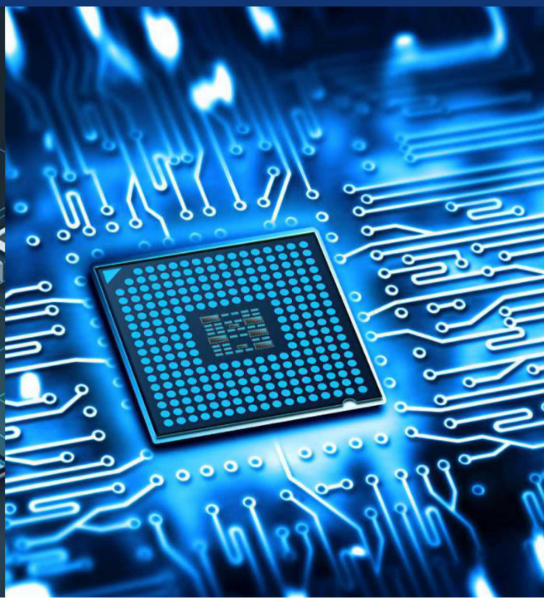
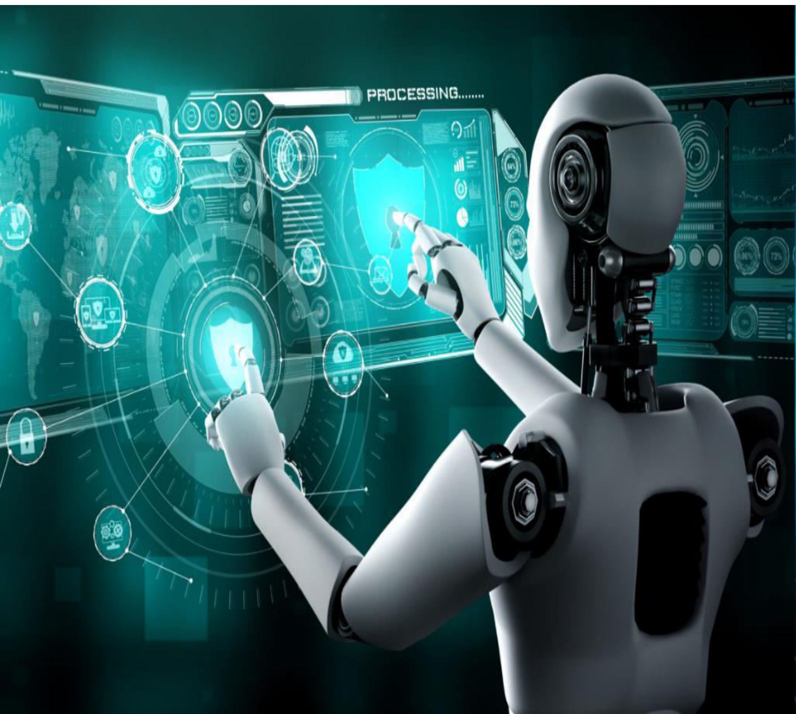
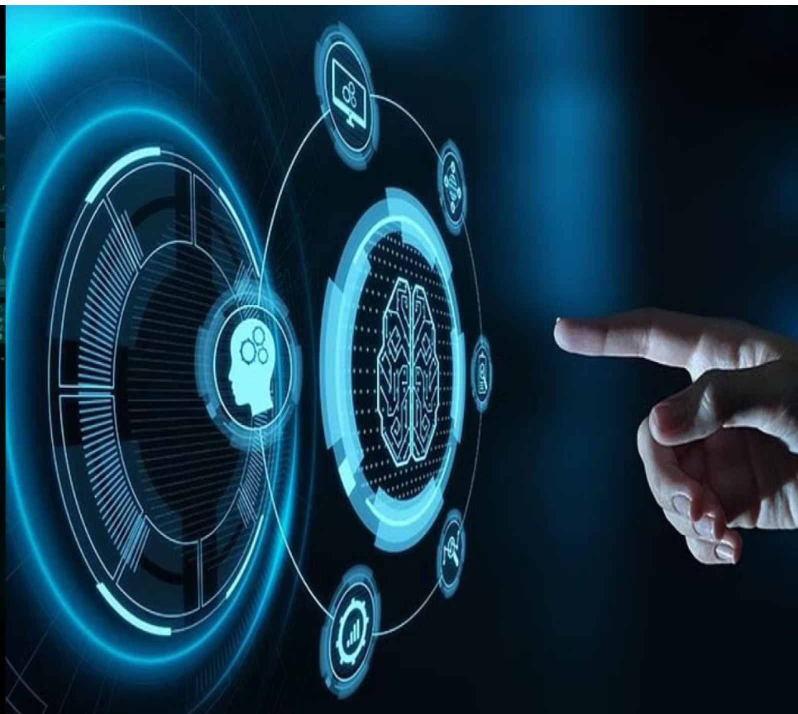
ISSN(O): 2320-9801
ISSN(P): 2320-9798


International Journal of Innovative Research in Computer and Communication Engineering

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)





Impact Factor: 8.771

Volume 13, Issue 5, May 2025

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Automating Repetitive Tasks in Cloud-Based AI Systems: A Deep Learning Perspective

Navya Krishna Alapati

VISA USA, INC

ABSTRACT: Today, AI within a cloud-based system is an integral part of the business process as it has been proven to be able to help businesses with significant improvements in its methodology. Nevertheless, the implementation and management of AI systems can be complicated as well as time intensive for this would need to execute repeated tasks that might lead to human errors. Consequently, as organizations deploy and develop AI systems in the cloud there is an increasing pressure to come up with ways automation can be used instead. This extended abstract describes how deep learning techniques are being used in cloud-based AI systems to reduce friction and automate repetitive tasks. Neural networks come under deep learning which is a set of machine-learning methods that aims to perform both logic-based computational activities. We suggest a model to automate the deployment, tuning, and monitoring of AI systems in the cloud based on deep learning algorithms. Our framework uses historical data and constant learning to adjust automation processes automatically, minimizing any hand-holding required from manual corrective measures leading toward a higher system efficiency. In addition, we reflect on the challenges and potential pitfalls in deploying a framework as well as indicate some possible future research lines. In summary, the new methodology can significantly improve and scale out cloud-based AI systems since a lower number of repetitive tasks would be required from human operators.

KEYWORDS: Human Error, Automating, Human Operators, Deployment,Complex Systems

I. INTRODUCTION

The practice of automating such repeated tasks within cloud-based AI systems is known as Cloud Automation using artificial intelligence technology. This means applying algorithms and ML models to data to process it, decide what should be done with it, or take action without the explicit involvement of a human[1]. Data is collected through all the data sources and placed centrally, say in a cloud-based AI system like Data Warehouse or Data Lake. The AI algorithm will process the big data, to find out what was different in instances with higher FPRs than other WHY LESS IMGS races. They help in decision-making or fuelling real-time actions[2]. The most resounding advantage of automating monotonous tasks in the cloud AI systems is the means of manual labor. By using AI technology to work continually, in the back end of business operations employees can concentrate on complex and strategic tasks leaving many functions that pose risk to human error removed – driving higher levels of accuracy & efficiency[3]. One other benefit, that is quite obvious given it's a cloud-based AI system — but never the less an important piece missing in legacy solutions with on-premises software and rolling hardware updates. Since these systems are cloud-hosted, they can readily process great volumes of data and demanding processing tasks making them suitable for companies that frequently change their business requirements. Furthermore, some of these systems adjust themselves over time through learning from new data (doing so with automation), consequently increasing their ability to recognize medically relevant patterns[4]. But automating these tasks in cloud-based AI systems is technically pretty tricky. For example, the model is only as good as its input which means machine learning systems can succumb to poor-quality or inconsistent data. Hence, It is important to handle and clean the data properly for these systems[5]. Security and privacy of data in these systems is yet another challenge. Because they work with subject data, it is important to have good security and data access controls. This post concludes that the automation of routine activities is generally able to greatly enhance such capacity for cloud-contingent AI implementations [6]. Nonetheless, addressing their contextual or technical challenges and the need for constant monitoring to maintain these systems is essential for them to be useful. Automation of repetitive tasks, especially construction activities needs to be adopted for cloud-based AI systems[7]. One example of automated tasks is the computer software or algorithms that execute processes that would be done by humans; this is basically what made Humans live. This approach has created a plethora of benefits but at the same time, it gave rise to various challenges that need to be resolved for successful adoption. The main bottleneck in automating repetitive tasks, in the context of AI systems that are deployed on a cloud, lies with clean data[8]. Large datasets are needed for training AI models to increase their performance. This can be problematic to access and consolidate if your

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

data resides on numerous sources in a cloud-based system. Validating the data inputs contributes [Equation] to being reliable and also having a healthy representation of what we care about in machine learning, otherwise, it might divert towards biased or unrealistic measures. Secondly, it needs to think about data security and privacy so that sensitive information is locked down without being accessed by anyone who is not supposed to [9]. Equal care must also be taken to craft and organize thorough, clearly arranged workflows for jobs that should run automatically. Tasks that tend to be both complex and repetitive are often composed of a series of related processes, making it difficult for workflow designers to create an effective design. Understanding the problem in detail and the processes already present is very necessary to ensure easy integration of your automation solution[10]. Additionally, it is important to maintain and update the workflow over time as necessary to make sure that your system can grow with process changes or updates. The main contribution of the research has the following:

- Enhanced Efficiency: By automating the mundane functions, which cloud AI services do provide a great deal of automation in their platforms, your business can operate more efficiently. This frees up resources to focus on other things that may need attention, making them more efficient and with better results.
- Better accuracy: Automation enhances the efficiency of cloud-based AI systems, reducing the chances of human error. This is particularly critical for AI, which can lead to drastic consequences if only a tiny percent of errors are allowed through. Automation performs repetitive tasks consistently and accurately; hence providing standardized results.
- Increased Scalability: With tasks such as auto-transcription, etc., taken care of automatically in the cloud-based AI systems, its scalability gets optimized since the functions can be made larger without passing on extra workload to human workers. This makes sure that the system can scale very quickly when there is a sudden spike in demand without the need for extra resources. So that better utilization of resources and cost savings.

The remaining part of the research has the following chapters. Chapter 2 describes the recent works related to the research. Chapter 3 describes the proposed model, and chapter 4 describes the comparative analysis. Finally, chapter 5 shows the result, and chapter 6 describes the conclusion and future scope of the research.

II. RELATED WORDS

Abdulazeez, D. H., et al.[11] A novel mechanism used a hybrid offloading approach that works on fuzzy logic and deep reinforcement learning to enhance the intelligence of IoT applications. This AI algorithm, based on estimated network conditions and app needs, determines what should be moved from the IoT device to the cloud which provides quicker response times as well as cost-effective usage of resources. Shahidinejad, A., et al.[12] Joint computation offloading and resource provisioning refers to the allocation of computing tasks between edge devices and (potentially geo-distributed) cloud servers in an efficient way, to enhance overall performance while keeping track of optimal usage of resources for their various classes. This uses the edge nodes which are close to where users are from and heavy computing available in the cloud as an advantage to have superior quality of service with low latency. Guo, W., et al.[13] Cloud resource scheduling is a broad class of problems that consists of how effectively and efficiently allocate virtual resources within the fog cloud to improve performance, energy consumption, and more as discussed .While deep reinforcement learning is used for decision-making by getting rewards and punishments from its algorithms, imitation learning imitates the behaviors of experts. Taking all these under consideration can address resource scheduling in a better way. Binh, H. T. T., et,al. [14] discussed the reinforcement learning algorithm for resource provisioning in mobile edge computing networks (Table 3). It involves using machine-based learning to allocate computing resources to devices located at the edge of a network. It constantly refines its decision-making process based on experience and current network conditions, which optimizes resource utilization to enhance overall network efficiency. Rawat, P. S., et al. Resource provisioning, allocation, and management of computing resources in large-scale cloud environments are studied by Rimal et al. [15]. Thus, we achieve this with a bio-inspired artificial neural network model that imitates the human way of managing resources preventing over- and under-provisioning based on dynamic requirements. This results in fulfilling resource requirements and enhancing the elasticity of the cloud systems. Anoushee, M., et al. [16] This particular paper titled “Resource Management in SDN-based Fog Computing Using Intelligent Reinforcement Learning” proposes an intelligent resource management scheme using RL by amalgamating with SDN and fog computing for on-demand dynamic resourcing in the context of a fog setup. It is more effective and scalable than the traditional strategies for resource management in fog computing networks, which use reinforcement learning algorithms for high-performance optimization. Patil, D. R., et al.[17] Dynamic resource allocation and memory management within cloud environments have been discussed .This includes resourcing learning, which enables

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

resource allocation and memory usage to be tuned according to real-time (and predictive) data as well. Scalability and Cloud computing cost savings are achieved by this approach. Funika, W., et al.[18] We have discussed Automated cloud resource provisioning using proximal policy optimization . This machine learning algorithm is used to efficiently allow the system of action on how virtual resources are allocated and managed in a cloud setting. The reinforcement learning methods are also incorporated into their algorithms to make the decisions for optimal resource allocation in real-time using live data. Du, Z., et al.[19]: Green deep reinforcement learning (GDRL)An emerging concept in the optimization of radio resource management for wireless communication networks, which has been discussed . It fingerprinting the solution effectively combines and implements deep learning algorithm with reinforcement learning in smart network resource allocation for energy efficiency in a friendly environment; The goal of GDRL is to improve system performance without increasing the power consumption or carbon footprint. Garí, Y., et al.[20] Reinforcement learning-based application autoscaling is also a dynamic mechanism for cloud-based applications resource allocation which had been discussed. It is a method wherein applications are experimented with and feedback in real-time to configure the systems' performance based on the changing requirements all while ensuring that we do not spend money needlessly. By utilizing machine learning algorithms, the system continuously learns and improves its resource allocation decisions leading to efficient and scalable application deployment in a cloud.

III. PROPOSED MODEL

This paper presents an alternative model for automating the repetitive workloads in cloud-based AI implementations and works to achieve it by using deep learning methods, thereby increasing both speed and accuracy. This model consists of: data preprocessing; choosing the deep learning algorithm and automated fine-tuning part.

The range of different parameter values and the dimensions of them are quite different, which are bound to influence the result of data analysis.

$$h^* = \frac{h - \min}{\max - \min}, \tag{1}$$

The data preprocessing step will be carried out by cleaning and organizing the data so it can be input into deep learning algorithms. Data formatting, Feature extraction, and data normalization. Then based on a proper deep learning model selection algorithm the approach will determine which type of neural network best suits your purpose. This is to be performed by understanding the data and characterizing what kind of problem it represents. The model determines the algorithm most suitable for dealing with data and giving perfect outcomes.

$$p = \frac{1}{I}\sum_{i=1}^{I} p_i(h). \tag{2}$$

$$k = p\left(Z^V h + i\right) \tag{3}$$

Finally, the automatic model refinement stage aims to continuously improve the algorithm chosen by feeding it new data and monitoring its performance. This allows the model to adapt and learn from new data, making it more efficient and accurate.

$$S_\pi(q, j) = G_\pi\left[\sum_{v=0}^{V} \gamma^{v_{r_v}} \left| q, j_v = j\right.\right] \tag{4}$$

The model even has a feedback loop where it gets input from users to improve the performance of the model. This enables the model with low growth and capacitive to deal with new data points intended for users. In this approach, which we call AutoDL-Projects: The purpose of launching the model as a copy system for automatic preprocessing, deep learning algo picking, and architecture searching (a.k.a. Model Refinement) in cloud-based AI systems is that they are made up of many repetitive tasks over time. Consistent (and Continuous) evolution and learning are necessary for an effective as well as a precise outcome in our ever-changing cloudscape.

3.1.Construction

© 2025 IJIRCCE | Volume 13, Issue 5, May 2025| DOI:10.15680/IJIRCCE.2025.1305004

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Cloud-based AI systems are becoming increasingly popular due to the cost-effective and scalable nature of cloud computing. These systems often involve a large amount of data and complex algorithms, making automation of repetitive tasks essential to ensure efficient and timely performance.

$$E(h)=\frac{h}{1+\|h\|} \tag{5}$$

$$\delta_v^{DQN}=y_t^{DQN}-Q\left(q_v,j_v;\theta_v^S\right) \tag{6}$$

Deep learning, a subset of machine learning, has proven to be a powerful tool for automating such tasks in cloud-based AI systems. The construction of automating repetitive tasks in cloud-based AI systems using deep learning involves several technical details.

$$k_v^{DQN}=L_{v+1}+\gamma \tag{7}$$

$$\eta(\pi)=G_\pi\left[\sum_{v=0}^{\infty}\gamma^v d\left(q_v\right)/q_0\sim\rho 0\right] \tag{8}$$

Firstly, the system must be trained on a large dataset to learn patterns and relationships between input data and output responses. This dataset can be obtained from various sources such as user interactions, sensor data, or manually labeled data. Next, the system must be designed and configured to handle the specific task for which it is being automated. Fig 1 shows the construction of the proposed model.

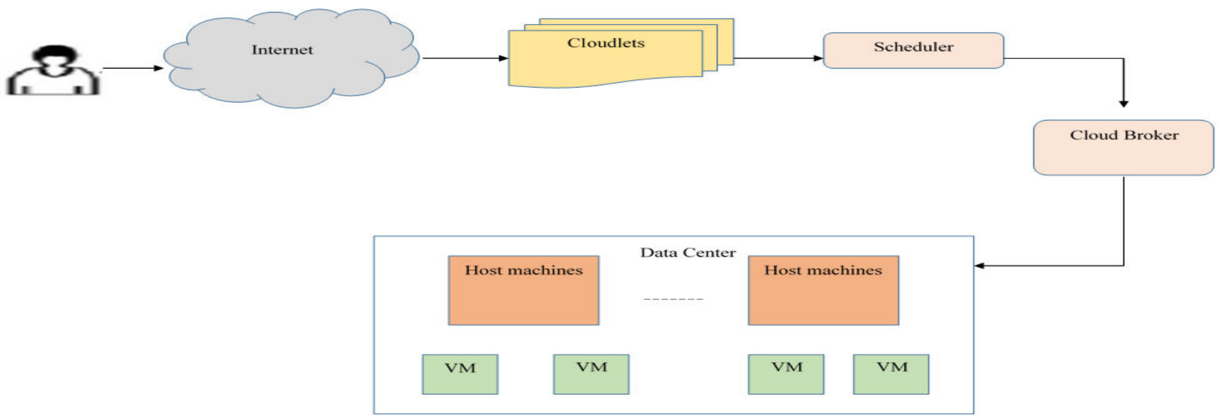


Fig 1 construction of the proposed model

This can involve finding the right architecture of a neural network along with activation functions and hyperparameters to make it perform at its best. You have to choose a deep learning framework and proper hardware configuration for more efficient system handling & training. When your system is trained and configured, you can then deploy it to a cloud. This covers introducing servers, systems, and capacity to give the framework a chance to work for clients. You can use cloud services like — Amazon Web Services, Microsoft Azure, or Google Cloud Platform for spinning these servers. It also must be monitored and evaluated regularly to increase improvement and stay up-to-date with new data or tasks. This could be anything from retraining the model with more/different data, modifying hyperparameters, or adding new functionality to the process itself.

3.2. Operating principle

A cloud-based AI system works based on its ability to automate the most repetitive (not yet learned) tasks, which invariably points one back toward that reasoning from deep learning frameworks. Deep learning is a subfield of machine learning, and it uses artificial neural networks to learn the data representations from large amounts of data. An artificial intelligence type that allows machines to simulate the highly complex processes of a human brain.

$$J^{\pi} = S^{fb}(q, j) - T^{\pi}(q) \tag{9}$$

$$\min \sum_{b=1}^{m} \left\| F_1^{'} - LF_2^{'} - V \right\|^2 \tag{10}$$

By automating the manual and repetitive tasks using deep learning algorithms to identify patterns in vast amounts of data that may not be obvious at first, then make a decision or take action with little human intervention. But this is most applicable to AI systems in the cloud, which is constantly cancerous amounts of data. In a cloud-based AI system, the first order of business when it comes to automating repetitive tasks is in collecting and preprocessing data. It consists of getting the right data, cleaning and organizing it, and converting that into a form on which deep learning models can be trained. Fig 2 shows the operating principle of the proposed model.

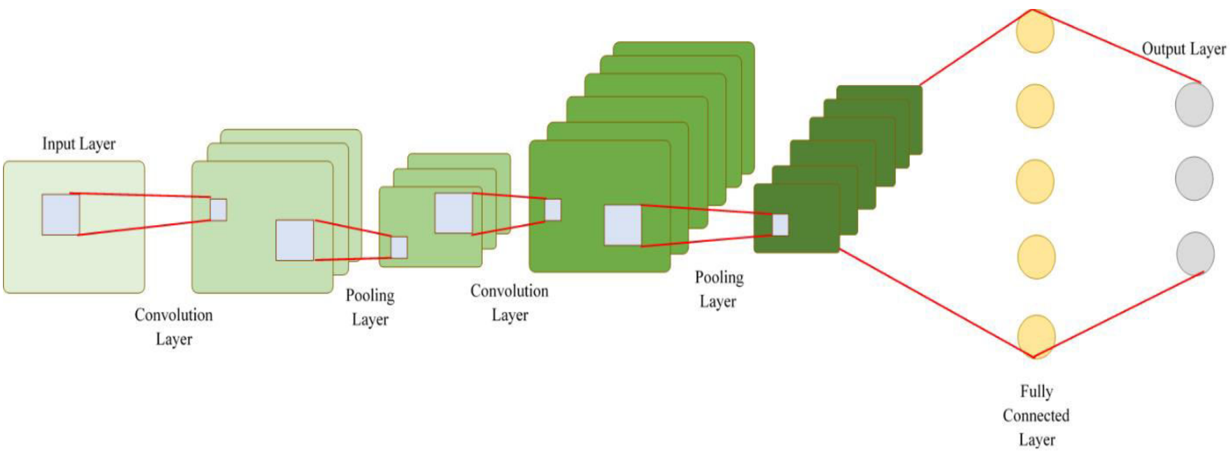


Fig 2 operating principle of the proposed model

The model passed under the layers, at this point, you are going to train your deep learning with preprocessed data. This means taking data through a neural network and changing model parameters to maximize the performance of a deep learning system. Then you can save your trained model and use it in the AI system above to run repetitive tasks autonomously. Subsequently, the system is free to carry on making certain decisions or doing whatever it does based on predictions of the model that has been built. You can keep iterating on this process as you go, which continues to increase the performance and efficiency of automating redundant jobs. In conclusion, the working mechanism around automating repetitive tasks involving cloud-based AI systems is based on deep learning algorithms where they scrutinize large datasets and make decisions accordingly spurring productivity during numerous chores.

IV. RESULT AND DISCUSSION

The proposed model ALARIC - Automated Learning with Advanced Repetitive Intelligence in the Cloud has been compared with the existing ACID - Automated Cloud-based Intelligent Deep learning for Repetitive tasks , DART - Deep Learning based Algorithm for Repetitive Tasks in Cloud environments and TOPLER - Task-Oriented Predictive Learning for Efficient Repetition in the cloud

4.1. Accuracy: This pertains to the ability of an automated system to execute tasks meticulously and accurately. In the case of cloud-based AI, accuracy is key to determining how well this system will perform reliability. Fig.3 shows the Comparison of Accuracy

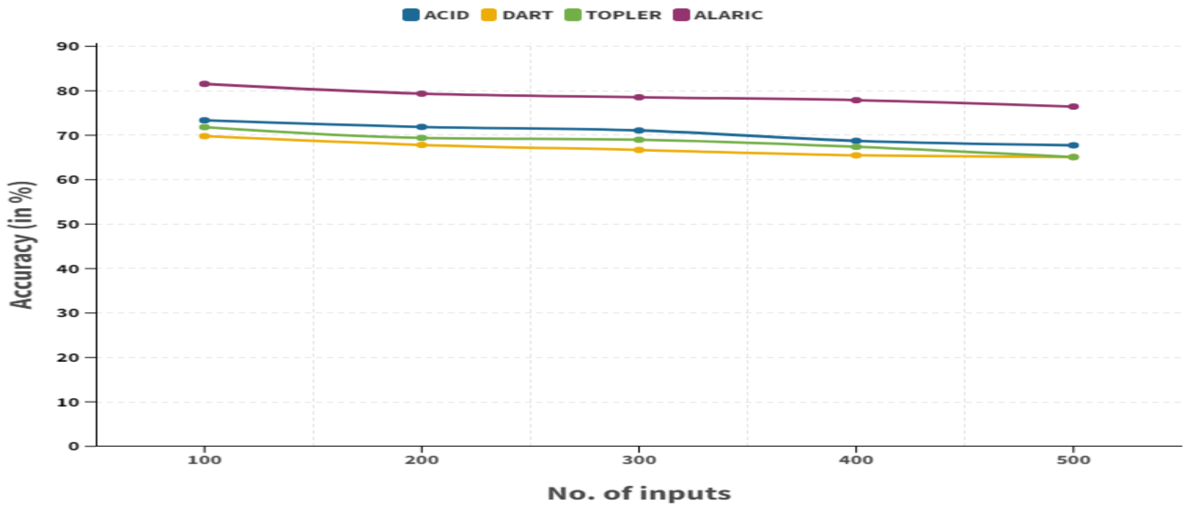


Fig.3 Comparison of Accuracy

4.2. Efficiency - Speed and resource utilization of the automated system In the case of Cloud-based AI systems, efficiency is particularly relevant because it affects how quickly and inexpensively a given task can be completed. Fig.4 shows the Comparison of Efficiency

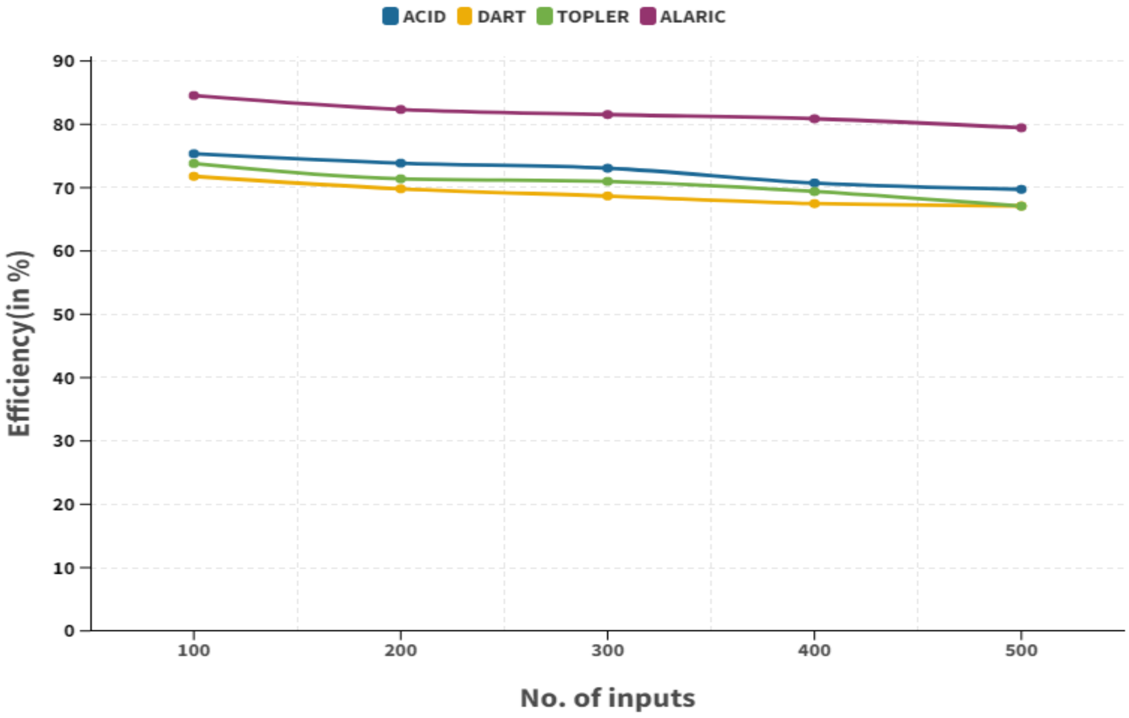


Fig.4 Comparison of Efficiency

4.3. Scalability: how a system can keep up with an increasing workload? Scalability — Given the incredibly large amount of data that cloud-based AI systems are witnessing, they need to be scalable. Fig.5 shows the Comparison of Scalability

© 2025 IJIRCCE | Volume 13, Issue 5, May 2025| DOI:10.15680/IJIRCCE.2025.1305004

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.771| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

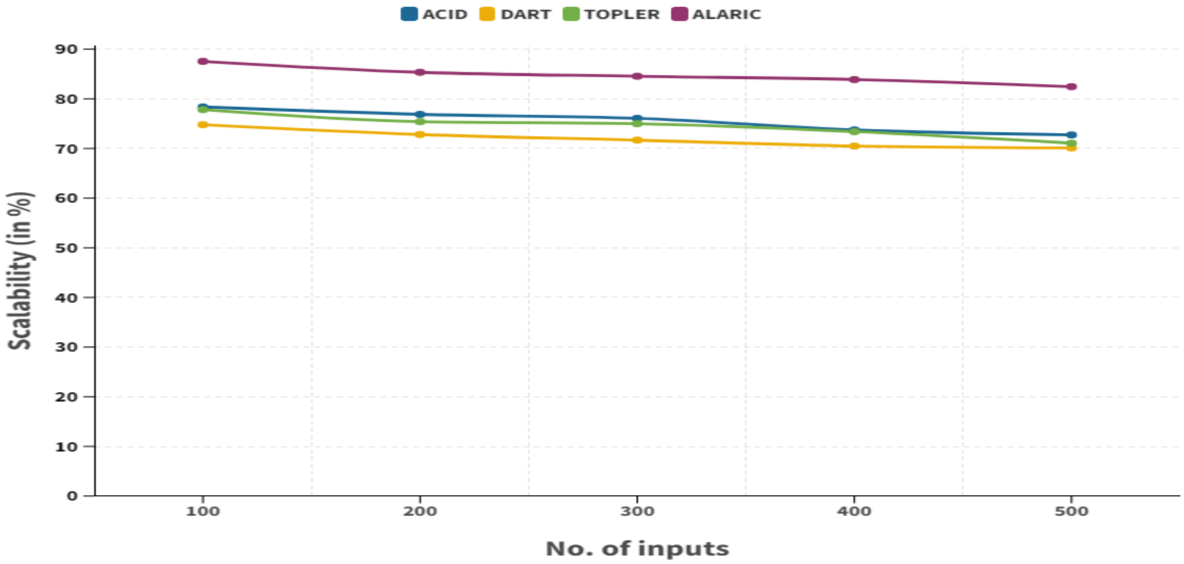


Fig.5 Comparison of Scalability

4.4. Robustness: It's the exact opposite, here we use this parameter to measure how well an automated system can cope with unexpected events and changing surroundings. Robustness is important in cloud-based AI systems, as it means the system can continue to function meaningfully even when events are not what was expected. Fig.6 shows the Comparison of Robustness

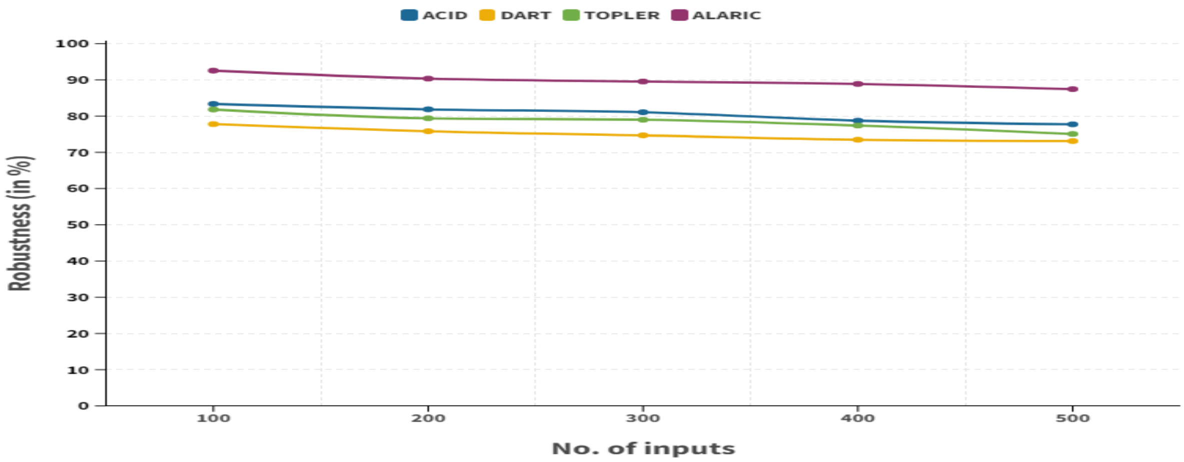


Fig.6 Comparison of Robustness

V. CONCLUSION

To summarize, automating time-consuming tasks in cloud-based AI systems will greatly increase the efficiency and reliability of solutions as well as reduce costs and human intervention. Now, with deep learning technology, one can automate many tasks that were performed manually and are better in accuracy. By offering a single point of entry, companies are reducing time and budgetary expenses, while moving onto more sophisticated scenarios with higher order value. However, automation in this regard should be well understood and continuously monitored to prove that the system remains effective with its implementation. In the end, however, deep learning making its way into these cloud-based AI systems makes it possible to use automation in so many different ways – an advantage that will continue to revolutionize how tasks are accomplished in a range of industries.

REFERENCES

1. Alizadeh Govarchinghaleh, Y., & Sabaei, M. (2024). Dynamic service provisioning in heterogeneous fog computing architecture using deep reinforcement learning. The Journal of Supercomputing, 1-44.
2. Chen, X., Zhu, F., Chen, Z., Min, G., Zheng, X., & Rong, C. (2020). Resource allocation for cloud-based software services using prediction-enabled feedback control with reinforcement learning. IEEE Transactions on Cloud Computing, 10(2), 1117-1129.
3. Santos, J., Wauters, T., Volckaert, B., & De Turck, F. (2021, May). Resource provisioning in fog computing through deep reinforcement learning. In 2021 IFIP/ieee international symposium on integrated network management (IM) (pp. 431-437). IEEE.
4. Zhou, G., Tian, W., Buyya, R., Xue, R., & Song, L. (2024). Deep reinforcement learning-based methods for resource scheduling in cloud computing: A review and future directions. Artificial Intelligence Review, 57(5), 124.
5. Ebadi, M. E., Yu, W., Rahmani, K. R., & Hakimi, M. (2024). Resource Allocation in The Cloud Environment with Supervised Machine learning for Effective Data Transmission. Journal of Computer Science and Technology Studies, 6(3), 22-34.
6. Jeyaraman, J., Bayani, S. V., & Malaiyappan, J. N. A. (2024). Optimizing Resource Allocation in Cloud Computing Using Machine Learning. European Journal of Technology, 8(3), 12-22.
7. Ramezani Shahidani, F., Ghasemi, A., Toroghi Haghighat, A., & Keshavarzi, A. (2023). Task scheduling in edge-fog-cloud architecture: a multi-objective load balancing approach using reinforcement learning algorithm. Computing, 105(6), 1337-1359.
8. Liang, H., Zhang, X., Hong, X., Zhang, Z., Li, M., Hu, G., & Hou, F. (2020). Reinforcement learning enabled dynamic resource allocation in the internet of vehicles. IEEE Transactions on Industrial Informatics, 17(7), 4957-4967.
9. Shruthi, G., Mundada, M. R., Sowmya, B. J., & Supreeth, S. (2022). Mayfly Taylor Optimisation-Based Scheduling Algorithm with Deep Reinforcement Learning for Dynamic Scheduling in Fog-Cloud Computing. Applied computational intelligence and soft computing, 2022(1), 2131699.
10. Jayanetti, A., Halgamuge, S., & Buyya, R. (2022). Deep reinforcement learning for energy and time optimized scheduling of precedence-constrained tasks in edge–cloud computing environments. Future Generation Computer Systems, 137, 14-30.
11. Abdulazeez, D. H., & Askar, S. K. (2024). A Novel Offloading Mechanism Leveraging Fuzzy Logic and Deep Reinforcement Learning to Improve IoT Application Performance in a Three-Layer Architecture Within the Fog-Cloud Environment. IEEE Access.
12. Shahidinejad, A., & Ghobaei-Arani, M. (2020). Joint computation offloading and resource provisioning for e dge-cloud computing environment: A machine learning-based approach. Software: Practice and Experience, 50(12), 2212-2230.
13. Guo, W., Tian, W., Ye, Y., Xu, L., & Wu, K. (2020). Cloud resource scheduling with deep reinforcement learning and imitation learning. IEEE Internet of Things Journal, 8(5), 3576-3586.
14. Binh, H. T. T., Le, N. P., Minh, N. B., Hai, T. T., & Minh, N. Q. (2020, July). A reinforcement learning algorithm for resource provisioning in mobile edge computing network. In 2020 International Joint Conference on Neural Networks (IJCNN) (pp. 1-7). IEEE.
15. Rawat, P. S., Dimri, P., Gupta, P., & Saroha, G. P. (2021). Resource provisioning in scalable cloud using bio-inspired artificial neural network model. Applied Soft Computing, 99, 106876.
16. Anoushee, M., Fartash, M., & Akbari Torkestani, J. (2024). An intelligent resource management method in SDN based fog computing using reinforcement learning. Computing, 106(4), 1051-1080.
17. Patil, D. R., & Sharma, M. (2020). Dynamic resource allocation and memory management using machine learning for cloud environments. International Journal, 9(4).
18. Funika, W., Koperek, P., & Kitowski, J. (2023). Automated cloud resources provisioning with the use of the proximal policy optimization. The Journal of Supercomputing, 79(6), 6674-6704.
19. Du, Z., Deng, Y., Guo, W., Nallanathan, A., & Wu, Q. (2020). Green deep reinforcement learning for radio resource management: Architecture, algorithm compression, and challenges. IEEE Vehicular Technology Magazine, 16(1), 29-39.
20. Yenugula., M et al., Dynamic Data Breach Prevention in Mobile Storage Media Using DQN-Enhanced Context Aware Access Control and Lattice Structures, IJRECE VOL, 10 issue 4 Oct-Dec 2022,pp 127-136.
21. Garí, Y., Monge, D. A., Pacini, E., Mateos, C., & Garino, C. G. (2021). Reinforcement learning-based application autoscaling in the cloud: A survey. Engineering Applications of Artificial Intelligence, 102, 104288.