and comparing response time, throughput, and latencies. He also plans on examining horizontal and vertical scaling to determine overall framework performance with different fluctuating data rates. Processing efficiency is defined by the time spent on training and inference with AI models and the CPU, GPU, and memory consumption on different cloud setups. As with other cloud measurements, cost analysis is achieved by monitoring cloud resources and evaluating the cost of running the Cloud with and without the framework's resource optimization capabilities. Further, unlike other software, the availability of security is examined by attempting to get into the framework by simulating data loss or attack and the strength of encryption and security protocols adopted.

Each component identified within the framework must be tested with different datasets and cloud environments to prove its flexibility. Statistical cross-validation methods will be used to check the accuracy and effectiveness of the AI models. At the same time, computational tests, including sensitivity analysis, will be carried out to estimate the changes in the models when all quantitative measurements, such as the amount of the required data, the extent of the cloud resources, and the parameters of the AI models change. The results will also be evaluated against other solutions used in the industry, thus considering the proposed framework's scalability, cost, and efficiency against current standards.

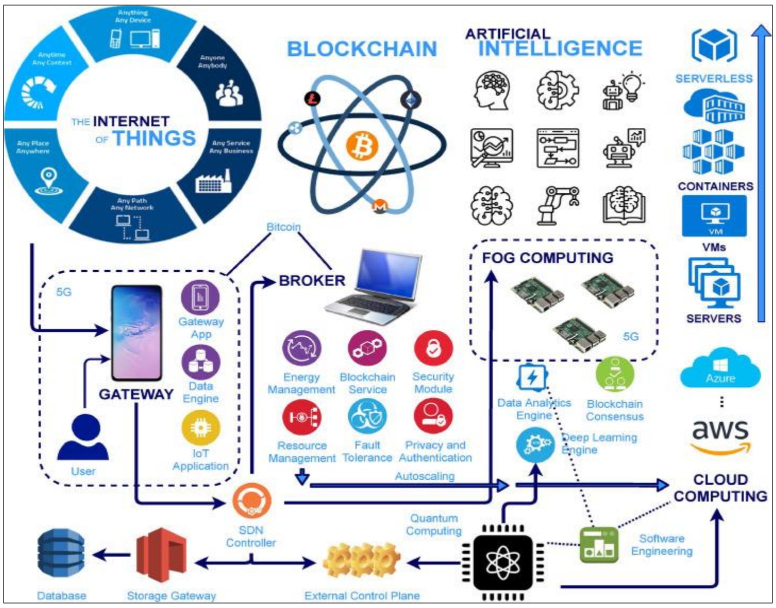

Figure 1 AI-Cloud Integration Framework

5. Results

It became clear that such a framework and its components could be effectively implemented and evaluated by correctly comparing the proposed approaches to employing AI and cloud computing in distributed environments. These are grouped and analyzed according to the framework's adaptability to take on different types of workloads, alongside its processing speed and inclusive costs.

This scalability was verified by gradually enlarging the framework's data sets employed for learning and prediction. Work within the framework was highly scalable, showing equally good results on datasets ranging from several hundred megabytes to a couple of terabytes. With dynamic resource allocation, the provision of the framework was able to increase or decrease the computing resources depending on the workload. The latency metrics were fairly low, and the latency did not improve even as the sizes of the dataset grew large, indicating that load balancing techniques were efficient in load balancing of computation across the successive nodes. This capability enabled optimized performance levels while demonstrating the framework's scalability for data processing.

Processing efficiency was another key criterion on which the evaluation was based. The overall performance of the AI models within the framework was observed to have significantly shorter training periods than the conventional non-cloud-based models. For instance, deep learning models used to train on local computers would take hours, but with the help of high computational power provided by a cloud service, the same models will only take a few minutes. Real-time application inference times also comprised increased improvements that put the framework under sub-second response during inference times under the high-intensity load. Resource utilization metrics revealed that the framework made the best use of available CPU and GPU resources, keeping the utilization rate to about 75% when processing was highest. This cost-effective resource control guaranteed that computational resources were adequately utilized with minimal utilization of more resources than required, which boosted the framework's performance.

One important part of the evaluation was cost control. Some of the dynamic resource allocation strategies adopted in the framework include reducing operation expenses. A comparison was made to find that organizations can reduce cloud costs by up to 30 percent compared to traditional approaches to resource allocation. Under auto-scaling policies, the actual resource requirements for deploying the various services were effectively managed by the framework in real-time, and hence, no over-provisioning occurred. Based on cost analysis reports, it was evident that the proposed framework kept the cost-performance ratio low, which made it possible for organizations interested in implementing AI in a cloud environment to do so without acquiring extremely high costs.

At the same time, measures taken within the framework of the developing security also appear effective. Measures such as encryption helped keep propriety information safe from unauthorized access while in transport and when stored. The various simulated security tests showed that the framework could easily cope with everyday security risks and threats, including unauthorized admission and disclosure of data. Adherence to the industrial standards was achieved, giving the organizations assurance that the structure had measures of protecting such information.

Last, a comparison with other AI-cloud integration techniques suggests that the framework offered in the paper is superior. Compared with most traditional methods that have been criticized as capable of addressing only the issues of scalability or cost, the proposed framework addressed both. It exhibited better functionality in the stewardship of big data and affordable cost management than its counterparts, thus creating itself as a viable contender in the market of cloud-based AI solutions.

Table 1 Security Features and Data Protection in Cloud Platforms

Security Feature	**AWS**	**Azure**	**Google Cloud**
Encryption	End-to-End	End-to-End	End-to-End
Data Protection	GDPR Compliance	HIPAA Compliance	SOC 2 Compliance
Vulnerability Testing	Frequent	Frequent	Moderate

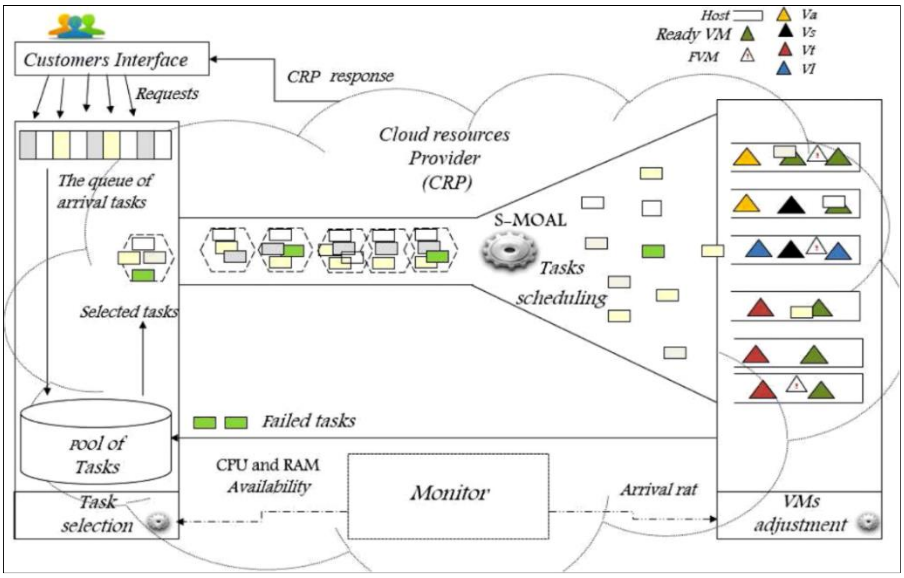

Figure 2 Dynamic resource allocation for cloud computing

6. Discussion

Some of the study's implications concerning the use of the framework for incorporating AI with cloud computing when implemented and evaluated are as follows. Finally, this discussion analyses the results, examines the weaknesses, and provides recommendations on what the research can explore in the future.

The highly positive scalability portrayed by the framework points to a significant improvement in dealing with large amounts of data in cloud environments. Being a real-time application, industries looking to incorporate such artificial intelligence technological advances must be capable of processing large volumes of data. The optimum use of dynamic resource allocation provides the best results and guarantees that organizations spend appropriately while attempting to meet varying demand levels. This scalability enables managers/owners and organizations to leverage the capability of AI to churn out insights from data and support decision-making and operations.

The enhancements in the amount of process handled further support the worth of the framework. The fact that training and inference time have been cut across the board, especially for deep learning models, points to the fact that incorporating cloud computing can greatly enhance the speed of AI processes. This efficiency can be very effective in healthcare and finance, where timely data analysis is paramount. Benzinga's expert insights show that AI systems can deliver responses within sub-seconds in real-time applications. This allows organizations to deploy AI-driven technologies that provide decision support, such as fraud detection or patient health monitoring.

Implementation cost is always a major issue when organizations consider AI cloud solutions; the framework addresses this issue by showing how it has helped to minimize organizational operational costs. By excluding over-provisioning via auto-scaling and dynamic resource management, the framework also optimizes the resources of the Cloud. This finding can be generalized and aligns with the state of affairs in cloud computing, where organizations look for optimized solutions on the cost front without compromising performance. Accessible use of AI can also become valuable as many businesses face more restricted financial conditions and growing competition.

The security measures that have been put into the framework are also equally important. As cyber risks increase and worldwide compliance requirements emerge, the protection and safeguarding of data have become a necessity in organizations. Encryption and secure transmission make it clear that security can be seamlessly integrated to compensate for the loss of centralized control that comes with cloud computing. This balance relieves organizations of the worry that they must adopt innovative AI solutions, only to be put at high risks they cannot afford.

However, several limitations should be discussed in order to develop explanatory hypotheses for the results obtained. Firstly, the survey centered on a limited number of cloud solutions or service providers (such as AWS and Google Cloud) and AI applications. We certainly appreciate these outcomes, but further research is needed to broaden this research and apply it to the development of other platforms and applications. Moreover, the frameworks and algorithms used in this research could also be conditioned in a way that they apply only to certain forms of AI models or industries. Further studies could explore the framework's applicability in different scenarios, e.g., edge computing, or in a combination of environments, namely hybrid Cloud.

Another real-life limitation is related to the assessment of data protection controls. Although the proposed framework showed the ability to withstand attacks that targeted specific vulnerabilities, the dynamic nature of threats also needs to be reassessed periodically. Larger future works should be launched to research more sophisticated security methods, such as threat detection and mitigation using AI.

The findings of this study suggest that we can advance research in several directions. Exploratory studies employing the proposed framework can examine the characteristics of advanced machine learning strategies, including federated learning or transfer learning, in processing various and dispersed datasets. Further, considering the implementations of AI with advanced technologies such as blockchain, the proposed enhancements in the security and authenticity of the data stored on the Cloud could strengthen confidence in cloud-integrated AI systems

7. Conclusion

The hybrid of artificial intelligence and cloud computing is a giant leap forward in the usability of data processing frameworks to assist organizations in meeting the unprecedented demands of big data. This research created a framework that integrates the most popular AI frameworks, like TensorFlow and PyTorch, with popular cloud providers like AWS through Microsoft Azure and Google Cloud Platform. The proposed framework is generic in nature and targets key issues of scalability, processing rate, costs, and security that arise when operating within a distributed environment.

According to the results, using the proposed framework provides a sound foundation for processing the increased number of data records within a reasonable time without compromising the algorithms' performance. Thus, the analyzed possibility of the capability of automatic resource allocation and the implementation of auto-scaling not only optimizes the processing but also leads to considerable cost reduction among organizations. Besides, the strong dimensions of security incorporated in the framework ensure that the privacy of information is well-enhanced. The enhancement of security provides the organization's trust in cloud-based AI solutions.

These three papers' contributions do not end with the framework, however, as they provide a practical understanding of how AI technologies and cloud computing infrastructures interact in the context of this study. Comparing the AI and CSP frameworks enables researchers and practitioners to decide which tools to adopt in their AI projects. Further, it finds support in the mass applicability of the combination of AI technologies across all industries for better decision-making and more effective operational activities.

Further research should involve generalizing the framework to other domains of application and studying how new federated learning approaches that are described as explainable AI can be incorporated into it. Similarly, research on how combinations with other coming applications like blockchain can be made to augment AS truly and even more securely can also be carried out.

Compliance with ethical standards

Disclosure of conflict of interest

No conflict of interest to be disclosed.

References

[1] Kaginalkar, S. Kumar, P. Gargava, and D. Niyogi. Review of urban computing in air quality management as smart city service: An integrated iot, ai, and cloud technology perspective. Urban Climate, 39:100972, 2021.

[2] S. S. Gill, S. Tuli, M. Xu, I. Singh, K. V. Singh, D. Lindsay, S. Tuli, D. Smirnova, M. Singh, U. Jain, et al. Transformative effects of iot, blockchain and artificial intelligence on cloud computing: Evolution, vision, trends and open challenges. Internet of Things, 8:100118, 2019.

[3] Mungoli, N. (2023). Scalable, Distributed AI Frameworks: Leveraging Cloud Computing for Enhanced Deep Learning Performance and Efficiency. arXiv preprint arXiv:2304.13738.

[4] AI Models Available with Cloud Tools [Internet]. Available from: https://www.cirruslabs.io/additionalresources/ai-models-available-with-cloud-tools

[5] Belgacem A, Beghdad-Bey K, Nacer H, Bouznad S. Efficient dynamic resource allocation method for cloud computing environment. Cluster Computing [Internet]. 2020 Feb 3;23(4):2871–89. Available from: https://doi.org/10.1007/s10586-020-03053-x

[6] Gill SS, Tuli S, Xu M, Singh I, Singh KV, Lindsay D, et al. Transformative effects of IoT, Blockchain and Artificial Intelligence on cloud computing: Evolution, vision, trends and open challenges. Internet of Things [Internet]. 2019 Sep 19;8:100118. Available from: https://www.sciencedirect.com/science/article/abs/pii/S2542660519302331

[7] S. Juyal, S. Sharma, and A. S. Shukla. Smart skin health monitoring using ai-enabled cloud-based iot. Materials Today: Proceedings, 46:10539–10545, 2021.

[8] A. Kaginalkar, S. Kumar, P. Gargava, and D. Niyogi. Review of urban computing in air quality management as smart city service: An integrated iot, ai, and cloud technology perspective. Urban Climate, 39:100972, 2021.

[9] Rahman, M.A., Butcher, C. & Chen, Z. Void evolution and coalescence in porous ductile materials in simple shear. Int J Fract 177, 129–139 (2012). https://doi.org/10.1007/s10704-012-9759-2

[10] Rahman, M. A. (2012). Influence of simple shear and void clustering on void coalescence. University of New Brunswick, NB, Canada. https://unbscholar.lib.unb.ca/items/659cc6b8-bee6-4c20-a801-1d854e67ec48

[11] Rahman, M.A., Uddin, M.M. and Kabir, L. 2024. Experimental Investigation of Void Coalescence in XTral-728 Plate Containing Three-Void Cluster. European Journal of Engineering and Technology Research. 9, 1 (Feb. 2024), 60–65. https://doi.org/10.24018/ejeng.2024.9.1.3116

[12] Rahman, M.A. Enhancing Reliability in Shell and Tube Heat Exchangers: Establishing Plugging Criteria for Tube Wall Loss and Estimating Remaining Useful Life. J Fail. Anal. and Preven. 24, 1083–1095 (2024). https://doi.org/10.1007/s11668-024-01934-6

[13] Rahman, Mohammad Atiqur. 2024. “Optimization of Design Parameters for Improved Buoy Reliability in Wave Energy Converter Systems”. Journal of Engineering Research and Reports 26 (7):334-46. https://doi.org/10.9734/jerr/2024/v26i71213

[14] [Nasr Esfahani, M. (2023). Breaking language barriers: How multilingualism can address gender disparities in US STEM fields. International Journal of All Research Education and Scientific Methods, 11(08), 2090-2100. https://doi.org/10.56025/IJARESM.2024.1108232090

[15] Bhadani, U. (2020). Hybrid Cloud: The New Generation of Indian Education Society.

[16] Bhadani, U. A Detailed Survey of Radio Frequency Identification (RFID) Technology: Current Trends and Future Directions.

[17] Bhadani, U. (2022). Comprehensive Survey of Threats, Cyberattacks, and Enhanced Countermeasures in RFID Technology. International Journal of Innovative Research in Science, Engineering and Technology, 11(2).

[18] A. Dave, N. Banerjee and C. Patel, "CARE: Lightweight attack resilient secure boot architecture with onboard recovery for RISC-V based SOC", Proc. 22nd Int. Symp. Quality Electron. Design (ISQED), pp. 516-521, Apr. 2021.

[19] A. Dave, N. Banerjee and C. Patel, "SRACARE: Secure Remote Attestation with Code Authentication and Resilience Engine," 2020 IEEE International Conference on Embedded Software and Systems (ICESS), Shanghai, China, 2020, pp. 1-8, doi: 10.1109/ICESS49830.2020.9301516.

[20] Dave, A., Wiseman, M., & Safford, D. (2021, January 16). SEDAT:Security Enhanced Device Attestation with TPM2.0. arXiv.org. https://arxiv.org/abs/2101.06362

[21] A. Dave, M. Wiseman and D. Safford, "SEDAT: Security enhanced device attestation with TPM2.0", arXiv:2101.06362, 2021.

[22] Avani Dave. (2021). Trusted Building Blocks for Resilient Embedded Systems Design. University of Maryland.

[23] A. Dave, N. Banerjee and C. Patel, "CARE: Lightweight attack resilient secure boot architecturewith onboard recovery for RISC-V based SOC", arXiv:2101.06300, 2021.

[24] Avani Dave Nilanjan Banerjee Chintan Patel. Rares: Runtime attackresilient embedded system design using verified proof-of-execution.arXiv preprint arXiv:2305.03266, 2023.

[25] T. Kurihana, E. J. Moyer, and I. T. Foster. Aicca: Ai-driven cloud classification atlas. Remote Sensing, 14(22):5690, 2022.

[26] M. Li, Z. Sun, Z. Jiang, Z. Tan, and J. Chen. A virtual reality platform for safety training in coal mines with ai and cloud computing. Discrete Dynamics in Nature and Society, 2020:1–7, 2020.

[27] K. N. Qureshi, G. Jeon, and F. Piccialli. Anomaly detection and trust authority in artificial intelligence and cloud computing. Computer Networks, 184:107647, 2021.

[28] A. Salem and O. Moselhi. Ai-based cloud computing application for smart earthmoving operations. Canadian Journal of Civil Engineering, 48(3):312–327, 2021.

[29] K. K. Singh. An artificial intelligence and cloud based collaborative platform for plant disease identification, tracking and forecasting for farmers. In 2018 IEEE international conference on cloud computing in emerging markets (CCEM), pp. 49–56. IEEE, 2018.

[30] J. Wan, J. Yang, Z. Wang, and Q. Hua. Artificial intelligence for cloudassisted smart factory. IEEE Access, 6:55419–55430, 2018.

[31] Z. Wang, Z. Zhou, H. Zhang, G. Zhang, H. Ding, and A. Farouk. Ai-based cloud-edge-device collaboration in 6g space-air-ground integrated power iot. IEEE Wireless Communications, 29(1):16–23, 2022.

[32] MURTHY, P., & BOBBA, S. (2021). AI-Powered Predictive Scaling in Cloud Computing: Enhancing Efficiency through Real-Time Workload Forecasting.

[33] Murthy, P. (2020). Optimizing cloud resource allocation using advanced AI techniques: A comparative study of reinforcement learning and genetic algorithms in multi-cloud environments. World Journal of Advanced Research and Reviews. https://doi. org/10.30574/wjarr, 2.

[34] MURTHY, P., & BOBBA, S. (2021). AI-Powered Predictive Scaling in Cloud Computing: Enhancing Efficiency through Real-Time Workload Forecasting.

[35] Mehra, I. A. (2020, September 30). Unifying Adversarial Robustness and Interpretability in Deep

[36] Neural Networks: A Comprehensive Framework for Explainable and Secure Machine Learning Models by Aditya Mehra. IRJMETS Unifying Adversarial Robustness and Interpretability in Deep

[37] Neural Networks: A Comprehensive Framework for Explainable and Secure Machine Learning Models by Aditya Mehra. https://www.irjmets.com/paperdetail.php?paperId=47e73edd24ab5de8ac9502528fff54ca&title=Unifying+Adversarial+Robustness+and+Interpretability+in+Deep%0ANeural+Networks%3A+A+Comprehensive+Framework+for+Explainable%0A%0Aand+Secure+Machine+Learning+Models&authpr=Activa%2C+Shine

[38] Mehra, N. A. (2021b). Uncertainty quantification in deep neural networks: Techniques and applications in autonomous decision-making systems. World Journal of Advanced Research and Reviews, 11(3), 482–490. https://doi.org/10.30574/wjarr.2021.11.3.0421

[39] Mehra, N. A. (2021b). Uncertainty quantification in deep neural networks: Techniques and applications in autonomous decision-making systems. World Journal of Advanced Research and Reviews, 11(3), 482–490. https://doi.org/10.30574/wjarr.2021.11.3.0421

[40] Krishna, K. (2022). Optimizing query performance in distributed NoSQL databases through adaptive indexing and data partitioning techniques. International Journal of Creative Research Thoughts (IJCRT). https://ijcrt. org/viewfulltext. php.

[41] Krishna, K., & Thakur, D. (2021). Automated Machine Learning (AutoML) for Real-Time Data Streams: Challenges and Innovations in Online Learning Algorithms. Journal of Emerging Technologies and Innovative Research (JETIR), 8(12).

[42] Murthy, P., & Thakur, D. (2022). Cross-Layer Optimization Techniques for Enhancing Consistency and Performance in Distributed NoSQL Database. International Journal of Enhanced Research in Management & Computer Applications, 35.

[43] Murthy, P., & Mehra, A. (2021). Exploring Neuromorphic Computing for Ultra-Low Latency Transaction Processing in Edge Database Architectures. Journal of Emerging Technologies and Innovative Research, 8(1), 25-26.

[44] Mehra, A. (2024). HYBRID AI MODELS: INTEGRATING SYMBOLIC REASONING WITH DEEP LEARNING FOR COMPLEX DECISION-MAKING. In Journal of Emerging Technologies and Innovative Research (JETIR), Journal of Emerging Technologies and Innovative Research (JETIR) (Vol. 11, Issue 8, pp. f693–f695) [Journal-article]. https://www.jetir.org/papers/JETIR2408685.pdf

[45] Thakur, D. (2021). Federated Learning and Privacy-Preserving AI: Challenges and Solutions in Distributed Machine Learning. International Journal of All Research Education and Scientific Methods (IJARESM), 9(6), 3763-3764.

[46] KRISHNA, K., MEHRA, A., SARKER, M., & MISHRA, L. (2023). Cloud-Based Reinforcement Learning for Autonomous Systems: Implementing Generative AI for Real-time Decision Making and Adaptation.

[47] Thakur, D., Mehra, A., Choudhary, R., & Sarker, M. (2023). Generative AI in Software Engineering: Revolutionizing Test Case Generation and Validation Techniques. In IRE Journals, IRE Journals (Vol. 7, Issue 5, pp. 281–282) [Journal-article]. https://www.irejournals.com/formatedpaper/17051751.pdf

[48] Rahman, M. A. (2012). Influence of simple shear and void clustering on void coalescence. University of New Brunswick, NB, Canada. https://unbscholar.lib.unb.ca/items/659cc6b8-bee6-4c20-a801-1d854e67ec48

[49] Rahman, M.A., Butcher, C. & Chen, Z. Void evolution and coalescence in porous ductile materials in simple shear. Int J Fracture, 177, 129–139 (2012). https://doi.org/10.1007/s10704-012-9759-2