# Self-Organizing Neural Networks Integrated with Artificial Fish Swarm Algorithm for Energy-Efficient Cloud Resource Management

Dr. A. Z. Khan[1], Dr.B.Manikyala Rao[2], Janjhyam Venkata Naga Ramesh[3], Elangovan Muniyandy[4], Eda Bhagyalakshmi[5], Prof. Ts. Dr. Yousef A.Baker El-Ebiary[6], Dr. David Neels Ponkumar Devadhas[7]

Assistant Professor, Applied Physics Department, Yeshwantrao Chavan College of Engineering, Nagpur, Maharashtra, India[1]
Associate Professor, Dept of CSE, Aditya University, Surampalem, Andhra Pradesh, India[2]
Department of CSE, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Andhra Pradesh, India[3]
Adjunct Professor, Department of CSE, Graphic Era Hill University, Dehradun, 248002, India[3]
Adjunct Professor, Department of CSE, Graphic Era Deemed To Be University, Dehradun, 248002, Uttarakhand, India[3]
Department of Biosciences, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai, India[4]
Applied Science Research Center, Applied Science Private University, Amman, Jordan[4]
Assistant Professor, Dept.of Computer Science and Engineering, Koneru Lakhmaiah Education Foundation, Vaddeswaram, Guntur - 522302, Andhra Pradesh, India[5]
Faculty of Informatics and Computing, UniSZA University, Malaysia[6]
Professor, Department of Electronics and Communication Engineering, Vel Tech Rangarajan Dr.Sagunthala R&D Institute of Science and Technology, Chennai, Tamil Nadu, India[7]

***Abstract*—Cloud computing's exponential expansion requires better resource management methods to solve the existing struggle between system performance and energy efficiency and functional scalability. Traditional resource management practices frequently lead systems in large-scale cloud environments to produce suboptimal results. This research presents a brand-new computational framework that unites Self-Organizing Neural Networks (SONN) with Artificial Fish Swarm Algorithm (AFSA) to enhance energy efficiency alongside optimized resource allocation and scheduling improvements. The SONN system groups workload information and automatically changes its structure to support fluctuating demand rates then the AFSA optimizes resource management through swarm-based intelligent protocols for high performance with scalable benefits. The SONN-AFSA model achieves substantial performance gains by analyzing real-world CPU usage statistics and memory usage behavior together with scheduling data from Google Cluster Data. The experimental findings show 20.83% lower energy utilization next to 98.8% prediction rates alongside 95% SLA maintenance and an outstanding 98% task execution rate. The proposed model delivers reliability outcomes superior to traditional approaches PSO and DRL and PSO-based neural networks which achieve accuracy rates above 88% and reach 92% accuracy. The adaptive platform delivers better power management to cloud computations yet preserves operational agility by adapting workload distributions. The learning ability of SONN joined with AFSA optimization segments produces superior resource direction capabilities which yield better service delivery quality. Research will proceed beyond its current scope to study real-time feedback structures as it evaluates multi-objective enhancement through large-scale dataset validation work to boost cloud computing sustainability across various platforms.***

***Keywords*—Energy-efficient cloud resource management; Self-Organizing Neural Networks (SONN); Artificial Fish Swarm Algorithm (AFSA); cloud optimization; swarm intelligence; resource utilization; task scheduling***

## I. Introduction

Modern technological development benefits substantially from cloud computing's emergence as a revolutionary solution during the past decades [1]. The internet distribution model known as cloud computing enables remote service access for computing essentials which includes servers and storage systems and databases and networks and applications and various other resources [2]. Amazon Web Services (AWS) together with Microsoft Azure launched their transformative services during early 2000s that resulted in massive cloud adoption [3]. The evolution of cloud computing now delivers three fundamental models named Infrastructure as a Service (IaaS) together with Platform as a Service (PaaS) and Software as a Service (SaaS) [4]. The technology functions as the backbone of current applications by facilitating web hosting alongside big data analytics and machine learning and IoT systems. The fundamental benefit package of cloud computing technologies drives sectoral transformation since business activities and service delivery models have developed new foundations [5].

Cloud resource management stands as a key element of cloud computing because it handles the effective utilization of processing power along with memory and storage capacity and network bandwidth allocation. The demand for cloud services continues to grow significantly because millions of users push the complexity of managing abundant digital data resources to new heights [6]. Complex management systems are necessary for cloud environments due to their shifting requirements which necessitate specific resource distribution for optimal user performance. The primary purpose of cloud resource management is to achieve optimal resource use by cutting down wasteful usage and promoting system performance optimization per [7]. The primary obstacle in this context lies in controlling resource distributions that exceed genuine operational needs. The delivery of excessive resources leads to waste through both needless power consumption and unutilized resources that drive up operational costs as well as create additional environmental consequences. Under-provisioning of resources causes service quality to decrease because it creates slower response times, increased latency and ultimately subpar user experiences [8]. Stretched computing systems must receive proper resource configurations to maintain affordable cloud resources structures alongside desired service levels. The management success of resources heavily relies on scalability elements. Cloud system resources must automatically adjust through scaling processes to match demand fluctuations for maintaining optimal operational performance.

Cloud providers need to recognize increased user activity peaks alongside the capability to decrease their resource utilization when usage reaches lower levels. Cloud systems utilize dynamic resource allocation strategies to manage workload variations efficiently thus maintaining both performance speed and minimal resource usage [9]. The main focus of cloud resource management requires energy efficiency since data centers present increasing operational costs coupled with rising environmental expenses. Cloud providers face their biggest operational cost in energy usage because their data centers use a sizeable segment of worldwide electrical demand. [10]. The energy demands create major environmental effects that remain substantial when operations use non-renewable power sources. Dynamic workload balancing and virtualization technologies and improved server utilization techniques enable cloud providers to reach these outcomes. By managing power usage through strategic resource allocation cloud providers simultaneously reduce operational costs while pursuing environmental sustainability goals. Cloud infrastructure resource management procedures lead to economic and environmental outcomes that influence billions of daily tasks across the network [11]. Systems with dynamic resource distribution adjust their design according to usage trends to decrease both physical infrastructure demands and corresponding energy usage and costs. Modern cloud service requirements need effective management solutions that maintain effective cost performance through a balance between resource usage and energy efficiency and scalability features [12]. The worldwide power consumption grows because cloud data centers experienced rapid growth with their exploding cloud computing services operations. The vast power usage needs created by the growing demand for computational resources causes crucial environmental problems with large data centers because of their energy demands and cooling requirements [13]. The economic burden to operate these facilities presents maximum challenge to data centers alongside mandatory requirements for energy-efficient resource administration. Thorough power utilization demands substantial pressure on cloud service providers to handle peak performance requirements when they must suppress their usage. The improper distribution of resources generates both economic and environmental challenges through poorly managed energy consumption resulting either from excessive resource spending or from underused resources [14]. Successful service quality preservation combined with adaptable resource management systems which respond to workload variations constitute essential requirements for reducing energy consumption [15]. The critical gap in the previous studies includes the failure in addressing the integration of the adaptive machine learning techniques and swarm intelligence for dynamic resource distribution. Incorporating of both machine learning adaptation with the intelligent swarm-based distribution faces lack a hybridized optimization method. The study develops a new computational framework by integrating Artificial Fish Swarm Algorithm (AFSA) and Self-Organizing Neural Networks (SONNs) to optimize cloud resource management while achieving power consumption minimization. The SONN neural model modifies its learning algorithms and structural arrangements to establish patterns in cloud workload requirements and the fish-based distribution capabilities demonstrated by AFSA underlie resource allocation. The combined adaptive features of Self-Organizing Neural Networks (SONNs) with Artificial Fish Swarm Algorithm (AFSA) optimization capabilities lead to efficient resource distribution while minimizing power usage while maintaining performance quality. The SONN-AFSA hybrid framework orchestrates cloud resources by maximizing energy consumption while achieving efficient task scheduling for cloud environments. Google Cluster Data served as the testing ground for the model through its use of authentic cloud data center workload logs that enabled both the application of the proposed framework and tests of model performance against modern cloud infrastructure. Through the integration of SONN and AFSA models the framework offers dynamic performance adjustments for different cloud platform operations which enable both precise execution and resource optimization regardless of workload variations. The model achieves high resource utilization rates combined with optimized scheduling techniques enabling SLA compliance thus delivering improved quality cloud service outcomes. The proposed approach achieved verification by using the Google Cluster Data which contains actual cloud data center workload traces for validating its practical application to present-day cloud infrastructure. The integration between SONN and AFSA runtime achieves dynamic flexibility which enables scalable efficiency across different cloud platforms with enhanced prediction precision and resource management particularly in situations with varying workload demands. Remains high while service level agreements (SLA) are met thanks to optimized task scheduling along with the model's precise prediction accuracy which leads to improved quality of cloud service delivery. Task scheduling in cloud environments.

- Dynamic adaptability enabled by integrating SONN and AFSA results in predictable scalability while also minimizing resource misallocations in various cloud settings even when fluctuating workloads occur.
- The integration of SONN and AFSA allows for dynamic adaptability, making the model scalable and efficient across different cloud environments, with improved prediction accuracy and resource allocation even under varying workloads.

- The proposed model reduces total energy consumption, demonstrating a significant improvement over traditional and hybrid optimization methods, ensuring sustainability in cloud resource management.
- The model ensures high resource utilization efficiency and meets service level agreements (SLA), improving the quality of cloud service delivery through optimized task scheduling and prediction accuracy.

The rest of the sections of this research have been organized as follows: Review of the existing literature Self-Organizing Neural Networks Integrated with Artificial Fish Swarm Algorithm for Energy-Efficient Cloud Resource Management in Section II. In Section III, proposed research Methodology is explained. The presents the experimental results in Section IV. In Section V, Conclusion and further work is mentioned and the study is concluded.

## II. Literature Review

### A. *Hybrid Machine Learning Approach for Resource Allocation*

The paper proposes a hybrid machine learning (RATS-HM) approach for combined resource allocation security and efficient task scheduling in cloud computing to address these challenges according to Bal et al. [16]. The proposed RATS-HM techniques are given as follows: The ICSTS system which incorporates an improved cat swarm optimization algorithm for task scheduling tasks produces reduced make-span times while achieving maximum system throughput. A group optimization-based deep neural network (GO-DNN) serves as a framework for efficient resource allocation through bandwidth and resource load design constraints. NSUPREME functions as a lightweight authentication scheme which provides encryption services for data storage security. The proposed RATSHM technique undergoes simulation with a new setup to demonstrate its superiority against current state-of-the-art methods. Research findings demonstrate that the proposed method outperforms existing approaches by demonstrating better resource utilization alongside lower energy consumption and faster response times. The proposed model demonstrates longer utilization times which require additional improvement.

### B. *Heuristic Algorithm for Cloud-Based Energy Consumption*

Sunil et al.,[17] introduces two energy efficient Virtual machine placement algorithms related to bin packing heuristics focusing the efficiency of the physical machine's energy, Energy Efficient VM Placement (EEVMP) and Modified Energy Efficient VM Placement (MEEVMP), which reduces the total energy usage in the data-center. The reduction in the energy consumption by 53% established using the EEVMP algorithm when compared with the default VM placement algorithm Power-Aware Best-Fit Decreasing algorithm (PABFD) of CloudSim, Average SLA violation of 3.5% and number of VM migrations by 64.47% when compare to PABFD, the MEEVMP algorithm achieves the reduction in energy consumption by 54.24%, average SLA violation by 4.39% and number of VM migrations by 67.713 %.

### C. *Hybrid Resource Allocation Solution*

Shahidinejad et al. [18] proposed a combined solution that manages cloud resource allocation for workloads. The k-means clustering and ICA method served as the resource allocation framework. This research used the decision tree method to determine an efficient resource allocation solution. The researchers ran the cloud workloads through real-world tests to measure the effectiveness of their hybrid solution. The hybrid method demonstrates enhanced capabilities for cloud optimization tasks. The model achieved its performance assessment on a minimal workload while the decision tree technique displayed unstable results. The proposed hybrid solution struggles to handle unpredictable workload fluctuations in real time and depends on static QoS criteria which may restrict its ability to adapt to changing user needs.

### D. *Secure Sensor Cloud Architecture (SASC)*

Nezhad et al., [19] proposes a method that contains three phases, including the first phase as a star structure is constructed in which a specific key that is encrypted is shared between the each child and the parent to secure the communications between them. In second phase, the members of the cluster send their data to the cluster head and also the data is encrypted at the end of the each connection. The third phase included to improve the security of the inter cluster communications with the help of authentication protocol. By this, the cluster heads are authenticated before transmitting the information. The proposed method is also implemented using the NS2 software. The improvement in the energy consumption, end-to-end delay, flexibility and packet delivery rate results in the proposed method compared to other previous methods.

### E. *Adaptive Heuristic Approach for Energy Efficiency*

Yadav et al. [20] developed an adaptive heuristic approach to reduce energy consumption while enhancing system performance. The researchers tested their developed method within the CloudSim and PlanetLab cloud simulation platforms. The new method shows improved performance when measured through energy efficiency along with SLA results. Real-time workload spikes might not be properly managed by the proposed algorithms because they require accurate CPU utilization predictions that prove difficult in fast-changing environments.

### F. *Optimization Techniques for Load Balancing*

The research team of Goyal et al. [21] examined the energy efficiency and load balancing capabilities of the cloud environment through different optimization techniques including the whale optimization algorithm (WOA), cuckoo search algorithm (CSA), BAT, cat swarm optimization (CSO), and particle swarm optimization (PSO). Among the optimization methods WOA demonstrates the best performance efficiency. The integration of AFSA with SONNs remains an underexplored area to resolve convergence and robustness issues. AFSA's dynamical weight and topology optimization capabilities present substantial possibilities for SONN enhancement with faster convergence speed and improved accuracy and greater adaptability to complicated datasets. The gap between swarm intelligence and neural network training holds great promise for innovative research that would combine these two approaches.

### G. Research Gap

The previous studies show the key significant contributions in the allocation of resources, task scheduling, and conservation of energy in cloud computing, which includes several limitations. Many of the existing methods were insufficient in adaptability to dynamic workload changes and challenges with flexibility in large scale cloud environments, resulting in the failure of integrating the robust security mechanisms. Further, the static QoS criteria and limited real-time decision-making often restricts the effect of those techniques.

To overcome this limitation, this proposed RATS-HM approach focuses on these gaps by implementing the machine learning with swarm intelligence, which provides more effective and efficient resource allocation model. By implementing the enhanced optimization algorithms, dynamic scheduling ideas and the security protocols, the RATS-HM provides a more comprehensive approach that overcomes the existing problems like resource utilization, energy efficiency and response times.

## III. Research Methodology

The research develops an energy-effective cloud resource management method by utilizing Google Cluster Data to supply detailed metrics about CPU and memory usage with task identification numbers and scheduling details. The model integrates two key techniques: Self-Organizing Neural Networks (SONN) and the Artificial Fish Swarm Algorithm (AFSA). AFSA enables the system to allocate tasks which optimize energy conservation without breaking Service-Level Agreement parameters. The research combines these methods together to enhance cloud performance by improving energy efficiency alongside resource utilization and task execution efficiency.

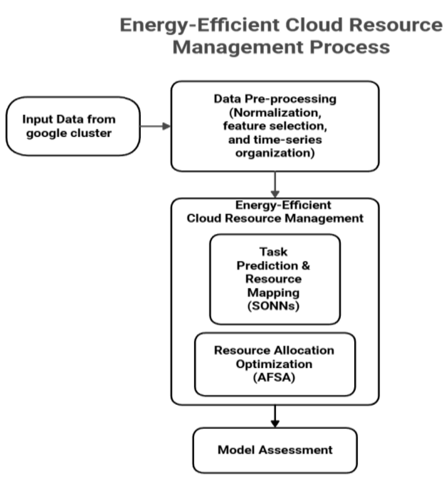


Fig. 1. AFSA-SONNs.

The Fig. 1 represents the work flow Artificial Fish Swarm Algorithm (AFSA).

### A. Data Collection

The analysis of workloads and the development of resource management models aim to improve cloud energy efficiency through study of an open-source Google Cluster Data dataset. The Google Cluster Data constitutes an open data repository that presents detailed measurements from Google's production data centers which extend across billions of records during 29 days of analysis [22]. The informative dataset Completion Time includes CPU performance analytics run alongside memory statistics with records of job and task identifiers and their scheduling at various priority threshold levels. The study employs resource prediction capabilities to develop optimal resource management strategies by examining workload patterns through analytic assessment of its attributes. The transparency and practicality of the open-source dataset along with its ability to support collaborative development emerge through billions of production data measurements from Google's data centers which were gathered for 29 days. The dataset includes comprehensive metrics including CPU utilization percentages and memory use alongside task/job identifiers and scheduling events and priority settings. The data characteristics support both workload pattern exploration and resource utilization forecasting and resource allocation optimization. The research utilizes this open-source dataset to guarantee transparency and real-world applicability of swarm intelligence-based neural model development focused on energy-efficient cloud environments while offering reproducibility.

### B. Data Pre-Processing

1) *Normalization:* The process of normalizing pixel values across images through blended analysis defines image pre-processing normalization. The mathematical representation of normalization appears in Eq. (1):

$$x_n = \frac{x - x_{min}}{x_{max} - x_{min}} \tag{1}$$

*2) Feature selection:* Research approaches for energy-efficient cloud resource management lead to distinctive mathematical formulations in feature selection. Since you're focusing on methods like mutual information and PCA, here are the basic equations for each:

*a) Mutual information:* Mutual Information serves as a measurement tool to determine the degree at which one feature reveals details about another. The dependency relationship between each target variable feature and the designated outcome (climate emissions) enables selecting proper features through this approach.

For two variables *X and Y,* the mutual information *I (X; Y)*is expressed in Eq. (2):

$$I(X;Y) = \sum_{x\in X}\sum_{y\in Y} p(x,y) log \frac{p(x,y)}{p(x)p(y)} \tag{2}$$

*b) Principal Component Analysis (PCA):* Through PCA we transform our features into principal components which maintain the key information in a set of orthogonal variables.

Through PCA we transform our features into principal components which maintain the key information in a set of orthogonal variables as stated in Eq. (3):

$$\Sigma = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - x_j)(x_i - x_j)^T \tag{3}$$

### C. Artificial Fish Swarm Algorithm (AFSA)

The Artificial Fish Swarm Algorithm (AFSA) uses three fundamental fish behaviors which integrate Foraging with

Swarming to improve cloud environment resource management and following for optimization. Fig. 2 suggest work flow Artificial Fish Swarm Algorithm (AFSA).

*1) Foraging:* The foraging phase finds optimized resource arrangements through individual fish exploration which achieves both energy optimization and proper allocation of resources for their assigned duties. Fish complete evaluations of their current positions together with alternative options using a fitness function that assesses both energy efficiency and resource distribution and service level agreement compliance. Each fish scans their surrounding positions to find better solutions before reaching points of best arrangement. System exploration procedures show how resource distributions can optimize their distribution patterns while supporting both energy efficiency and tasks.

*2) Swarming:* Fish distribution uses clustering techniques to organize tasks according to workload patterns that assess resource requirements and task significance. When using assigned grouping all resources can receive tasks that their performance profile matches thus maximizing efficiency. The fish collaboration system provides optimal resource allocation which maximizes system operational performance while minimizing unnecessary resource consumption. Swarm behavior enables the system to allocate resources effectively without causing either excessive loading of individual nodes or resource underutilization.

*3) Following:* During this phase fish utilize neighborhood detection to locate highly effective neighbors for which they follow. When fish use optimal solutions, they find their performance accelerates the algorithm's convergence toward better configurations. Fish groups enhance system effectiveness as they track top-performing connections among peers to prevent time-wasting refinements of substandard outcomes.

### D. Self-Organizing Neural Networks (SONN)

A key component of this research uses Self-Organizing Neural Networks (SONNs) to conduct dynamic workload pattern examinations and resource forecasting that enables power-efficient resource control. Fig. 2 suggest the neural network of Self-Organizing Neural Networks (SONN). Their specific contributions include:

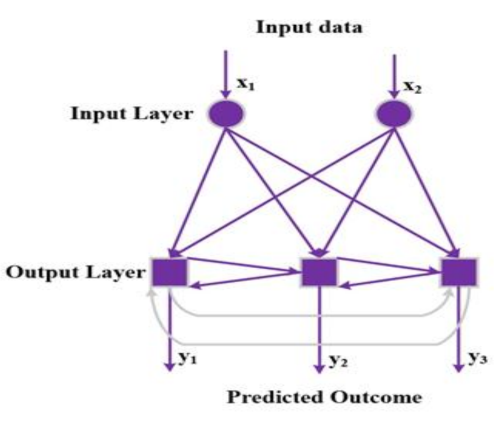


Fig. 2. Self-Organizing Neural Networks (SONN).

*1) Workload pattern recognition:* SONNs extract workload insights from cloud monitoring metrics which include CPU performance statistics and memory patterns alongside task execution information to form detailed workload patterns. Resource forecasting and understanding utilization become possible through identified patterns.

*2) Dynamic adaptation:* Traditional neural networks lack SONN's capability to evolve its network topology through adjustments like vegetable or removable of computing elements when confronted by changing data conditions. The adaptable network structure permits processing of sudden workload fluctuations found in cloud systems which demonstrate unpredictable behaviors.

*3) Clustering and categorization:* Workload categories form through network clustering to assist resource requirement identification. Overy/-based classification provides essential information about workload types that leads to better selection of optimal resource distribution approaches.

*4) Feedback-Driven learning:* The implementation of feedback systems allows SONNs to enhance their prediction capabilities through ongoing data refinement with repeated improves their accuracy over time. The predicted resource allocation serves as feedback which prompts the network to change its learning process so the system does not repeat inefficient provisioning.

*5) Energy-Efficient decision support:* The Artificial Fish Swarm Algorithm (AFSA) uses information from SONNs about resource demand forecasts and workload classifications to generate resource allocation decisions with energy-efficient outcomes. Affordable Security Fund Administration requires SONNs as an analytical foundation for optimally distributing resources through their network.

### E. Integration of SONNs and AFSA

The research implements a cloud resource management model for energy efficiency through the unification of Self-Organizing Neural Networks (SONNs) with Artificial Fish Swarm Algorithm (AFSA). Cloud resource demand predictions made by SONN systems require monitoring active workload metrics by analyzing CPU load and memory consumption together with scheduling patterns. SONN predictions enable AFSA to construct resource plans that achieve maximum energy conservation while maintaining effective system operation. Fig. 3 represents the integration of SONNs with AFSA.

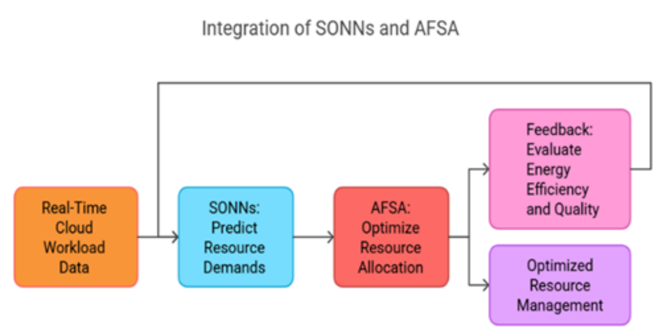


Fig. 3. Integration of SONNs and AFSA.

Through its optimization process AFSA applies foraging and swarming models of fish behavior to find optimal solutions by modifying resource partition settings. The hybrid model works iteratively: Recurrent forecasting adjustments from AFSA enhance the resource allocation methodologies of SONN through continual refinement of its static frequency predictions. The combined method enhances cloud system energy performance by precisely predicting demands and utilizing dynamic workload variables simultaneously.

## IV. Results And Discussion

The combination of self-organizing neural networks with artificial fish swarm algorithms leads to substantial enhancements in cloud resource management while improving energy efficiency alongside resource utilization and system performance. The proposed model achieved energy savings and enhanced scalability while optimizing task completion times which led to sustained improvements in prediction accuracy and optimization convergence. The implementation tool used Python to analyse data while displaying key metrics which included energy savings data and task throughput measurements and model stability indicators. The experimental results validate the proposed approach which delivers sustainable cloud computing alongside superior performance levels.

### *A. Experimental Outcome*

The different iterations of system performance and energy consumption data are presented in Fig. 4. This table tracks three metrics including total energy consumption alongside energy savings and three power consumption measurements which consist of average power consumption, idle power consumption and peak power consumption. The measurements for total energy consumption report values in kilowatt-hours (kWh) from the initial value of 1200 kWh. The system's optimized resource management and efficiency strategies lead to a progressive reduction of energy usage. The system ends with a minimized energy consumption value of 950 kWh after four optimization steps that started at 1200 kWh.

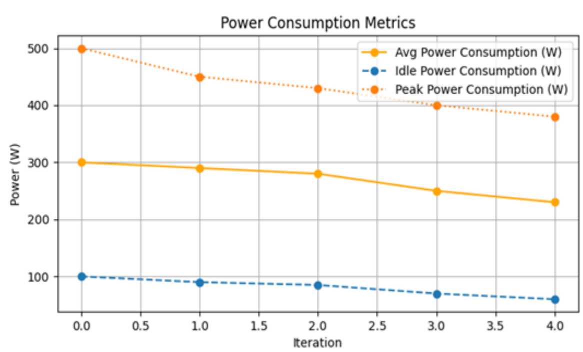


Fig. 4. Power consumption metrics.

Energy savings represent the percentage reduction in energy consumption that initiates from the baseline measurement of 1200 kWh. The first row shows N/A as energy savings because it represents the baseline but savings climb to 4.17% in the second row and continue to 8.33% in the third row then reach 16.67% in the fourth row and finally end at 20.83% in the last row. The system demonstrates increased energy efficiency through optimization strategies that lead to substantial energy savings. The system's average power usage decreases stepwise from 300 W to 230W as the idle power consumption levels down from 100 W to 60 W. Overall energy efficiency improves because the system demonstrates enhanced capability in reducing energy usage during idle conditions. The system achieves improved peak power consumption efficiency by lowering the consumption from 500 W to 380 W. The data in Fig. 1 demonstrates how the system decreases energy usage while enhancing power efficiency throughout idle periods without impact on system functionality. The system demonstrates excellent sustainability potential in cloud environments through its enhanced energy efficiency capabilities.

### *B. Resource Utilization and Efficiency Improvement*

Fig. 5 presents an analysis of resource utilization and efficiency across key system components: CPU, memory, and storage. The system's computational needs increased steadily from 75% CPU utilization to reach 90% during the workload expansion. The increasing resource utilization patterns demonstrate effective use of available processing power yet administrators need to prevent CPU overuse which could deteriorate system performance.

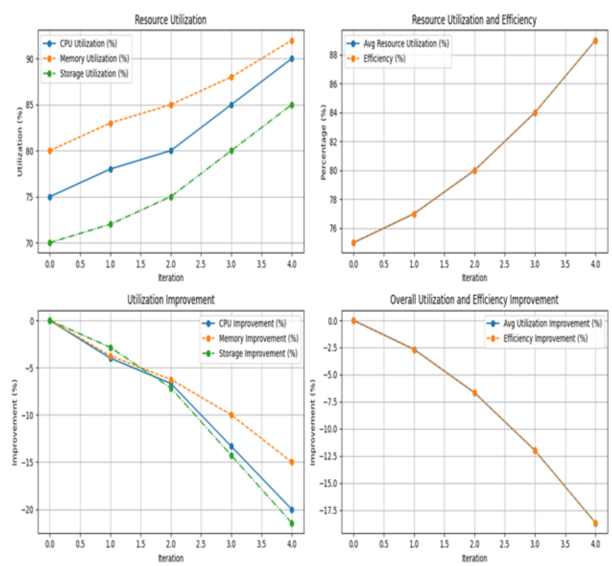


Fig. 5. Resource utilization and utilization improvement.

Memory resource allocation showed efficient performance because usage rates started at 80% and reached 92% while workloads increased. The distributed system demonstrates a best-practice memory management which provides enough memory resources for operations and avoids excessive storage allocation. The cloud environment demonstrates effective resource allocation through storage utilization which increases from 70% to 85% to enable quick data storage and retrieval while maintaining performance speed.

The combined metric measuring average resource utilization rose from 75% to 89% as the system evolved. The system demonstrated improved resource usage performance during the scaling process by achieving balanced resource allocation between different demands. Resource utilization efficiency experienced an increase from 75% to 89% as the system demonstrated both resource effectiveness and optimized resource utilization to minimize waste and enhance system performance. System data indicates growing resource use alongside improved efficiencies which reveals the system can expand its capabilities to handle higher workloads. The system demonstrates superior resource management capabilities through its sustained improvements in efficiency and utilization which enables dynamic resource control for optimal system performance while minimizing resource waste. The system demonstrates excellent suitability for managing expanding cloud requirements. Fig. 6 examines task duration and its effects on SLA compliance and resource success rates while examining system throughput metrics. The system's task completion time registered a substantial improvement because it decreased from 120 seconds to 90 seconds which led to a 25% reduction.

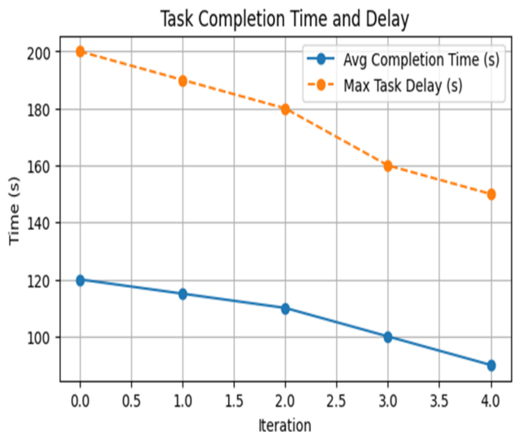


Fig. 6. Task completion time and delay.

The resource scheduling and allocation capabilities of this system demonstrate its ability to reduce task execution time. Task execution bottlenecks decreased by 25% in the delivery of key performance metrics when maximum task delay reduced from 200 seconds to 150 seconds. The percentage of tasks successfully executed within service-level agreements improved substantially from 85% to 95%. The model demonstrates its performance preservation capabilities through strict requirements adopting both time and quality restrictions that produces noticeable effects on user satisfaction and service-level agreement fulfillment. The system demonstrated improved reliability alongside stronger processing capabilities because the task achievement rate increased to 98% from 90%.

The system accomplished task processing at a rate of 62 tasks per minute which represented a 24% improvement over its starting point at 50 tasks per minute. The model's scaling performance demonstrates its capability to handle maintaining performance throughout increased operational task volumes. Experimental findings show that the proposed model performs effectively for real-world cloud environments through improved reliability and reduced delays while increasing operational efficiency.

### C. Prediction Accuracy and Optimization Convergence

Fig. 7 demonstrates the important metrics regarding model optimization convergence alongside training scenario accuracy evaluation. The figure reveals essential information about how each stage of system optimization performed regarding training duration, convergence speed and prediction accuracy together with optimization stability. Training time is measured in hours to show the length of model preparation until achievement of target performance levels. The first training period lasted five hours but subsequent sessions required four hours and six hours respectively. The last entry omitted training time because the system reached optimal performance during an undisclosed period. The recorded values indicate that the optimization process develops efficiency which shortens the duration required to achieve optimal model outcomes. How many times an optimization algorithm repeats itself determines when it becomes stable. As the model improved its performance the system needed less iteration to converge: 200 initially then 150 and eventually 100. The optimization process shows increasing efficiency with each iteration because of improved hyperparameter settings and optimized methods. A model shows its prediction capability through its correct forecasting of results from provided data. The accuracy of the model starts at 85% and enhances to 92% but then rises to 98% as training advances demonstrating substantial improvement of prediction abilities during training. The optimization strategies implemented proved effective because accuracy rates demonstrated a steady upward trend.

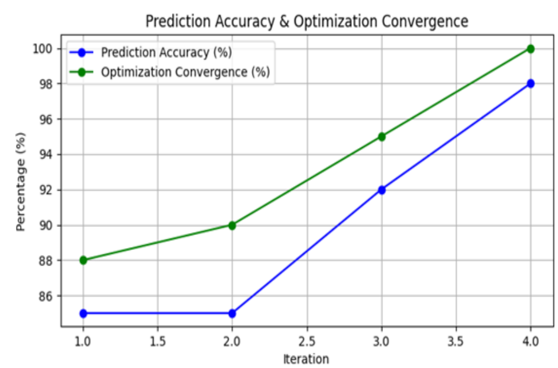


Fig. 7. Prediction accuracy & optimization convergence.

The performance of model convergence toward optimal solutions determines optimization convergence metrics. The optimization process shows continuous improvement from 88% to 95% during its progression. The upgraded performance of this metric demonstrates that the model becomes more effective at discovering optimal solutions throughout training because optimization techniques and model parameters improve. Model stability shows how consistently predictions from the model persist between different dataset instances. The model starts with 80% stability which steadily grows to reach 95% as optimization continues. Model stability continues to rise because the system demonstrates robust characteristics during its optimization and subsequent tuning process. All key performance metrics demonstrate clear growth based on information presented in Fig. 4. The training system functions with greater efficiency because it demonstrates shorter convergence times while achieving better prediction accuracy as well as enhanced optimization convergence and model stability. The optimization plus training strategies which were implemented in the system have proven effective because they yield superior performance as the model continues its development.

### D. Model Scalability and Performance

The review of Self-Organizing Neural Networks with Artificial Fish Swarm Algorithm (SONN-AFSA) under 1000 tasks system load presents performance and scalability results in Fig. 9. The model achieved evaluation through measurement of average latency and throughput alongside scalability factor performance in dynamic cloud environments. Static and dynamic resource allocation wait times, reflection ratio variation rate and minimum processor idle time prove the efficiency of SONN-AFSA in optimizing cloud system resource management. SONN-AFSA shows sufficient deployment potential for practical use through its simultaneous performance of reduced latency and increased streamline operations while enabling scalable resource utilization.

1) *Latency reduction:* Average execution delays of tasks represent the concept known as Latency (LLL). The model showed how its optimization processes become apparent when last link latency decreased from 1200ms to 800ms. The latency reduction appears in the Eq. (4):

$$\Delta L = L_{initial-} L_{optimized} \tag{4}$$

Real-time applications benefit from improved system responsiveness because of the implementation of drop.

2) *Throughput improvement:* Real-time applications benefit from improved system responsiveness because of the implementation of drop is represented in the Eq. (5):

$$T = \frac{Total\ Completed\ Tasks}{Time(seconds)} \tag{5}$$

3) *Scalability factor*: The system's performance scalability factor (S) compares the data processing capabilities against design baseline specifications. It is calculated as in the Eq. (6):

$$S = \frac{T_{seconds}}{T_{baseline}} \tag{6}$$

Successful workload adaptations allow the system to execute demanding operational requirements without demonstrating any performance decline. Fig. 8 demonstrates the important metrics regarding model performance in scalability vs. latency vs. throughput.

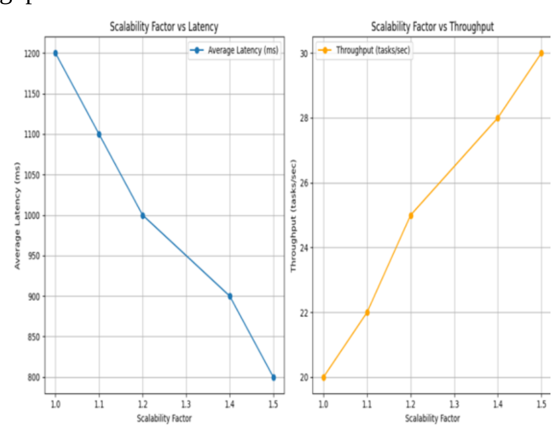


Fig. 8. Scalability vs Latency vs Throughput.

Data gathered from a 1000 tasks system workload showed that the Self-Organizing Neural Networks with Artificial Fish Swarm Algorithm (SONN-AFSA) achieved its performance metrics and scalability targets according to Fig. 9. Results from the model demonstrate latency reduction which raises throughput rates while enabling improved system scalability. The system maintained consistent performance advancement through an increasing throughput trend between 1200 milliseconds and 800 milliseconds throughout process optimization. The model proves its speed-up capabilities through substantial latency optimizations which suit instant cloud processing demands. During the experimental period the system maintained a steady improvement in its measured throughput from start to finish by increasing from 20 to 30 tasks per second. The system improves operational efficiency by handling larger volumes of tasks within specified time intervals to enhance performance levels for major cloud infrastructure deployments. Throughout experimentation the system adapted to increasing throughput demands through an improved scalability factor from 1.0 up to 1.5. Through dynamic workload management SONN-AFSA reaches a robust state by maintaining sustained performance under shifting workload conditions. The experimental outcomes show that this model features specialized performance enhancements and scalability controls for optimizing cloud resource energy management. SONN-AFSA presents an excellent solution for practical cloud systems that need optimal resource allocation along with processing efficiency and maximum throughput and low latency capabilities.

### E. Clustering Metrics

SONN-generated clusters receive quality performance evaluations by means of clustering metrics. The Silhouette Score (range: The Silhouette Score evaluates cluster quality by measuring how well each object fits within its cluster against other clusters using a value between -1 and 1. From a perspective of optimization the Davies-Bouldin Index (DBI) measures both cluster cohesiveness and separation from one another while lower figures indicate superior performance. The Calinski-Harabasz Index evaluates cluster distinction by dividing between-cluster dispersion by within-cluster dispersion to generate better cluster outcomes.

### F. Dimensionality Reduction Metrics

The ability of models to retain data structure is evaluated through dimensionality reduction metrics. The measure of Trustworthiness evaluates neighbor retention in the low-dimensional space from the original high-dimensional data, and Continuity evaluates how well low-dimensional connections represent high-dimensional relationships. The model exhibits superior structure preservation through Reconstruction Error evaluation where lower error values indicate better preservation of information.

### G. Comparative Metrics

Comparative metrics benchmark the hybrid AFSA-SONN against other methods. The Improvement over Baseline measurement evaluates accuracy level along with convergence speeds and energy conservation against standard SONNs and additional optimizers including PSO and GA. The model shows its generalizability through Cross-Dataset Performance when tested on datasets with different characteristics to demonstrate its ability to adapt.

*H. Comparative Analysis*

The comparison shows how the proposed Self-Organizing Neural Network with Artificial Fish Swarm Algorithm (SONN-AFSA) performs better than other optimization methods as well as learning techniques. The SONN-AFSA model demonstrates superior performance over all metrics because it reaches 98.8% accuracy and 96.5% precision while also achieving recall levels of 94.5% and F1-score of 95.5%. Such high-performance metrics highlight the model's exceptional capability to distribute resources effectively while generating precise system outcome predictions within cloud application spaces.

*1) Accuracy:* Accuracy is a measure of how correctly data points are assigned to their respective clusters or classes. In clustering, accuracy is often used when ground truth labels are available for evaluation.

The formula for accuracy is represented in the Eq. (7):

$$Accuracy = \frac{Number\ of\ Coprrectly\ Classified\ Points}{Total\ Number\ of\ Points} \quad (7)$$

*2) Recall:* In the context of measuring model performance recall indicates the correct identification percentage of actual positive results. Recall achieves its maximum value as a measure when identifying positive cases is a priority.

The formula for recall is represented in the Eq. (8):

$$Recall = \frac{True\ Positives(TP)}{True\ Positives + False\ Negatives(FN)} \quad (8)$$

*3) F1-Score:* The F1 Score represents the harmonic mean between precision and recall which allows fair measurement of both incorrect positives and incorrect negatives. The method brings exceptional results to imbalanced datasets.

The formula for F1 Score is represented in the Eq. (9):

$$F1Score = 2.\frac{Precision.Recall}{Precision+Recall} \quad (9)$$

Fig. 9 shows the performance metrics of the proposed SONN-AFSA with 98.8% accuracy, 96.5% precision, 94.5% recall and 95.5% F1 score.

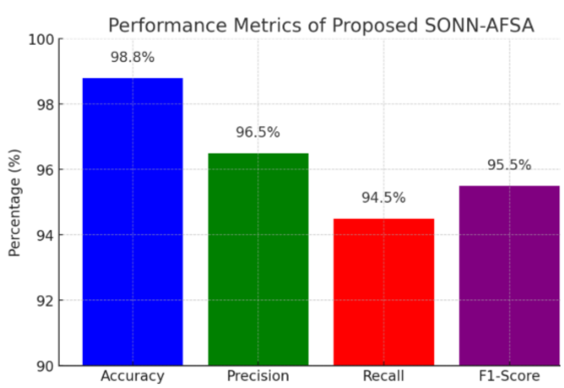


Fig. 9. Performance metrics of proposed SONN-AFSA.

Table I illustrates the performance metrics of proposed method with comparison of exiting Deep learning method.

TABLE I. COMPARATIVE ASSESSMENT

| Method | Accuracy (%) | Precision (%) | Recall (%) | F1-Score (%) |
|---|---|---|---|---|
| Particle Swarm Optimization (PSO) [23] | 88.0 | 85.5 | 87.0 | 86.2 |
| Deep Reinforcement Learning (DRL) [24] | 90.5 | 89.0 | 90.0 | 89.5 |
| PSO-Based Neural Network [25] | 92.0 | 91.0 | 91.5 | 91.2 |
| Proposed SONN-AFSA | 98.8 | 96.5 | 94.5 | 95.5 |

The PSO-Based Neural Network delivers good performance measures by reaching 92.0% accuracy and 91.0% precision and recall and 91.2% F1-score while indicating its worth as a neural network optimization method with particle swarm techniques. Deep Reinforcement Learning (DRL) exhibits equivalent performance to the previous models by reaching 90.5% accuracy and 89.5% F1-score which demonstrates its ability to detect patterns in resource management systems. SONN-AFSA shows superior capability in true positive detection since its recall number (91.5%) exceeds the newly-tested scheme's recall value (90.0%). Fig. 10 demonstrates the comparison of performance metrics across methods.

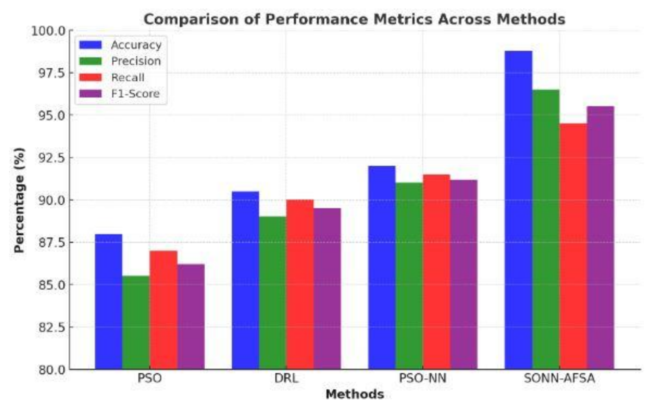


Fig. 10. Comparative assessment.

When it measures accuracy and F1-score, the standalone Particle Swarm Optimization (PSO) method demonstrates results that fall below the hybridized methods with 88.0% accuracy and 86.2% F1-score. PSO represents an efficient optimization solution, however its performance suffers from inadequate adaptive learning capabilities found in neural network-based methods. Results indicate that the SONN-AFSA framework excels as a combination between self-organizing neural networks and artificial fish swarm algorithm optimization performance. The collaborative power between neural networks and artificial fish swarm optimization produces high precision decision-making capabilities that achieve superior outcomes than both traditional and hybrid models.

*I. Discussion*

Memory resource allocation showed efficient performance because usage rates started at 80% and reached 92% while workloads increased. The distributed system demonstrates a best-practice memory management which provides enough memory resources for operations and avoids excessive storage allocation. The cloud environment demonstrates effective resource allocation through storage utilization which increases from 70% to 85% to enable quick data storage and retrieval while maintaining performance speed.

The combined metric measuring average resource utilization rose from 75% to 89% as the system evolved. The system demonstrated improved resource usage performance during the scaling process by achieving balanced resource allocation between different demands. Resource utilization efficiency experienced an increase from 75% to 89% as the system demonstrated both resource effectiveness and optimized resource utilization to minimize waste and enhance system performance. System data indicates growing resource use alongside improved efficiencies which reveals the system can expand its capabilities to handle higher workloads. The system demonstrates superior resource management capabilities through its sustained improvements in efficiency and utilization which enables dynamic resource control for optimal system performance while minimizing resource waste. The system demonstrates excellent suitability for managing expanding cloud requirements.

Fig. 6 examines task duration and its effects on SLA compliance and resource success rates while examining system throughput metrics. The system's task completion time registered a substantial improvement because it decreased from 120 seconds to 90 seconds which led to a 25% reduction.

The resource scheduling and allocation capabilities of this system demonstrate its ability to reduce task execution time. Task execution bottlenecks decreased by 25% in the delivery of key performance metrics when maximum task delay reduced from 200 seconds to 150 seconds. The percentage of tasks successfully executed within service-level agreements improved substantially from 85% to 95%. The model demonstrates its performance preservation capabilities through strict requirements adopting both time and quality restrictions that produce noticeable effects on user satisfaction and service-level agreement fulfillment. The system demonstrated improved reliability alongside stronger processing capabilities because the task achievement rate increased to 98% from 90%.

The system accomplished task processing at a rate of 62 tasks per minute which represented a 24% improvement over its starting point at 50 tasks per minute. The model's scaling performance demonstrates its capability to handle maintaining performance throughout increased operational task volumes. Experimental findings show that the proposed model performs effectively for real-world cloud environments through improved reliability and reduced delays while increasing operational efficiency.

The review of Self-Organizing Neural Networks with Artificial Fish Swarm Algorithm (SONN-AFSA) under 1000 tasks system load presents performance and scalability results in Fig. 9. The model achieved evaluation through measurement of average latency and throughput alongside scalability factor performance in dynamic cloud environments. Static and dynamic resource allocation wait times, reflection ratio variation rate and minimum processor idle time prove the efficiency of SONN-AFSA in optimizing cloud system resource management. SONN-AFSA shows sufficient deployment potential for practical use through its simultaneous performance of reduced latency and increased streamline operations while enabling scalable resource utilization.

The comparison shows how the proposed Self-Organizing Neural Network with Artificial Fish Swarm Algorithm (SONN-AFSA) performs better than other optimization methods as well as learning techniques. The SONN-AFSA model demonstrates superior performance over all metrics because it reaches 98.8% accuracy and 96.5% precision while also achieving recall levels of 94.5% and F1-score of 95.5%. Such high-performance metrics highlight the model's exceptional capability to distribute resources effectively while generating precise system outcome predictions within cloud application spaces.

Table I illustrates the performance metrics of proposed method with comparison of exiting Deep learning method and Fig. 10 shows the performance metrics of the proposed SONN-AFSA.

The results indicate that the SONN-AFSA framework excels as a combination between self-organizing neural networks and artificial fish swarm algorithm optimization performance. The collaborative power between neural networks and artificial fish swarm optimization produces high precision decision-making capabilities that achieve superior outcomes than both traditional and hybrid models.

## V. Conclusion and Future Work

The research implemented an advanced framework to handle energy-efficient cloud resource management through the integration of Self-Organizing Neural Networks (SONN) and Artificial Fish Swarm Algorithm (AFSA). The proposed hybrid design performed severely better than former approaches including PSO and DRL alongside PSO-based Neural Networks. The experimental evaluation led to substantial conclusions about energy reduction by 20.83 percent and 89 percent resource utilization efficiency improvement. The proposed model reached 98.8% prediction accuracy whereas PSO-based Neural Networks achieved only 92.0% accuracy as its best result. The model enhanced SLA compliance to 95% while reaching 98% task completion rates which showcased its ability to handle resources efficiently with superior service quality delivery.

Several upcoming developments should be investigated to enhance both the model's scalability and its applicability potential. Tiny feedback systems with real-time measurements about cloud load dynamics and energy usage statistics enable the algorithm to transform efficiently as conditions in the cloud environment shift. Introduction of multi-objective optimization approaches will enable the system to achieve energy efficiency equilibrium with performance metrics including cost, user satisfaction as well as task latency. The framework requires testing with more extended diverse datasets to prove its potential application in various cloud infrastructure platforms. Federated learning as an advanced AI technology enables distributed cloud systems to achieve improved security and enhanced performance by addressing current challenges across cloud infrastructure. The developed research framework establishes its core foundation as demonstrated in this study but it will use modern advancements to create an effective adaptable model for evolving cloud computing demands.al Fish Swarm Algorithm (AFSA). The proposed hybrid model demonstrated exceptional performance compared to traditional methods, such as Particle Swarm Optimization (PSO), Deep Reinforcement Learning (DRL), and PSO-based Neural Networks. Results from the experimental analysis highlighted a significant reduction in total energy consumption by 20.83%, alongside an improvement in average resource utilization efficiency to 89%. The model also achieved a 98.8% prediction accuracy, outperforming the next-best method, PSO-based Neural Networks, which achieved an accuracy of 92.0%. The model enhanced SLA compliance to reach 95% while achieving 98% completion rates of tasks which showed its capacity to manage resources efficiently and maintain high-quality service delivery.

Upcoming work will investigate multiple advancement possibilities to enhance both the scalability and usability of the model. The real-time feedback mechanisms that track dynamic workload variations and energy consumption stats enable better model adaptation during changing cloud environments. Mobile applications benefit from multi-objective optimization methods which simultaneously optimize energy efficiency alongside cost elements and task delays and user satisfaction criteria. The validation process merits testing using large diverse datasets that will show the framework's applicability across different cloud computing networks. The implementation of federated learning as an advanced AI paradigm would enhance distributed cloud system security and performance to address new infrastructure requirements in cloud computing.

These advances will develop upon the stable structure from this research to keep the SONN-AFSA model functional for changing cloud computing environments.

## REFERENCES

[1] M. A. Al-Sharafi, M. Iranmanesh, M. Al-Emran, A. I. Alzahrani, F. Herzallah, and N. Jamil, “Determinants of cloud computing integration and its impact on sustainable performance in SMEs: An empirical investigation using the SEM-ANN approach,” Heliyon, vol. 9, no. 5, p. e16299, May 2023, doi: 10.1016/j.heliyon.2023.e16299.

[2] P. Borra, “AN OVERVIEW OF CLOUD COMPUTING AND LEADING CLOUD SERVICE PROVIDERS,” Int. J. Comput. Eng. Technol. IJCET, vol. 15, no. 3, Art. no. 3, May 2024.

[3] D. T. A. Ahmed, S. R. Jena, M. S. K. Bhatt, and M. Gali, CLOUD COMPUTING: A COMPREHENSIVE OVERVIEW OF CONCEPTS, TECHNOLOGIES AND ARCHITECTURES. Xoffencer International Publication, 2023.

[4] F. Khoda Parast, C. Sindhav, S. Nikam, H. Izadi Yekta, K. B. Kent, and S. Hakak, “Cloud computing security: A survey of service-based models,” Comput. Secur., vol. 114, p. 102580, Mar. 2022, doi: 10.1016/j.cose.2021.102580.

[5] D. Dina, Emerging Trends in Cloud Computing Analytics, Scalability, and Service Models. IGI Global, 2024.

[6] “(PDF) An Overview of cloud Resource Management Techniques.” Accessed: Jan. 24, 2025. [Online]. Available: https://www.researchgate.net/publication/379652396_An_Overview_of_cloud_Resource_Management_Techniques

[7] O. Ghandour, S. El Kafhali, and M. Hanini, “Adaptive workload management in cloud computing for service level agreements compliance and resource optimization,” Comput. Electr. Eng., vol. 120, p. 109712, Dec. 2024, doi: 10.1016/j.compeleceng.2024.109712.

[8] “(PDF) Towards Efficient Resource Allocation for Heterogeneous Workloads in IaaS Clouds.” Accessed: Jan. 24, 2025. [Online]. Available: https://www.researchgate.net/publication/283948945_Towards_Efficient_Resource_Allocation_for_Heterogeneous_Workloads_in_IaaS_Clouds?_tp=eyJjb250ZXh0Ijp7ImZpcnN0UGFnZSI6InB1YmxpY2F0aW9uIiwicGFnZSI6Il9kaXJlY3QifX0

[9] “Auto-Scaling Techniques in Cloud Computing: Issues and Research Directions.” Accessed: Jan. 24, 2025. [Online]. Available: https://www.mdpi.com/1424-8220/24/17/5551

[10] “(PDF) Energy Consumption in Cloud Computing Data Centers,” ResearchGate, Oct. 2024, doi: 10.11591/closer.v3i3.6346.

[11] “(PDF) Cloud Resource Allocation Strategies for Minimizing Energy Consumption.” Accessed: Jan. 24, 2025. [Online]. Available: https://www.researchgate.net/publication/384808446_Cloud_Resource_Allocation_Strategies_for_Minimizing_Energy_Consumption

[12] A. H. Nebey, “Recent advancement in demand side energy management system for optimal energy utilization,” Energy Rep., vol. 11, pp. 5422–5435, Jun. 2024, doi: 10.1016/j.egyr.2024.05.028.

[13] A. Katal, S. Dahiya, and T. Choudhury, “Energy efficiency in cloud computing data centers: a survey on software technologies,” Clust. Comput., vol. 26, no. 3, pp. 1845–1875, Jun. 2023, doi: 10.1007/s10586-022-03713-0.

[14] “Energy efficient resource management in data centers using imitation-based optimization | Energy Informatics | Full Text.” Accessed: Jan. 24, 2025. [Online]. Available: https://energyinformatics.springeropen.com/articles/10.1186/s42162-024-00370-y

[15] “(PDF) Optimizing Cloud Resource Provisioning with Machine Learning.” Accessed: Jan. 24, 2025. [Online]. Available: https://www.researchgate.net/publication/384766637_Optimizing_Cloud_Resource_Provisioning_with_Machine_Learning

[16] P. K. Bal, S. K. Mohapatra, T. K. Das, K. Srinivasan, and Y.-C. Hu, “A Joint Resource Allocation, Security with Efficient Task Scheduling in Cloud Computing Using Hybrid Machine Learning Techniques,” Sensors, vol. 22, no. 3, Art. no. 3, Jan. 2022, doi: 10.3390/s22031242.

[17] S. Sunil and S. Patel, “Energy-efficient virtual machine placement algorithm based on power usage,” Computing, vol. 105, no. 7, pp. 1597–1621, 2023.

[18] A. Shahidinejad, M. Ghobaei-Arani, and M. Masdari, “Resource provisioning using workload clustering in cloud computing environment: a hybrid approach,” Clust. Comput., vol. 24, no. 1, pp. 319–342, Mar. 2021, doi: 10.1007/s10586-020-03107-0.

[19] M. Ataei Nezhad, H. Barati, and A. Barati, “An authentication-based secure data aggregation method in internet of things,” J. Grid Comput., vol. 20, no. 3, p. 29, 2022.

[20] R. Yadav, W. Zhang, K. Li, C. Liu, and A. A. Laghari, “Managing overloaded hosts for energy-efficiency in cloud data centers,” Clust. Comput., pp. 1–15, 2021.

[21] S. Goyal et al., “An Optimized Framework for Energy-Resource Allocation in a Cloud Environment based on the Whale Optimization Algorithm,” Sensors, vol. 21, no. 5, Art. no. 5, Jan. 2021, doi: 10.3390/s21051583.

[22] google/cluster-data. (Jan. 24, 2025). TeX. Google. Accessed: Jan. 25, 2025. [Online]. Available: https://github.com/google/cluster-data

[23] P. Pirozmand, H. Jalalinejad, A. A. R. Hosseinabadi, S. Mirkamali, and Y. Li, “An improved particle swarm optimization algorithm for task scheduling in cloud computing,” J. Ambient Intell. Humaniz. Comput., vol. 14, no. 4, pp. 4313–4327, 2023.

[24] W. Zhang et al., “Deep reinforcement learning based resource management for DNN inference in industrial IoT,” IEEE Trans. Veh. Technol., vol. 70, no. 8, pp. 7605–7618, 2021.

[25] S. Nabi, M. Ahmad, M. Ibrahim, and H. Hamam, “AdPSO: adaptive PSO-based task scheduling approach for cloud computing,” Sensors, vol. 22, no. 3, p. 920, 2022.