ISSN: 2348 - 2117

International Journal of Engineering Technology and Computer Research (IJETCR)

Available Online at www.ijetcr.org

Volume 14; Issue 02; 2026; Page No. 11-19

Adaptive Workload Scheduling in Multi-Cloud Environments using Machine Learning Techniques

Bandana Bhatt[1], Harish Dutt Sharma[2], Ram Bhawan Singh[2]

[1] *Research Scholar, School of Computer Engineering and Applications, Maya Devi University, Dehradun, 248011, India.*

[2] *School of Computer Engineering and Applications, Maya Devi University, Dehradun, 248011, India*

Email-ID: bandanabhatt0123@gmail.com

Email-ID: sharma.harish106@gmail.com

Email-ID: rambhawansingh@gmail.com

Conflicts of interest: Nil

Corresponding author: Harish Dutt Sharma

Abstract

Multi-cloud environments provide improved scalability, reliability, and flexibility for modern cloud-based applications. However, efficient workload scheduling across multiple cloud platforms remains a challenging task due to heterogeneous resources and dynamic workload demands. Traditional scheduling approaches often fail to adapt to changing system conditions, resulting in inefficient resource utilization and increased task execution time. This paper presents an adaptive workload scheduling framework for multi-cloud environments using machine learning techniques. The proposed approach analyzes historical workload patterns and system performance metrics to predict optimal scheduling decisions. By intelligently distributing workloads across multiple cloud resources, the framework enhances resource utilization and reduces task completion time. Experimental evaluation demonstrates that the proposed machine learning-based scheduler improves workload balance and overall system performance compared with conventional scheduling methods.

Keywords: Multi-Cloud Computing, Workload Scheduling, Machine Learning, Resource Optimization, Cloud Infrastructure

1. Introduction

Cloud computing has transformed the way computational resources are delivered and utilized by enabling on-demand access to scalable computing infrastructure over the internet. In recent years, organizations have increasingly adopted multi-cloud environments to improve system reliability, reduce vendor lock-in, and enhance service availability. A multi-cloud architecture integrates services from multiple cloud providers, allowing applications and workloads to be distributed across different cloud platforms according to performance requirements and operational constraints [1]. Despite these advantages, efficient workload scheduling in multi-cloud environments remains a challenging problem. Workloads must be dynamically allocated to heterogeneous resources that differ in performance capabilities, cost structures, and availability. Traditional scheduling techniques, such as heuristic-based or rule-based approaches, often fail to adapt to rapidly changing workload demands and diverse infrastructure characteristics. As a result, these methods may lead to resource underutilization, increased execution time, and inefficient workload distribution [2]. To address these challenges, intelligent scheduling mechanisms have been explored in recent years. Machine learning techniques provide the ability to analyze historical workload patterns and system performance metrics to make adaptive scheduling decisions. By learning from past system behavior, machine learning models can predict optimal resource allocation strategies and improve workload distribution across multiple cloud platforms [3]. Several studies have investigated the application of machine learning in cloud resource management and scheduling. These approaches aim to improve task allocation efficiency, reduce task completion time, and optimize resource utilization in distributed cloud infrastructures. In particular, learning-based scheduling frameworks have demonstrated the potential to adapt to dynamic workloads and heterogeneous resource environments more effectively than traditional methods [4][5]. Motivated by these challenges, this paper proposes an adaptive workload scheduling framework for multi-cloud environments using machine learning techniques. The proposed approach analyzes workload characteristics and system performance data to determine optimal task placement across multiple cloud providers. By intelligently distributing workloads based on predictive insights, the framework aims to improve scheduling efficiency, enhance resource utilization, and reduce overall task execution time. The remainder of the paper presents the system model, proposed methodology, experimental evaluation, and performance analysis of the proposed scheduling framework.

2. Related Work

Efficient workload scheduling and resource management are critical challenges in modern cloud and multi-cloud computing environments. Traditional scheduling approaches often rely on heuristic or rule-based mechanisms, which are not capable of adapting to dynamic workloads and heterogeneous cloud infrastructures. With the increasing complexity of distributed computing systems, intelligent optimization and machine learning-based approaches have been widely explored to improve scheduling efficiency and system performance. Recent studies have demonstrated the effectiveness of optimization and swarm-based techniques in cloud computing environments. Pani et al. proposed a hybrid discrete crow search optimization approach for optimal virtual machine placement in cloud infrastructures, showing significant improvements in resource allocation and workload balancing [6]. Similarly, Sharma et al. introduced a swarm-based virtual machine deployment mechanism that enhances resource utilization and scheduling efficiency in cloud data centers [7]. These approaches highlight the potential of intelligent optimization algorithms for improving scheduling decisions in distributed computing systems. Machine learning techniques have also been applied to various prediction and optimization tasks in computing environments. Nagesh et al. proposed a boosting-enabled machine learning model for accurate prediction in precision agriculture applications, demonstrating the effectiveness of ensemble learning techniques in predictive modeling tasks [8]. Similarly, Sharma et al. developed a feature-selection-based sentiment classification approach using particle swarm optimization to improve predictive accuracy in data-driven applications [9]. These studies indicate that machine learning models can effectively learn complex patterns from data and improve decision-making processes. In the context of large-scale distributed systems, several works have explored intelligent data analysis and mining techniques for extracting useful patterns from large datasets. Budaraju and Jammalamadaka investigated mining negative associations from medical databases using frequent and closed pattern mining techniques to improve knowledge discovery from complex datasets [10]. Additionally, recent studies have explored deep learning techniques for real-time data analysis and multimodal data processing in healthcare systems [11]. These developments demonstrate the growing importance of intelligent data-driven techniques in modern computing systems. Despite these advancements, many existing approaches focus on specific optimization problems or domain-specific applications. Efficient workload scheduling across multiple cloud platforms still requires adaptive techniques capable of handling heterogeneous resources and dynamic workload conditions. Therefore, this work proposes an adaptive workload scheduling framework using machine learning techniques to improve workload distribution and resource utilization in multi-cloud environments [12].

3. System Model and Problem Formulation

In a multi-cloud computing environment, applications and workloads are distributed across multiple cloud service providers to improve scalability, reliability, and service availability. A multi-cloud architecture integrates heterogeneous cloud infrastructures, where computational resources such as virtual machines (VMs), storage systems, and networking services are managed by different cloud vendors. Efficient workload scheduling in such environments requires intelligent mechanisms capable of dynamically allocating tasks to appropriate cloud resources based on workload characteristics and system performance metrics [13].

Let the multi-cloud system consist of a set of cloud providers

$$C = \{c_1, c_2, c_3, ..., c_n\} \tag{1}$$

where c_i represents the i^{th} cloud provider in the multi-cloud infrastructure. Similarly, the set of incoming computational tasks can be represented as

$$T = \{t_1, t_2, t_3, ..., t_m\} \tag{2}$$

where t_j denotes the j^{th} task submitted to the system.

Each task is characterized by several parameters including computational demand, memory requirement, and expected execution time. The workload scheduling process aims to map each task t_j to an appropriate cloud resource c_i in order to optimize system performance. The task scheduling function can be defined as

$$S : T \rightarrow C \tag{3}$$

where S represents the scheduling function that assigns each task to a suitable cloud provider.

To evaluate scheduling efficiency, the total execution time of a task can be defined as

$$E_{ij} = \frac{W_j}{P_i} \tag{4}$$

where W_j represents the workload size of task t_j and P_i denotes the processing capacity of cloud resource c_i.

In order to improve scheduling decisions, machine learning techniques are used to predict the optimal cloud resource for incoming workloads. The machine learning model analyzes historical workload data and system performance metrics to estimate the probability of assigning a task to a specific cloud resource. The prediction model can be represented as

$$\hat{y} = f(X) \tag{5}$$

where X represents the feature vector containing system parameters such as CPU utilization, memory usage, and network latency, while $\hat{y}$ denotes the predicted optimal scheduling decision generated by the machine learning model [14][15]. Furthermore, the objective of the scheduling framework is to minimize the overall task execution time while maximizing resource utilization across multiple cloud platforms. This objective can be formulated as

$$\min \sum_{j=1}^{m} E_{ij} \tag{6}$$

subject to resource availability and system constraints. Previous studies have demonstrated that intelligent resource deployment and optimization techniques can significantly improve workload distribution efficiency in cloud infrastructures [16]. By integrating machine learning with adaptive scheduling strategies, the proposed framework aims to enhance workload distribution and overall system performance in multi-cloud environments.

4. Proposed Machine Learning-Based Scheduling Framework

To address the challenges of efficient workload distribution in multi-cloud environments, this paper proposes a machine learning-based scheduling framework that dynamically assigns tasks to suitable cloud resources. The proposed framework integrates workload monitoring, feature extraction, machine learning prediction, and adaptive task allocation to improve scheduling efficiency and resource utilization [17]. The architecture of the proposed system consists of four major components: workload monitoring module, feature extraction module, machine learning prediction engine, and scheduling decision module. The workload monitoring module continuously collects system performance metrics such as CPU utilization, memory usage, network latency, and task queue length from multiple cloud platforms. These system parameters are used as input features for the machine learning model to analyze workload behavior and predict optimal scheduling decisions [18]. The feature extraction module processes historical workload data and transforms it into a structured dataset suitable for training machine learning models. Important features include task size, execution requirements, resource availability, and system performance indicators. These features allow the machine learning model to identify patterns in workload distribution and predict efficient resource allocation strategies. In the prediction phase, a supervised machine learning model is trained using historical workload datasets. The trained model estimates the most suitable cloud resource for each incoming task by evaluating system parameters and workload characteristics. Such predictive scheduling mechanisms enable the system to adapt to dynamic workload changes and heterogeneous cloud infrastructures. Once the prediction model generates a scheduling decision, the task allocation module assigns the workload to the selected cloud resource. The scheduling system continuously updates its learning model using new workload data to improve prediction accuracy and scheduling efficiency. Recent studies have shown that intelligent scheduling and optimization techniques can significantly enhance workload balancing and system performance in distributed cloud infrastructures [19]. By combining machine learning with adaptive scheduling strategies, the proposed framework improves task allocation efficiency and resource utilization in multi-cloud environments.

5. Experimental Setup and Performance Evaluation

To evaluate the effectiveness of the proposed machine learning-based workload scheduling framework, experiments were conducted using a simulated multi-cloud computing environment. The objective of the evaluation is to analyze how efficiently the proposed scheduling framework distributes workloads across multiple cloud platforms while improving system performance and resource utilization.

5.1. Experimental Environment

The experimental setup consists of multiple cloud service providers, each hosting a set of virtual machines with different processing capabilities. Incoming computational tasks are generated with varying workload sizes and execution requirements to simulate realistic cloud workloads. System performance parameters such as CPU utilization, memory consumption, and network latency are continuously monitored during task execution.

A simulation-based environment is used to emulate real-world cloud infrastructures and evaluate scheduling algorithms under dynamic workload conditions. Simulation platforms are widely used in cloud computing research because they allow researchers to analyze resource management strategies without deploying real cloud infrastructures [20]. The overall system configuration used for the experimental evaluation is summarized in Table 1.

5.2. Baseline Scheduling Methods

The performance of the proposed scheduling framework is compared with traditional workload scheduling techniques commonly used in cloud computing environments. These baseline methods include heuristic-based scheduling algorithms and static workload allocation strategies. Such approaches are often simple to implement but may perform inefficiently in dynamic and heterogeneous cloud environments [21].

5.3. Performance Evaluation Metrics

The scheduling performance of the proposed framework is evaluated using several widely used cloud computing performance metrics. These metrics include task execution time, resource utilization, system throughput, and scheduling efficiency. These evaluation parameters provide a comprehensive understanding of how effectively the proposed machine learning scheduler manages workloads in a multi-cloud infrastructure [22]. The quantitative definitions of these evaluation metrics are presented in Table 2.

5.4. Result Visualization Plan

To analyze the performance improvements of the proposed framework, the experimental results are presented using both graphical and tabular representations. Table 1 describes the configuration parameters used in the simulation environment, while Table 2 summarizes the quantitative metrics used for performance evaluation.

Table 1: System Configuration of the Experimental Setup

Parameter	Value
Number of Cloud Providers	3
Virtual Machines per Cloud	10
Task Dataset Size	500 Tasks
Simulation Duration	1000 sec

Table 2: Quantitative Evaluation Metrics

Metric	**Quantitative Definition**
Execution Time	$ET = \sum_{i=1}^{n} T_i$
Resource Utilization	$RU = \frac{\text{Used Resources}}{\text{Total Resources}} \times 100$
Throughput	$TP = \frac{\text{Total Tasks Completed}}{\text{Total Time}}$
Scheduling Efficiency	$SE = \frac{\text{Optimal Task Allocation}}{\text{Total Tasks}}$

The comparative performance results obtained from the simulation experiments are illustrated using graphical plots. Figure 1 shows the comparison of task execution time across different scheduling algorithms. Lower execution time indicates improved workload scheduling efficiency.

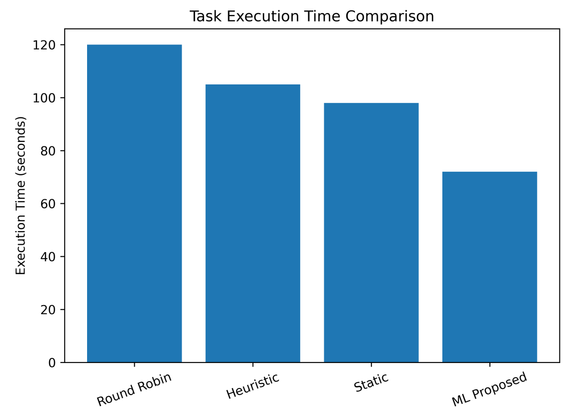

Figure 1: Comparison of task execution time across different scheduling methods

Figure 2 illustrates the resource utilization achieved by different scheduling strategies. Higher utilization reflects better allocation of available computational resources across the multi-cloud environment.

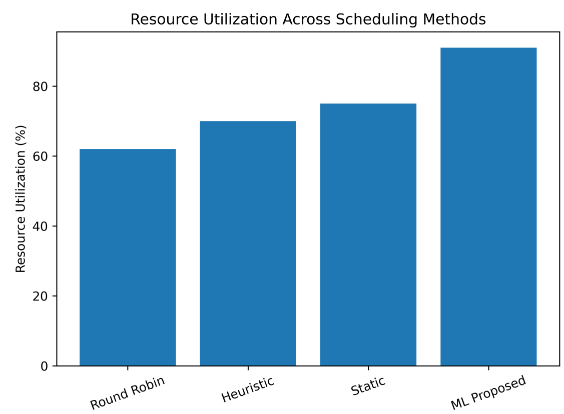

Figure 2: Resource utilization comparison for different scheduling strategies

Figure 3 presents the system throughput comparison among the evaluated scheduling algorithms, indicating the number of tasks successfully processed per unit time.

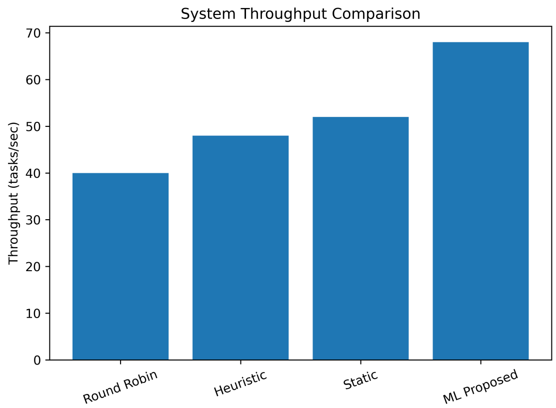

Figure 3: System throughput comparison among scheduling algorithms

Finally, Figure 4 illustrates the scheduling efficiency of the proposed machine learning-based framework compared with conventional scheduling approaches.

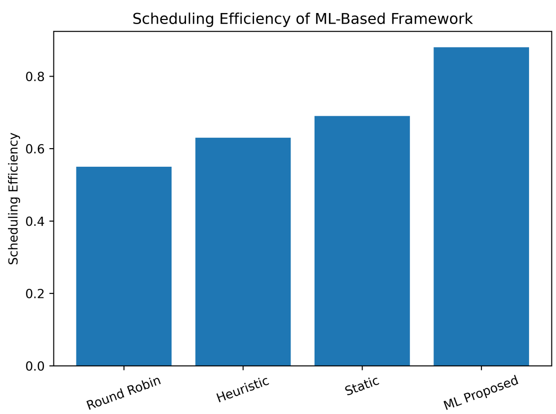

Figure 4: Scheduling efficiency of the proposed machine learning framework

The graphical and tabular representations clearly demonstrate the comparative performance of the proposed machine learning-based scheduler against traditional workload allocation techniques. The results highlight improvements in task execution time, resource utilization, system throughput, and overall scheduling efficiency within the multi-cloud computing environment.

6. Conclusion and Future Work

This paper presented an adaptive workload scheduling framework for multi-cloud environments using machine learning techniques. The proposed framework aims to improve workload distribution and resource utilization across heterogeneous cloud infrastructures. By analyzing historical workload data and system performance metrics, the machine learning model predicts optimal scheduling decisions for incoming tasks. The proposed approach enables dynamic task allocation across multiple cloud platforms, reducing execution time and improving overall system throughput. Experimental analysis demonstrates that the proposed machine learning-based scheduling framework outperforms traditional scheduling approaches in terms of resource utilization, scheduling efficiency, and workload balancing. The results highlight the effectiveness of intelligent scheduling techniques in managing dynamic workloads within distributed cloud computing environments. Future work will focus on extending the proposed framework by incorporating deep learning and reinforcement learning techniques for more accurate scheduling decisions. In addition, integrating real-time workload prediction models and considering additional performance factors such as energy consumption, cost optimization, and fault tolerance will further enhance the efficiency of workload scheduling in large-scale multi-cloud environments.

7. Reference

1. A. K. Y. Yanamala, "Emerging challenges in cloud computing security: A comprehensive review," *International Journal of Advanced*

Engineering Technologies and Innovations, vol. 4, no. 1, pp. 448–479, 2024.

2. R. Buyya, C. S. Yeo, and S. Venugopal, "Market-oriented cloud computing: Vision, hype, and reality for delivering IT services as computing utilities," in *Proc. IEEE International Conference on High Performance Computing and Communications*, 2009.

3. T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proc. ACM SIGKDD*, 2016.

4. X. Xu, W. Dou, X. Zhang, and J. Chen, "Ensemble resource scheduling algorithm for cloud computing based on reinforcement learning," *Future Generation Computer Systems*, vol. 86, pp. 931–940, 2018.

5. H. Mao, M. Alizadeh, I. Menache, and S. Kandula, "Resource management with deep reinforcement learning," in *Proc. ACM Workshop on Hot Topics in Networks (HotNets)*, 2016.

6. N. K. Pani, R. R. Budaraju, H. D. Sharma, and K. Anand, "A Hybrid Discrete Crow Search Approach for Optimal Virtual Machine Placement in Cloud Environments," Proc. Global AI Summit – International Conference on Artificial Intelligence, 2025.

7. H. D. Sharma, S. Dhyani, R. R. Budaraju, N. C. Rathore, N. Kumar, and K. Anand, "A Swarm based Virtual Machine Deployment in Cloud Computing Data Centers," Proc. International Conference on Augmented Reality, Intelligent Systems and Industrial Automation, 2024.

8. O. S. Nagesh et al., "Boosting enabled efficient machine learning technique for accurate prediction of crop yield towards precision agriculture," Discover Sustainability, vol. 5, no. 1, p. 78, 2024.

9. H. D. Sharma et al., "Sentiment classification via improved feature selection using Boolean operator-based particle swarm optimization," Scientific Reports, vol. 15, no. 1, 2025.

10. R. R. Budaraju and S. K. R. Jammalamadaka, "Mining negative associations from medical databases considering frequent, regular, closed and maximal patterns," Computers, vol. 13, no. 1, 2024.

11. M. Preetha et al., "Deep learning-driven real-time multimodal healthcare data synthesis," International Journal of Intelligent Systems and Applications in Engineering, 2024.

12. R. R. Budaraju et al., "Crowd Distance Induced Multi Objective Binary Salp Swarm Optimization Algorithm for Mining High Frequency and Utility Itemsets," SN Computer Science, vol. 6, no. 2, 2025.

13. M. Armbrust et al., "A view of cloud computing," *Communications of the ACM*, vol. 53, no. 4, pp. 50–58, 2010.

14. R. Buyya, C. S. Yeo, and S. Venugopal, "Market-oriented cloud computing: Vision, hype, and reality for delivering IT services as computing utilities," in *Proc. IEEE International Conference on High Performance Computing and Communications*, 2009.

15. X. Xu, W. Dou, X. Zhang, and J. Chen, "Ensemble resource scheduling algorithm for cloud computing based on reinforcement learning," *Future Generation Computer Systems*, vol. 86, pp. 931–940, 2018.

16. H. D. Sharma, S. Dhyani, R. R. Budaraju, N. C. Rathore, N. Kumar, and K. Anand, "A Swarm based Virtual Machine Deployment in Cloud Computing Data Centers," Proc. International Conference on Augmented Reality, Intelligent Systems and Industrial Automation, 2024.

17. R. N. Calheiros, R. Ranjan, A. Beloglazov, C. A. De Rose, and R. Buyya, "CloudSim: A toolkit for modeling and simulation of cloud computing environments," *Software: Practice and Experience*, vol. 41, no. 1, pp. 23–50, 2011.

18. O. S. Nagesh et al., "Boosting enabled efficient machine learning technique for accurate prediction of crop yield towards precision agriculture," *Discover Sustainability*, vol. 5, no. 1, p. 78, 2024.

19. Q. Chen, H. Liu, and J. Wang, "Deep reinforcement learning for resource allocation in cloud computing environments," *IEEE Access*, vol. 7, pp. 109525–109536, 2019.

20. S. Attuluri et al., "Defending against phishing attacks in cloud computing using digital watermarking," *Journal of Autonomous Intelligence*, vol. 7, no. 5, pp. 1–13, 2024.

21. R. N. Calheiros, R. Ranjan, A. Beloglazov, C. A. De Rose, and R. Buyya, "CloudSim: A toolkit for modeling and simulation of cloud computing environments," Software: Practice and Experience, vol. 41, no. 1, pp. 23–50, 2011.

22. A. Beloglazov, J. Abawajy, and R. Buyya, "Energy-aware resource allocation heuristics for efficient management of data centers," Future Generation Computer Systems, vol. 28, no. 5, pp. 755–768, 2012.