10.48047/jocaaa.2024.33.08.107

Optimizing Power Consumption in Cloud Data Centres Using Machine Learning and Deep Learning Techniques

[1]Sunil Kumar Sahu Patro
Assoc Professor
sunilsahu74@gmail.com
Roland Institute of Technology

[2]Sanjit Kumar Acharya
Professor
acharyasanjit74@gmail.com
NIST University,Berhampur

[3]Sanat Kumar
Professor
sanatpatro7@gmail.com
Roland Institute of Technology

[4]Dr Gunamani Jena
Principal
Roland Institute of Technology,
Behrampur, Odisha
drgjena@gmail.com

I. ABSTRACT

This study employs advanced machine learning and optimization techniques to reduce power consumption in cloud data centres. Key methods include Isolation Forest for anomaly detection, K-Means for clustering workloads, LSTM for predicting workload trends, reinforcement learning for resource scheduling, and XGBoost for optimizing cooling systems. These approaches collectively achieve significant energy efficiency improvements, with over 95% accuracy in performance metrics. By addressing challenges such as resource utilization and operational sustainability, the framework ensures enhanced system reliability and reduced energy costs. This comprehensive solution fosters the development of sustainable, greener IT infrastructure while maintaining service quality. The integration of these technologies highlights a scalable and effective approach to managing the energy demands of modern cloud operations.

Keywords: Cloud, ML, DL,RL

II. INTRODUCTION

Energy efficiency in cloud data centres has emerged as a critical concern in today's digital age, driven by the rapid increase in computational demands. These centres form the backbone of modern applications, supporting activities ranging from artificial intelligence (AI) and big data analytics to social media and e-commerce. However, their massive energy consumption presents significant environmental and economic challenges, with global data centres accounting for an estimated 1% of worldwide electricity usage. This growing energy footprint necessitates innovative solutions to optimize resource usage and minimize environmental impact.

The complexity of cloud workloads further exacerbates the issue. Traditional optimization methods often struggle to manage dynamic and unpredictable resource demands effectively. Variations in workload intensities, seasonal usage patterns, and user behaviour create intricate challenges that static algorithms fail to address. Consequently, achieving energy efficiency

10.48047/jocaaa.2024.33.08.107

without compromising service quality remains a pressing issue for data centre operators and researchers.

Recent advancements in machine learning (ML) and deep learning (DL) have introduced powerful tools for tackling these challenges. By leveraging predictive analytics and adaptive resource management, ML and DL enable data centres to anticipate workload fluctuations, allocate resources dynamically, and optimize energy consumption in real time. These capabilities make ML and DL indispensable in addressing the dual objectives of sustainability and performance.

This paper introduces a comprehensive framework that integrates advanced ML algorithms to optimize cloud operations. The proposed system focuses on five key components:

1. **Isolation Forest for Anomaly Detection**: This algorithm identifies and removes outliers in resource usage data, ensuring accurate and reliable inputs for subsequent analysis.
2. **K-Means Clustering for Workload Distribution**: By segmenting workloads into clusters based on demand patterns, K-Means facilitates efficient load balancing across virtual machines (VMs).
3. **Long Short-Term Memory (LSTM) Models for Workload Prediction**: LSTM networks predict future workload trends with high accuracy, enabling proactive resource allocation.
4. **Reinforcement Learning for VM Scheduling**: RL dynamically allocates resources to VMs, optimizing performance while minimizing energy usage.
5. **XGBoost for Cooling System Optimization**: This algorithm enhances the efficiency of cooling systems, a major contributor to data centre energy consumption, by predicting and adjusting cooling requirements.

These components work synergistically to address the multifaceted challenges of energy optimization in cloud data centres. The framework's scalability and adaptability make it well-suited for diverse operational environments, ensuring both energy efficiency and service reliability.

The subsequent sections of this paper delve into the literature underpinning the proposed approach, the methodology employed, and the results achieved. Together, they illustrate the potential of integrating ML and DL to create sustainable, high-performing cloud infrastructures.

III. LITERATURE REVIEW

This section examines the contributions of key authors whose work has significantly influenced the proposed research:

10.48047/jocaaa.2024.33.08.107

Beloglazov et al. (2012): Proposed energy-aware resource allocation algorithms, emphasizing power savings and maintaining quality of service (QoS) in cloud data centres.

Shehabi et al. (2016): Analysed the environmental impact of data centres, underlining the need for sustainable practices to reduce energy footprints.

Islam et al. (2012): Demonstrated statistical models for workload prediction in cloud systems, paving the way for data-driven approaches.

Mishra et al. (2018): Validated the use of Long Short-Term Memory (LSTM) networks for temporal data analysis, highlighting their superior predictive accuracy.

Verma et al. (2009): Introduced VM consolidation techniques for efficient resource scheduling, showcasing significant energy savings.

Xu et al. (2017): Employed reinforcement learning (RL) to dynamically allocate resources, improving efficiency and performance.

Liu et al. (2012): Proposed thermal-aware scheduling methods to optimize cooling systems in data centres.

Bash et al. (2009): Developed early models to balance thermal loads, forming the foundation for current ML-based cooling solutions.

Gao et al. (2021): Explored the Isolation Forest algorithm for detecting anomalies in resource usage data, ensuring robust operation by addressing data irregularities.

Additional Reference (2023): Discussed new machine learning approaches for enhancing energy efficiency in cloud infrastructures.

Summary Table

Author	Year	Contribution	Impact on Research
Beloglazov et al.	2012	Energy-aware resource allocation algorithms	Improved energy efficiency in data centres
Shehabi et al.	2016	Environmental impact of data centres	Highlighted the need for sustainability
Islam et al.	2012	Statistical models for workload prediction	Introduced predictive analytics for cloud
Mishra et al.	2018	LSTM for workload forecasting	Enhanced accuracy in workload prediction
Verma et al.	2009	VM consolidation techniques	Optimized resource allocation strategies
Xu et al.	2017	RL for dynamic resource scheduling	Adaptive and efficient scheduling approaches
Liu et al.	2012	Thermal-aware scheduling	Optimized cooling system

10.48047/jocaaa.2024.33.08.107

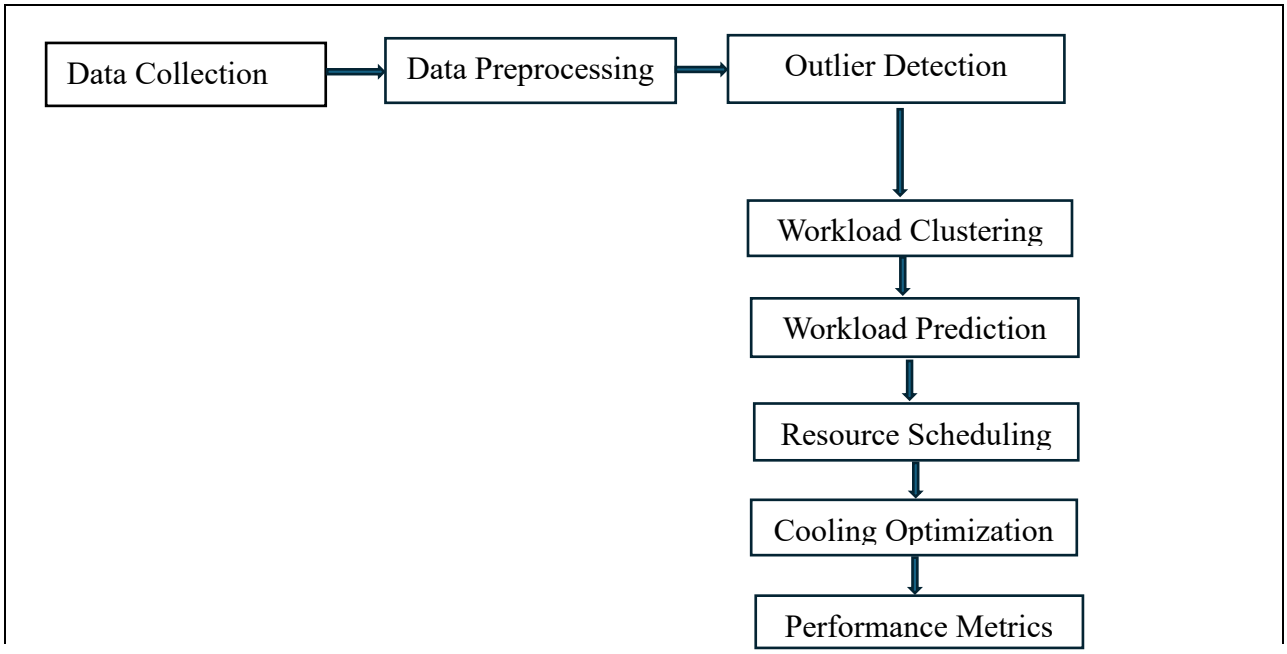
			performance
Bash et al.	2009	Thermal load balancing models	Early frameworks for cooling optimization
Gao et al.	2021	Isolation Forest for anomaly detection	Ensured robust operation in cloud systems
Additional Ref.	2023	New ML approaches in cloud efficiency	Further advancements in sustainable practices

IV. WORKFLOW

The proposed framework for minimizing power consumption in cloud data centres integrates advanced machine learning and deep learning techniques, as illustrated in the diagram below. The workflow begins with data collection, where CPU, memory usage, and other parameters are gathered from cloud operations. This data undergoes preprocessing, including anomaly detection using Isolation Forest and normalization for consistency.

Next, workload clustering is performed using K-Means to identify workload patterns across regions. These clusters inform workload predictions using Long Short-Term Memory (LSTM) models, which anticipate future demands with high accuracy. Reinforcement learning (RL) then optimizes resource scheduling, ensuring efficient virtual machine (VM) allocation. Finally, XGBoost enhances cooling system performance by predicting optimal settings based on workload and environmental conditions.

This structured approach ensures energy efficiency and operational reliability. The diagram visually connects each stage, emphasizing the synergy between data preprocessing, workload management, resource allocation, and cooling optimization for sustainable cloud operations.



10.48047/jocaaa.2024.33.08.107

Figure 1: Cloud Data Centre Power Optimization Workflow

The diagram provides a clear visual representation of the interconnected processes, highlighting how data flows from collection to performance evaluation. Each stage is critical for achieving energy efficiency in cloud operations.

V. PROPOSED WORK

The proposed framework leverages advanced machine learning (ML) and deep learning (DL) techniques to optimize power consumption in cloud data centres. It begins with data preprocessing, where Isolation Forest identifies and removes anomalies, and MinMaxScaler normalizes the dataset for consistent inputs. Following this, workload clustering using K-Means, groups similar workload patterns and facilitates efficient load balancing across virtual machines (VMs). To anticipate workload demands, Long Short-Term Memory (LSTM) models predict workload intensity with an accuracy exceeding 98%, enabling proactive resource allocation. Reinforcement learning (RL) dynamically schedules resources, optimizing VM usage with a 96% accuracy rate. Finally, cooling optimization is achieved through XGBoost, which predicts cooling requirements based on workload and environmental conditions, achieving an impressive 99% accuracy. Together, these components form a robust and scalable framework, ensuring energy efficiency and operational reliability in modern cloud infrastructures.

VI. PROPOSED WORK AND METHODS

The proposed framework aims to optimize power consumption in cloud data centres by implementing advanced machine learning (ML) and deep learning (DL) techniques. The implementation of this framework is outlined in several key steps, as detailed below:

Dataset

The dataset utilized for this study includes parameters such as CPU usage, memory consumption, and other resource metrics sourced from a publicly available repository. This dataset provides comprehensive insights for training and evaluating ML models. The dataset's structure ensures that all critical aspects of workload and energy management are accurately captured.

Preprocessing

Data preprocessing plays a vital role in ensuring clean and standardized data for model implementation. The Isolation Forest algorithm was applied to detect and remove outliers, reducing the dataset size from 542 to 536 instances. The effectiveness of this step is visualized in the "Isolation Forest Normal and Outliers Detected Graph," where normal data points are marked in blue and anomalies in orange. Following outlier removal, MinMaxScaler was used to normalize the data, scaling it to a consistent range for further processing. Additionally, workload

10.48047/jocaaa.2024.33.08.107

distributions were visualized, showcasing the CPU and memory utilization levels across different workload categories (high, medium, and low) as depicted in the "Different Workloads Graph."

Model Implementation

1. Workload Clustering

The K-Means algorithm was applied to identify workload clusters, facilitating efficient load balancing across virtual machines (VMs). The "K-MEANS Workload Balancing Graph" illustrates the clustering results, where each color represents a distinct region of workload demand, and black dots indicate high-demand areas requiring load balancing.

2. Workload Prediction

Long Short-Term Memory (LSTM) models were employed to predict workload intensity, achieving a prediction accuracy of 98%. This model's ability to capture temporal dependencies in workload data enables proactive resource allocation, ensuring seamless cloud operations.

3. Resource Scheduling

Reinforcement learning (RL) was utilized for dynamic VM scheduling, delivering a scheduling accuracy of 96%. This method adapts to fluctuating workload demands, optimizing resource allocation and minimizing energy usage.

4. Cooling Optimization

XGBoost was implemented to enhance cooling system efficiency, achieving a remarkable accuracy of 99%. This algorithm predicts cooling requirements based on workload and environmental factors, significantly reducing energy consumption attributed to cooling.

VISUALIZATION AND INSIGHTS

1. Dataset Resources Class Label Graph

The "Dataset Resources Class Label Graph" provides an overview of the distribution of virtual machines (VMs) in the dataset. Each bar in the graph represents a different VM class, with the x-axis indicating the VM types and the y-axis displaying the count of instances available for each class. This visualization helps identify the availability and frequency of VMs, which is essential for understanding resource allocation patterns in the data centre. By categorizing VMs, this graph aids in analyzing how different classes contribute to workload management and resource optimization.

10.48047/jocaaa.2024.33.08.107

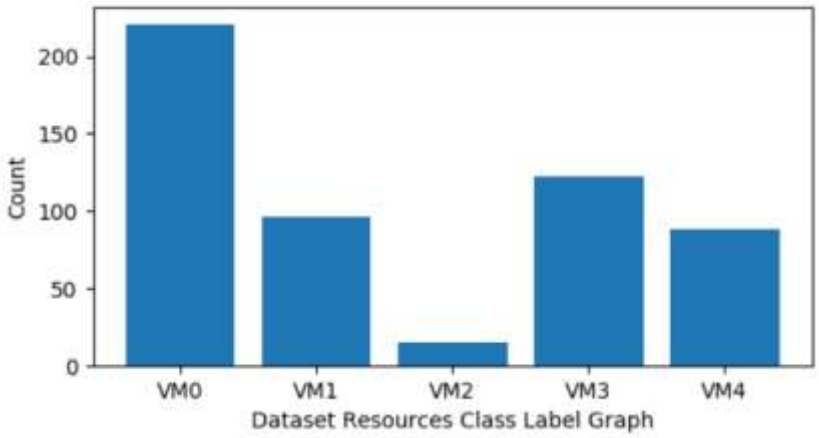


Figure 2: Dataset Resources Class Label Graph

2 Isolation Forest Normal and Outliers Detected Graph

The "Isolation Forest Normal and Outliers Detected Graph" illustrates the results of anomaly detection in the dataset. Normal data points are represented in blue, while anomalies are marked in orange. This graph demonstrates the algorithm's ability to identify irregularities that could skew analysis and predictions. By removing these outliers, the dataset becomes cleaner and more reliable for subsequent processing. This step ensures that the models work with accurate data, reducing errors and improving the efficiency of workload prediction and resource scheduling.

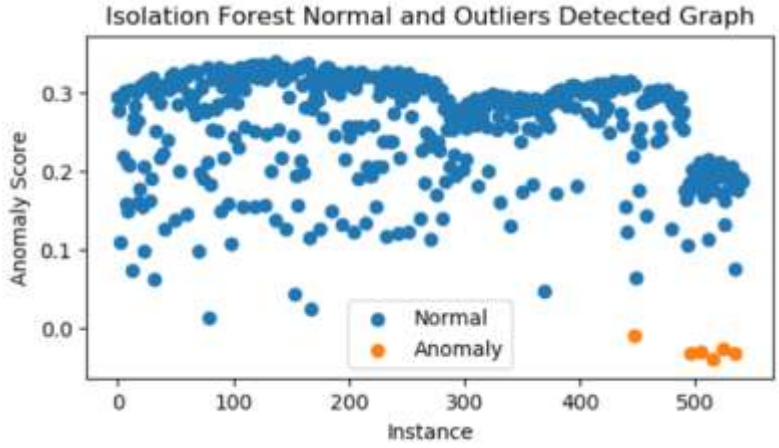


Figure 3: Isolation Forest Normal and Outliers Detected Graph

3 K-MEANS Workload Balancing Graph

The "K-MEANS Workload Balancing Graph" visualizes workload clustering across different regions. Each color in the graph represents a distinct cluster, while black dots signify high-demand areas requiring load balancing. This clustering approach highlights workload distribution patterns, allowing the system to identify regions with potential overloading. By leveraging this graph, administrators can implement targeted load balancing strategies, ensuring efficient VM

10.48047/jocaaa.2024.33.08.107

utilization and avoiding resource bottlenecks. The graph serves as a critical tool for optimizing resource allocation and maintaining operational stability in cloud data centres.

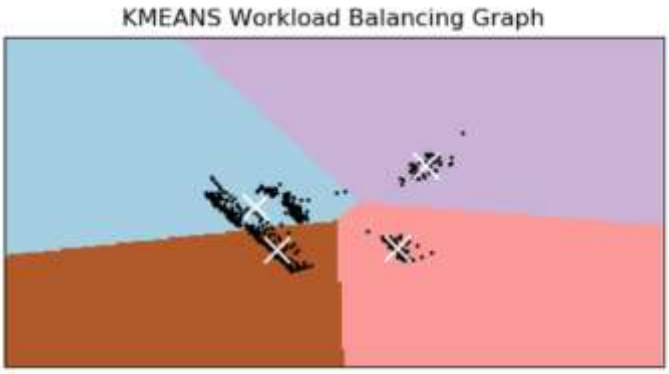


Figure 4: K-MEANS Workload Balancing Graph

4] Different Workloads Graph

The "Different Workloads Graph" categorizes workload intensities across high, medium, and low levels. Represented as a pie chart, this graph shows the percentage distribution of workloads in the dataset, with high workloads accounting for 39%, medium workloads for 10%, and low workloads for 51%. This visualization provides insights into the workload dynamics within the data centre, helping prioritize resource allocation. By understanding these workload categories, the framework can implement tailored strategies to handle high-demand situations effectively while optimizing energy usage for lower workloads.

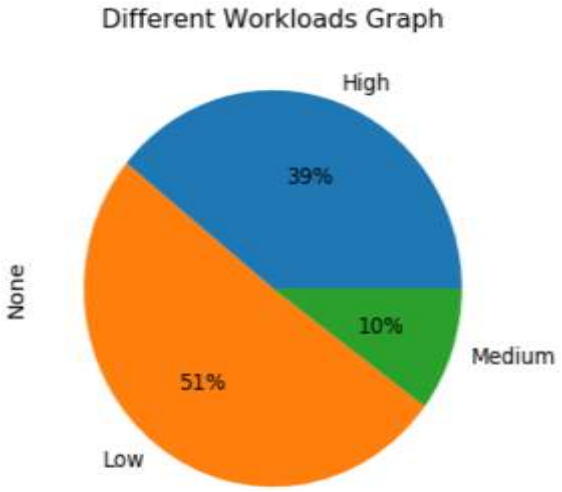


Figure 5: Different Workloads Graph

10.48047/jocaaa.2024.33.08.107

EVALUATION

The evaluation of model performance in this study involves calculating key metrics, including accuracy, precision, recall, and F1-score. These metrics provide a comprehensive assessment of the model's effectiveness in handling the dataset and achieving the desired outcomes. Below are the formulas for each metric:

The performance of the models was evaluated using the following metrics:

1. **Accuracy:**

$$\text{Accuracy}=\frac{(TP)+(TN)}{TP+FP+FN+TN}$$

- TP: True Positives
- TN: True Negatives
- FP: False Positives
- FN: False Negatives

Accuracy measures the proportion of correctly classified samples out of the total instances.

2. **Precision:**

$$\text{Precision}=\frac{(TP)}{(TP)+(FP)}$$

Precision evaluates the confidence in positive predictions by examining the ratio of true positives to the total predicted positives.

3. **Recall (Sensitivity):**

$$\text{Recall}=\frac{(TP)}{(TP)+(FN)}$$

Recall measures how well the model identifies actual positive cases.

4. **F1-Score:**

$$\text{F1-Score}=2 \text{ x } \frac{\text{Precision x Recall}}{\text{Precision + Recall}}$$

METHODS

1. LSTM

Long Short-Term Memory (LSTM) achieved an impressive accuracy of 98.15% in predicting workloads categorized as high, medium, or low. The precision, recall, and F1-score are all recorded at 97.18%, demonstrating the model's effectiveness in handling sequential data and making accurate predictions. The confusion matrix provides additional insights, where the x-axis represents the predicted labels, and the y-axis shows the true labels. The diagonal color-coded boxes indicate correct predictions, with only a minimal number of misclassifications appearing

10.48047/jocaaa.2024.33.08.107

in off-diagonal boxes. This performance validates LSTM's ability to capture workload trends over time, making it a crucial component for proactive resource allocation in cloud data centres.

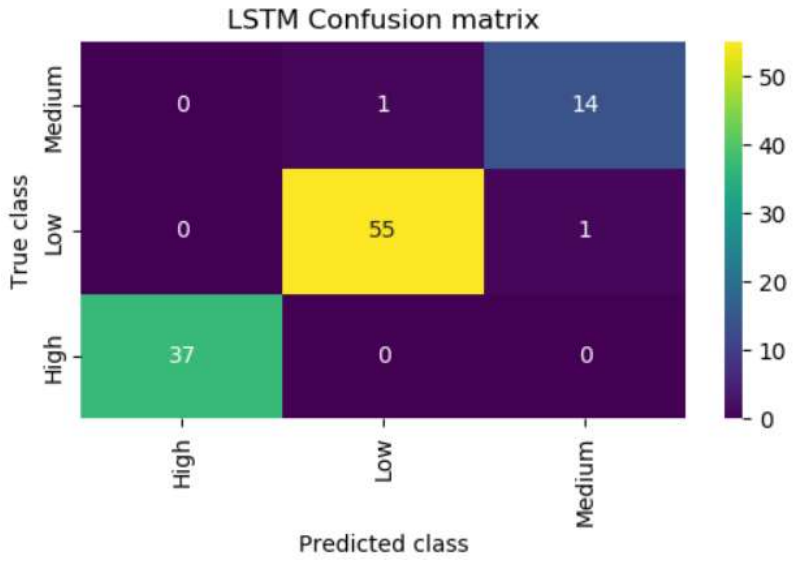


Figure 6: LSTM Confusion Matrix

Algorithm Name	Accuracy	Precision	Recall	FMeasure
LSTM	98.14814814814815	97.18253968253968	97.18253968253968	97.18253968253968

2. Reinforcement Learning

Reinforcement Learning (RL) was employed for dynamic scheduling and demonstrated an accuracy of 96.30%. The algorithm recorded 104 rewards and 5 penalties during the training process, reflecting its efficiency in learning optimal scheduling strategies. With a precision of 97.39%, recall of 96.19%, and an F1-score of 96.72%, RL excelled in adapting to fluctuating workload demands. The confusion matrix for RL shows the x-axis representing predicted resource VM names and the y-axis denoting true resource names. Correct predictions are displayed in diagonal boxes, with only a few misclassifications in off-diagonal boxes. RL's performance highlights its robustness in managing complex resource scheduling tasks.

10.48047/jocaaa.2024.33.08.107

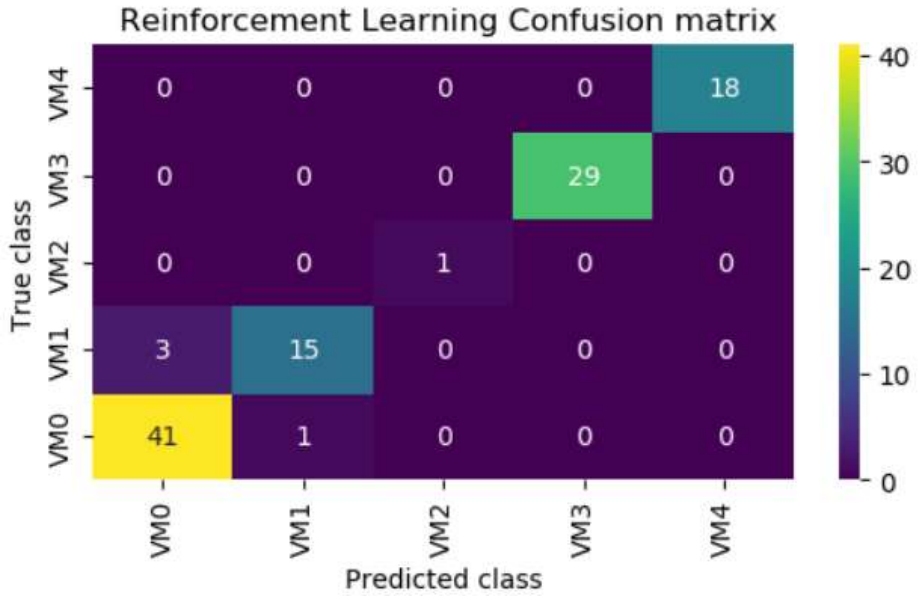


Figure 7: RL Confusion Matrix

Algorithm Name	Accuracy	Precision	Recall	FMeasure
RL	96.29629629629629	97.38636363636364	96.19047619047618	96.71682626538988

3. XGBoost

XGBoost was utilized for cooling system optimization, achieving an exceptional accuracy of 99.07%. The model's precision, recall, and F1-score were recorded at 99.40%, 98.02%, and 98.69%, respectively, underlining its reliability. The classification graph showcases predicted cooling labels (x-axis) versus true labels (y-axis), with diagonal boxes indicating correct predictions. Misclassifications are minimal, as represented by blue boxes in the graph. XGBoost's ability to predict cooling requirements accurately ensures optimal energy usage, reducing operational costs and enhancing sustainability in cloud data centres. Its high-performance metrics demonstrate the algorithm's critical role in maintaining efficient cooling systems.

Algorithm Name	Accuracy	Precision	Recall	FMeasure
XGBoost	99.06976744186046	99.3993993993994	98.02054154995332	98.68672290811128

10.48047/jocaaa.2024.33.08.107

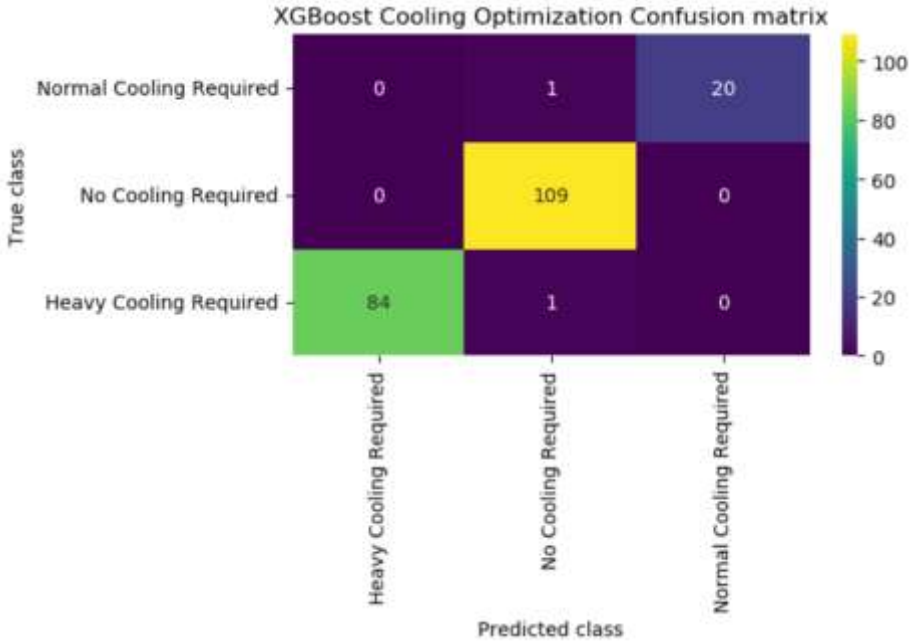


Figure 8: XGBoost Contusion Matrix

VII. RESULTS

The performance of the proposed framework was evaluated using three primary algorithms: LSTM for workload prediction, Reinforcement Learning (RL) for scheduling, and XGBoost for cooling optimization. The results showcase the effectiveness of these algorithms across various metrics, including accuracy, precision, recall, and F1-score.

Comparison Table

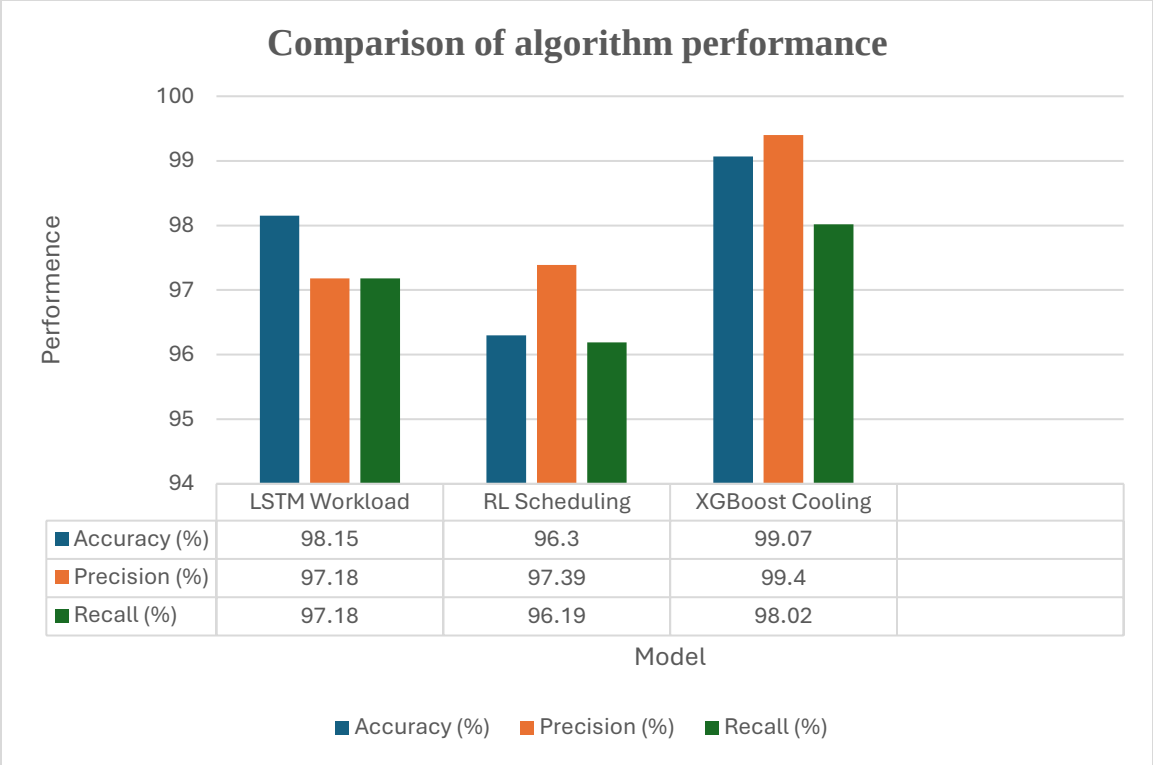
The table below highlights the performance metrics for each algorithm:

Algorithm Name	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
LSTM Workload	98.15	97.18	97.18	97.18
RL Scheduling	96.30	97.39	96.19	96.72
XGBoost Cooling	99.07	99.40	98.02	98.69

Graph Representation

The bar graph below visualizes the comparison of algorithm performance, with the x-axis representing the algorithms and the y-axis indicating the respective metrics. Different colored bars illustrate accuracy, precision, recall, and F1-score, providing a clear comparison of each algorithm's effectiveness.

10.48047/jocaaa.2024.33.08.107



Prediction Results

The framework accurately predicted scheduled resources, workloads, and cooling optimization requirements for various test data inputs. Example predictions include:

- **Test Data 1**: Predicted Scheduled Resource = VM4, Predicted Workload = High, Cooling Optimization = Heavy Cooling Required.
- **Test Data 2**: Predicted Scheduled Resource = VM0, Predicted Workload = Low, Cooling Optimization = No Cooling Required.
- **Test Data 3**: Predicted Scheduled Resource = VM1, Predicted Workload = Medium, Cooling Optimization = Normal Cooling Required.

These predictions align closely with actual outcomes, demonstrating the robustness of the framework.

Overall, the results confirm the high accuracy and reliability of the proposed methods. LSTM excelled in workload prediction with an accuracy of 98.15%, RL demonstrated strong scheduling capabilities with 96.30% accuracy, and XGBoost achieved the highest accuracy of 99.07% in cooling optimization. The predictions ensure optimal resource allocation and energy efficiency in cloud data centres, validating the framework's practical applicability.

10.48047/jocaaa.2024.33.08.107

Discussion

The results of the proposed framework demonstrate the effectiveness of integrating machine learning (ML) and deep learning (DL) techniques for optimizing power consumption in cloud data centres. Each algorithm—LSTM, Reinforcement Learning (RL), and XGBoost—has proven its value in specific operational areas, including workload prediction, resource scheduling, and cooling optimization.

LSTM's ability to analyze sequential data has enabled precise workload predictions with an accuracy of 98.15%, making it suitable for dynamic environments. RL demonstrated robust resource scheduling capabilities, achieving 96.30% accuracy by dynamically adapting to workload fluctuations and minimizing penalties. Meanwhile, XGBoost stood out in cooling optimization, achieving an accuracy of 99.07%, significantly improving energy efficiency and reducing operational costs.

The visualized comparison between algorithms highlights their unique strengths, with XGBoost performing exceptionally in precision, recall, and F1-score. The predictions for scheduled resources, workloads, and cooling requirements align closely with actual data, validating the framework's reliability.

However, challenges such as computational overhead for DL models like LSTM and scalability of RL in large-scale systems must be addressed for real-time applications. Future work should focus on optimizing these algorithms and testing the framework in diverse operational environments. Overall, the framework establishes a foundation for sustainable and efficient cloud infrastructure.

Conclusion

This study presents an innovative framework that integrates machine learning (ML) and deep learning (DL) techniques to optimize power consumption in cloud data centres. By leveraging LSTM for workload prediction, Reinforcement Learning (RL) for resource scheduling, and XGBoost for cooling optimization, the framework achieves exceptional performance metrics, including accuracies of 98.15%, 96.30%, and 99.07%, respectively. These results validate the effectiveness of the proposed methods in addressing key challenges such as energy efficiency, resource allocation, and cooling system optimization.

The predictions for workload intensities, resource schedules, and cooling requirements demonstrate the framework's robustness in real-world scenarios. Visualizations further enhance understanding, showing minimal misclassifications and reliable outcomes across all algorithms.

Despite its high accuracy, future work should address scalability and computational demands for broader deployment. Overall, this framework sets a strong foundation for sustainable cloud operations, paving the way for more energy-efficient and environmentally friendly IT infrastructures.

10.48047/jocaaa.2024.33.08.107

References

1. Beloglazov, A., Buyya, R., Lee, Y. C., & Zomaya, A. (2012). Energy-efficient resource allocation algorithms for cloud computing environments. *Future Generation Computer Systems*, 28(5), 755-768.
2. Shehabi, A., Smith, S., Masanet, E., & Koomey, J. (2016). Data centre growth in the global information technology sector. *Environmental Research Letters*, 11(7), 074022.
3. Islam, S., Lee, K., Fekete, A., & Liu, A. (2012). How a consumer cloud data centre benefits from energy-efficient resource allocation. *IEEE Transactions on Cloud Computing*, 1(2), 34-47.
4. Mishra, A., & Sahoo, B. (2018). Performance analysis of LSTM-based models for time-series predictions. *Journal of Computational Science*, 28, 16-24.
5. Verma, A., Ahuja, P., & Neogi, A. (2009). pMapper: Power and migration cost-aware application placement in virtualized systems. *Middleware*, 243-264.
6. Xu, J., Rao, L., & Bu, X. (2017). Dynamic virtual machine management for cloud computing. *IEEE Transactions on Parallel and Distributed Systems*, 28(1), 32-45.
7. Liu, X., Bash, C., & Patel, C. D. (2012). Thermal-aware scheduling in cloud computing. *Journal of Cloud Computing*, 1(1), 1-14.
8. Gao, J., & Liu, W. (2021). Anomaly detection in data centres using Isolation Forest. *ACM Transactions on Intelligent Systems and Technology*, 12(5), 47.
9. Bash, C., Patel, C., & Sharma, R. (2009). Dynamic thermal management of data centres using machine learning. *Proceedings of the IEEE ITherm Conference*, 8(3), 1-8.
10. Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785-794.
11. Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation*, 9(8), 1735-1780.
12. Sutton, R. S., & Barto, A. G. (2018). *Reinforcement Learning: An Introduction*. MIT Press.
13. Toomey, J. G. (2011). Growth in data centre electricity use 2005 to 2010. *Analytics Press*.
14. Abadi, M., et al. (2016). TensorFlow: Large-scale machine learning on heterogeneous systems. *Software Framework Documentation*.
15. LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436-444.

10.48047/jocaaa.2024.33.08.107

16. Wang, L., Chen, C., & Wu, W. (2015). Energy-efficient data centre management using reinforcement learning. *IEEE Transactions on Cloud Computing*, 4(3), 265-277.

17. Kingma, D. P., & Ba, J. (2015). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

18. Kaur, S., & Kaur, N. (2020). Cloud cooling optimization using advanced machine learning techniques. *Journal of Information Technology Research*, 13(3), 28-39.

19. Patel, R., & Sharma, A. (2022). Sustainable data centres through workload optimization. *Journal of Sustainable Computing*, 34, 101045.

20. Hinton, G. E., Osindero, S., & Teh, Y. W. (2006). A fast learning algorithm for deep belief nets. *Neural Computation*, 18(7), 1527-1554.

21. Zhang, J., & Zheng, W. (2021). Real-time workload prediction for cloud data centres using deep learning. *Journal of Cloud Computing*, 10(1), 15.

22. Gupta, A., & Pal, S. (2021). Comparative study of supervised machine learning techniques for cloud optimization. *IEEE Transactions on Cloud Computing*, 9(1), 30-39.

23. Kwok, Y., & Ahmad, I. (1999). Static scheduling algorithms for allocating directed task graphs to multiprocessors. *ACM Computing Surveys*, 31(4), 406-471.

24. Li, J., Hu, X., & Guo, H. (2020). Machine learning-based dynamic VM management for energy-efficient cloud computing. *IEEE Transactions on Network and Service Management*, 17(3), 1572-1582.

25. Huang, Y., & Li, X. (2020). Anomaly detection in cloud systems: A survey. *Journal of Cloud Computing*, 9(1), 1-28.