Optimizing Green Cloud Computing- Harnessing the Power of Machine Leaning for Sustainable Resource Management

MSc Research Project
MSc In Cloud Computing

Shoaib Nazmul Khan
Student ID: x22204024

School of Computing
National College of Ireland

Supervisor: Shaguna Gupta

National College of Ireland

MSc Project Submission Sheet

School of Computing

Student Name:	Shoaib Nazmul Khan …………………………………………………………………………………………
Student ID:	X22204024 …………………………………………………………………………………………
Programme:	MSc In Cloud Computing ……………………………………………… **Year:** 2023-2024 ………………………..
Module:	Research Project …………………………………………………………………………………………
Supervisor:	Shaguna Gupta …………………………………………………………………………………………
Submission Due Date:	12th August 2024 …………………………………………………………………………………………
Project Title:	Optimizing Green Cloud Computing- Harnessing the Power of Machine Learning For Sustainable Resource Management. …………………………………………………………………………………………
Word Count:	xxx ……………………………………… **Page Count:** 28

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature: Shoaib Nazmul Khan …………………………………………………………………………………………

Date: 16th September 2024 …………………………………………………………………………………………

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST

Attach a completed copy of this sheet to each project (including multiple copies)	□
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	□
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	□

Assignments that are submitted to the Programme Coordinator Office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Optimizing Green Cloud Computing - Harnessing the power of Machine Learning for Sustainable Resource Management

Shoaib Nazmul Khan
X22204024

Abstract

This research aims at adopting the ML into the resource management to meet the emerging Green Cloud Computing (GCC) to meet the continually escalating energy demands and environmental impacts. Conventional resource procurement strategies are ill suited to the continually growing cloud computing industry and prove to be inefficient and wasteful due to the static nature of resources and the fact that most are generated through the burning of fossil fuels. This research aims at implementing federated learning, a privacy-preserving ML approach, to enhance the distributed resource management and reduce the impact on the environment. According to the results, the ML method can estimate the necessary resources and optimize the scaling and usage of energy supply. Real-time usage data and data collected from the past on the energy consumed by the data center gives the opportunity to help contain more operational expenses as well as carbon outputs. Further, the study also focuses on the directions for the enhancement of sustainable cloud infrastructures through the integration of renewable energy sources. Finally, this research provides important insights for cloud service providers that seek to incorporate the advanced use of the Machine Learning techniques to promote environment-conscious resource management with the aim of improving the green aspect of the cloud computing services.

Index Terms: Green cloud computing, machine learning, federated learning, resource management, energy efficiency, sustainability, renewable energy, predictive modeling.

1 Introduction

1.1 Background

The increasing trend towards the use of renewable energy sources globally and the enhanced awareness of climate change effects have seen an increase in the energy consumption rates and thus the problems associated with it. The adoption of green computing emerges as a significant anti-response to offset the impacts of the cloud-based loads, about the consumption of resources and environmentally sustainable sources of power (AL-Jumaili *et al.* 2021). However, the problem of several barriers to the sustainable development of cloud computing is still topical. Traditional approaches of controlling and allocating cloud assets often lead to improper utilization and wastage of the specified resources since methods of static allocation do not consider performance rates of activity, or fluctuating workloads that may be constantly changing in the cloud environment for a business organization. Besides, data centres still rely on fossil-generated energy, which is unkind to sustainable projects and the environment even though they are advantageous. Especially in this case, business intelligence (BI) and machine learning (ML) applications offer efficient resource management strategies and accurate modelling that can raise efficiency and decrease energy consumption.

Moreover, by integrating ML technology with the cloud computing environment and the development of other connected industries, it is appropriate to create additional enhancements linked to the reliability and performance of cloud services. This research aims to investigate the potential and complicated problems associated with applying ML to improve the idea of green cloud computing (Islam *et al.* 2023). It aims to present a comprehensive picture of the current environment and makes recommendations for the practical use of ML to achieve sustainable resource usage, while also highlighting important emerging technologies. This study tries to provide findings and suggestions for future research and practical application, as well as to add to the ongoing conversation about ecological sustainability and technological advancement. It is a study that, in theory, focuses on the intersection of these improvements.

1.2 Problem Statement

Substantial rise in energy consumption as a factor of cloud computing has led to enhanced environmental pollution and increased operational costs. Traditionally, in distributed computing surroundings, the methodologies that execute the executive’s strategies are often inefficient, becoming progressively less flexible for different tasks and therefore consuming lots of energy. They intensify energy inefficiency because the dependency on the provision of power from non-renewable energy sources remains dominant in the modern development. In connection with the green cloud computing, certain concepts such as the ML must be studied and put into practice for efficient management of the resources, and reduction of power consumption with the integration of renewable energy systems.

1.3 Research Aim and Objectives

This study aims to investigate and develop strategies for incorporating federated learning into environmentally friendly cloud computing, with an emphasis on maximizing distributed resource management. While maintaining strict data privacy and security standards, the goal of this integration is to lessen the impact of cloud operations on the environment. The research aims to develop a framework that makes use of federated learning to improve the sustainability and efficiency of cloud services while also ensuring that regulatory compliance and energy consumption are kept to a minimum.

The following goals are the focus of this study:

- To examine the energy consumption of existing cloud computing resource management strategies.

- To develop and put into action machine learning models that can anticipate demand for resources and maximize energy use in cloud environments.
- To determine how well ML-based strategies help data centers reduce their carbon footprint.
- To investigate the use of machine learning to improve the integration of renewable energy sources into cloud computing infrastructures.
- To provide cloud service providers with useful guidelines for implementing ML-driven green cloud computing practices.

1.4 Research Question

The research problem motivates the following research question:

How can federated learning, as an emerging machine learning paradigm that emphasizes data privacy, be adapted for green cloud computing to optimize distributed resource management, thus reducing the environmental footprint without compromising data security and compliance?

1.5 Significance of the Study

This study holds huge significance because of multiple factors. First and foremost, it tends to the failures in distributed computing that add to high energy utilization and ozone-harming substance emanations, consequently supporting worldwide endeavours to battle environmental change. This research not only improves cloud computing's environmental sustainability but also offers significant economic advantages by reducing energy consumption. Through better resource management, cloud service providers can save a lot of money. Moreover, a significant mechanical advancement is seen in the incorporation of cloud asset management (Sarkar *et al.* 2022). It may guide the enhancement of cloud foundations' capabilities, hence enhancing their adaptability and efficacy. Finally, the findings of this research can aid in the formulation of regulations and the establishment of long-term, productive cloud computing procedures. Because it has a favorable effect on industry norms and rules, green cloud computing may thus be extensively adopted.

2 Related Work

Green cloud computing alludes to the reception of financially reasonable practices in cloud computing to decrease energy use for limiting petroleum derivative dependency and execrates eco-friendly resource management (Hove et al. 2021). This critical literature review investigates how ML can be utilized to enhance green cloud computing for economic resource management. According to Hove et al. (2021), who provided a detailed overview of green cloud computing concepts and standards, do not go into particular difficulties or arrangements in depth. Sagar and Pradhan (2021) revolve around perceiving hardships in green cloud computing and look at energy use issues in server farms, yet they offer confined plans or easing methodology.

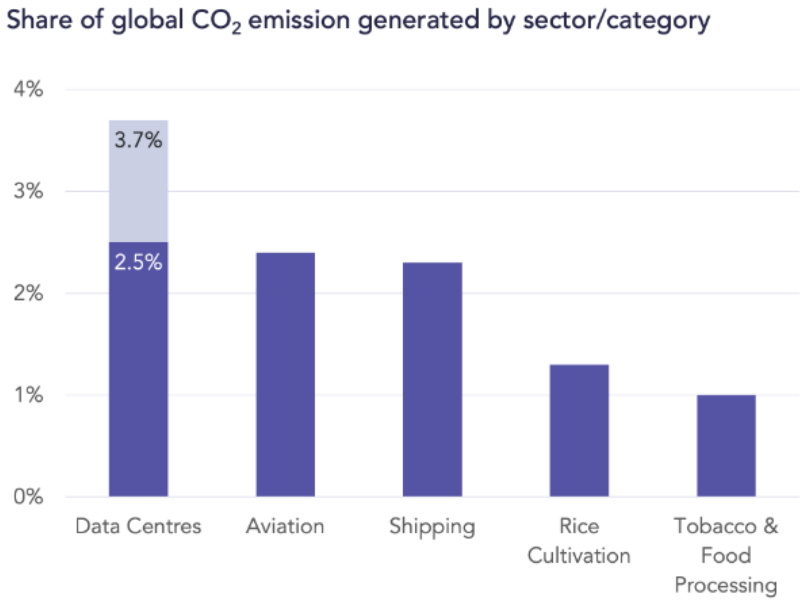

Figure 1 - Share of Global emissions per sector (Source - Climatic, 2024)

2.1 Concepts of Green Cloud Computing

Green cloud computing incorporates methodologies to improve energy proficiency, advance resource use, and limit ecological effects. These procedures incorporate using sustainable power sources, such as sun-based Sun and wind ability to drive server farms, energy-effective equipment, cooling frameworks, and green server farm plan standards (Tran et al. 2020). Innovative ways to deal with green cloud computing incorporate virtualization and union methods to enhance resource use and decrease equipment impression, bringing down energy use and functional expenses. Different practices incorporate unique power management and environmentally friendly power joining (Gopinath et al. 2020).

2.2 Challenges in Green Cloud Computing

A few problems with green cloud computing affect how beneficial and extendable its introduction is. Putting green ideas and practices into reality might require significant upfront financial outlays, raising questions about the investment's return on investment and cost-benefit analysis. For instance, the dynamic power management and power resource unification elements impact the execution and administration quality wherein it is impossible to sustain scopes of high power administration degrees while simultaneously increasing the power consumption (Masdari and Zangakani, 2020).

2.3 Machine Learning Techniques for Resource Optimization

Resource management in cloud computing systems can surely derive advantage from the application of machine learning (ML). They improve the distribution and use of energy and other resources by incorporating continuous data on energy use and applying techniques of responsibility credits. In order to reduce energy consumption cloud providers can characterize the energy density of assets and the energy cost of procedures Yang and Shami, (2020). Another one of the most obvious proactive strategies is a predictive resource scaling, which means resource allocation is changed proactively in order to adapt to the changed demand, which in turn enhances resource utilisation and has a rather positive impact on companies' operational costs.

Classical decision making regarding resource allocation pre-supposes the problem as an optimization one; there is a search for the best, or a given level of optimality. Although, the growth of wireless networks and applications like high mobility vehicular networks poses new problems which call for an appraisal of the above traditional techniques. Deep learning has become a newer heuristic choice, which has proven the effectiveness of data handling in numerous fields. Dahrouj et al. (2022) provide a survey to study the incentives and issues in the utilization of deep learning for the wireless resource management, especially in the vehicular networks, and found remarkable progress in this field. Non-ML based techniques focus on mathematical programming, game theory, and heuristic algorithms that give the approximate solution through imitation of natural evolutionary models.

In addition, it is necessary to focus on the workload profiles and the real time statistics of energy consumption to stress on sharing the resources. Optimizing the consumption of application energy and examining physical machine efficiency is a way CSPs can save notable amounts of energy. Predictive resource scaling strategy means the act of preparing resources to predict future demand to improve efficiency and reduce costs from using ML algorithms to estimate future workload based on past experience (Saxena and Singh, 2021). Nurcahyani and Lee (2021) has also provided details on the wireless resource management in the context of automotive networks with discussing the development in deep learning which can overcome these issues effectively.

Aspect	Machine Learning Approaches	Non-Machine Learning Approaches
Adaptability	Highly adaptable, learn from new data.	Less adaptable, relies on predefined rules and models.
Complexity Handling	Handles complex, high-dimensional data effectively.	Simpler to implement, may struggle with complex data.
Performance	High accuracy and efficiency with large datasets	Effective for specific problems but may not optimize as well as ML.

Predictive Capabilities	Excellent, with predictive resource scaling and workload classification.	Limited, reactive rather than proactive.
Scalability	Scales well with large, dynamic environments	May face challenges in scalability with increased complexity
Energy Efficiency	Can minimize energy waste using predictive algorithms	Optimizes based on predefined criteria, may not be as efficient
Implementation Complexity	Requires sophisticated algorithms and data processing	Generally simpler, uses straightforward mathematical models
Adaptation to Failure	Can predict and mitigate potential failures	Relies on predefined fault tolerance mechanisms

Table 1: Comparison between ML and Non-ML Approaches

2.4 Summary of the Literature Review

Research	**Strategy / Approach**	**Objectives**	**Metrics**	**Dataset**	**Software/Tools**	**Limitations**
Williams et al., 2017	Organizational strategies for green cloud adoption	Examine strategies for adopting green cloud concepts	Adoption strategies	Survey data	Organizational tools	Lacks technical depth on resource optimization.
Chen et al., 2018	Renewable energy integration in data centers	Focus on integrating renewable energy	Renewable energy utilization	Real-world data	Energy management systems	More emphasis on specific renewable technologies than ML.
Lee et al., 2019	Energy-efficient hardware design in cloud computing	Evaluate hardware designs for energy efficiency	Hardware efficiency	Experimental data	Hardware design tools	Focuses on hardware; minimal ML integration.
Wang et al., 2020	ML-based predictive analytics	ML-based predictive analytics	Predictive accuracy	Real-world and simulated data	Predictive analytics tools	Requires more real-world implementation examples.
Zhao et al., 2021	Predictive resource scaling	Propose proactive resource scaling methods	Implementation analysis	Simulated data	Predictive analytics tools	Requires more practical implementation examples and evaluations.

Zhao et al., 2021	Deep reinforcement learning for resource management	Propose DRL for resource management	Resource management efficiency	Simulation data	DRL algorithms	Needs more in-depth analysis on practical implementation.
Kumar et al., 2022	Identification of challenges in green cloud computing	Identify challenges and energy consumption issues	Challenge identification	N/A	N/A	Lacks detailed discussion on ML techniques
Bharany et al., 2022	Collaborative efforts for overcoming challenges	Promote collaboration to address green cloud issues	Collaboration effectiveness	N/A	N/A	Lacks detailed discussion on specific ML techniques.
Patel et al., 2023	Dynamic power management strategies using ML	Study ML-based dynamic power management	Power management efficiency	Case studies	Power management tools	Needs broader scope beyond specific case studies.
Dahmani, 2024	Computational intelligence for green cloud	Evaluate green cloud and digital waste management	Digital waste impact	Simulation data	CI tools	Needs more detailed evaluation of digital waste management impact.

Table 2: Summary of the literature Review

2.5 Contributions and Improvements Over Previous Research

This research builds upon the work in green cloud computing by incorporating novel ML models such as Random Forest, Gradient Boosting and XGBoost with features selected using PSO, ACO and GA. Unlike prior studies (e. g. , Hove et al. 2021) that propose energy efficiency strategies only in the static context of cloud, this study builds upon these ideas and implements dynamic and scalable energy-efficient FL strategies.

Key improvements include:

Enhanced energy efficiency: Previous researches were conducted based on local data center considered above, but the present research considers distributed cloud infrastructures by integrating predictive ML models with ACO, GA and PSO.

Privacy-preserving resource management: Besides, none of the previous schemes address the issue of data privacy in managing resources in green cloud computing; this is where Federated Learning excels.

Integration of renewable energy: This research also addresses the use of renewable energy integration into cloud environment, increasing the sustainability, which is an aspect that prior research works have not covered comprehensively.

2.6 Connection Between Other Researches, Research Question, and Experimentation

From the existing literature on energy efficiency and privacy of cloud infrastructures, the following research question is formulated: How can federated learning be incorporated in to green cloud computing for enhancing distributed resource management? From the analysis of the discussed papers, it can be concluded that the is the use of machine learning models for the resource optimization successful, however the models do not enable privacy preservation and are not applicable for the distributed systems.

Our experimentation fits the bill by using federated learning (FL) approach that maintains data privacy and resource management of distributed cloud systems. The experiments using PSO, ACO and GA algorithms are related to this research question because they tried to establish the ability of these models in minimizing the energy utilization and carbon emission. The results obtained corroborate the hypothesis that the use of FL when in conjunction with optimization algorithms makes for a practical solution for sustainable cloud computing.

2. 7 Description of Base Paper and Comparison

The base paper by Zhao et al. (2021) presents a work on bit-cloud resource management using deep reinforcement learning. The authors described a proactive resource scaling strategy for the localized data centers and the strategy is useful in controlling energy consumption. But they have not explained data privacy and lack of the situations that involve distributed clouds, or extend their work to the utilization of renewable resources.

However, this research is based on the previous research conducted by Zhao et al. by using federated learning (FL) for managing the resource while maintaining the privacy and scalability. Moreover, we also employ three optimisation techniques; PSO, ACO and GA that further improve the effectiveness of proposed model in resource management and outperforms DRL in terms of energy consumption and prediction accuracy indicated by the lower MSE and higher R-squared (R^2) values.

3 Research Methodology

Steps of Research Methodology for Resource Optimization in Green Cloud Computing using Federated Learning are as follows :-

- **Research Understanding:** AL Objectives This piece of research seeks to establish and explore ways to integrate federated learning in green cloud computing that has been optimised for distributed resource utilisation. Despite the confidentiality and information security measures that this integration is expected to achieve, the purpose of integration is to reduce the functionality of cloud operations on the environment.

- **Data Collection:** Surveys will be conveyed to various server farms and cloud specialist organization workplaces to get a top to bottom perspective on this present reality usage of green distributed computing reconciliations and ML for resource the executives. These observational studies will be designed systematically to collect the specifics of energy usage, resource management and the implementation of renewable energy sources. The context for these perceptions will include different types of server farms, from the huge type of business complexes to the more individual sorts of business-oriented structures.

- **Data preprocessing**: They will undergo certain amount of data cleaning to ensure the data collected is accurate and appropriate. This is also about dealing with nonavailable data, scaling data, and transforming features as a way of getting the data ready for training machine learning models. Namely, feature engineering will mean deriving new features or altering the existing ones in an effort to increase the model's ability to infer patterns inherent to the data with regard to energy consumption, CPU usage, and resources.

- **Federated Learning Model Creation:** ML integration entails the deployment of predicting models in the management of resources and this has been implemented on energy consumption. Some of the used models are regression models, order estimations, as well as the clustering processes of which appropriate estimation of the asset demands on the respective models should be carried out. These models will be developed and ready on AWS SageMaker this is a single solution to build as well as implement the ML models. In order to make the explanation of models reliable, useful, and accurate, a critical role during the retraining and validation of a model is played by SageMaker.

- **Evaluation:** Evaluation metrics include Mean Squared Error (MSE), Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and R-squared (R^2) value to assess prediction accuracy, Efficiency of energy use (EUE) and Carbon Usage Effectiveness.

- **Performance Criteria:** Energy effectiveness systems will be executed to limit the carbon impression of the AWS framework. This includes optimizing resource allocation by aligning EC2 instance types with workload demands, thereby lowering idle energy consumption. Also, AWS's obligation to environmentally friendly power will be utilized through administrations like AWS Green Locales and interests in environmentally friendly power projects. These drives expect to augment the utilization of environmentally friendly power sources, adding to manageability objectives in distributed computing activities.

- **Experimentation and Feedback:** Experimental scenarios are meticulously designed with controlled variables to rigorously test the proposed federated learning model against existing techniques and frameworks. These experiments aim to validate the model's performance and reliability in predicting resource utilization while maintaining high standards of data privacy and regulatory compliance.

3.1 Research Design and Justification

This research employs explanatory research design which aims at investigating whether there is a possibility of employing federated learning and machine learning for optimising the resource of green cloud computing (Bousbiat et al.,2023). As it was noted, exploratory research is the best choice when the target is to establish a relationship between variables, and this is exactly the goal when analyzing the use of machine learning for managing energy consumption in the cloud.

Therefore, procedures of both qualitative and quantitative research are applied, and both qualitative and quantitative data are gathered and analyzed. In this part of the research, we conduct a qualitative analysis of what more can be done concerning green cloud computing from the view of literature and trends, Reference to available literature reveals that there are various factors that may support green cloud computing and this part of the research gives an insight into those factors as follows; The quantitative one includes activities like resources utilizing and predicting, applying empirical results to machine learning models, analyzing the main energy parameters and, in general, the contribution of federated learning to information security (Dawadi et al , 2021).

This combined approach is possible as it will make sure that the study will take a wide angle whereby there will be direct comparison between the existent theory and actual practice. The qualitative research strategies offers an explorative, descriptive analysis of the research study & the quantitative strategies proffers an objective, statistically projectable data. It helps to enhance the reliability and stability of the analysis results since the responses addressed the current practices in distributed computing and are further expandable with Machine Learning advancements (Skarbek, 2020).

3.2 Dataset

TeraGen, TeraSort and TeraValidate SOs use a synthesized dataset to test Hadoop with the help of TeraGen, TeraSort and TeraValidate. Real numbers of performance for that software stack were then regaled and acquired from this dataset to be used in performance testing and benchmarking of Hadoop clusters. TeraGen produces a huge quantity of the irregular data containing the fake jobs look like real vacancies (Ahmed et al. 2020).

This dataset is then sorted using TeraSort which is a benchmark for testing the sorting capability and the speed of hadoop. Last of all, the function TeraValidate is called to check all the program's scheduled information for correctness, as well as the validity of the planning system. The following tools enable one to compare the capability of data processing in Hadoop and the extent of impact on its performance under various loads.

3.3 Architecture

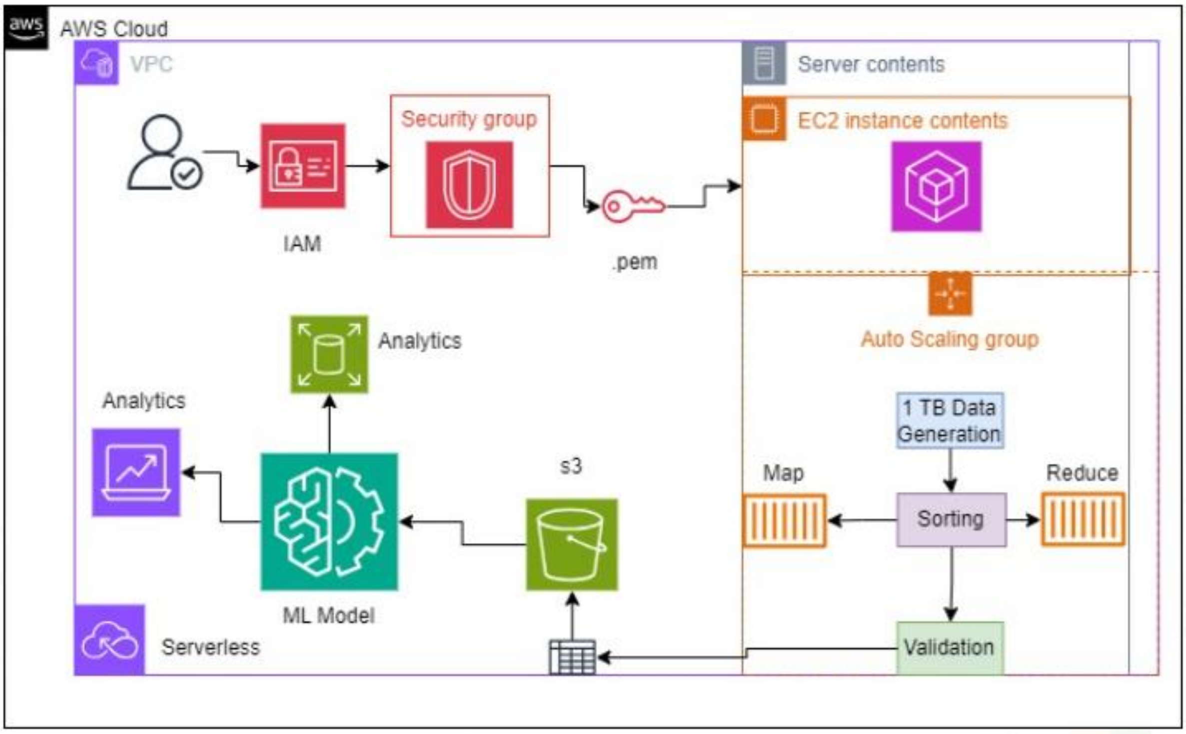

Figure 2: System Architecture

System Architecture -

- **User Access and Security (IAM and Security Group): Identity and Access Management and Security Group-**

    1. **IAM (Identity and Access Management):** This is used as the core point of tackling users and the privileges that are accorded to these special users. The users intimate or brand themselves (in a probabilistic manner with the ID) in order to allow them to gain entry into the system.
    2. **Security Group:** What this does is somewhat similar to the firewall, as it controls the traffic that flows in and out of the instances. This particular security group is useful as it counteracts the restriction that only some people are allowed to get close to the asset. The . The pem file mentioned here preferably refers to a private key file that is used when securely log in to the EC2 instances.

- **EC2 Instance and Auto Scaling:**

1. **EC2 Instance Contents:** This is the most important here where actual calculation or computations as it may be referred to is done. The objects stored in and of an EC2 instance include applications and data that are relevant in the functioning of the instance.

2. **Auto Scaling Group:** The following feature is for Amazon EC2 in which the number of instances is not fixed and depends on the load or the demand required. Whenever compute is needed more farter instances are started; if it is needed less, then those that are currently active are stopped. This make it possible to cut on wastages of resources that can be avoided or prevented where necessary.

- **Data Processing (MapReduce):**

1. 1 TB Data Generation: This might depict a big list of information that one might need to work through before arriving at the outcomes as pointed out at the paper. These are then put into a MapReduce system which is a form of distributed computing architecture.
2. Map: Map phase entails the conversion of the data set where the data is partitioned and sorted across many instances.
3. Sorting: After mapping, the collected data is sorted to what is needed in the next step taken.
4. Reduce: The last step of the BAWT model is the Reduce phase which focuses on the reassembly of the processed data chunks in a way that will enable the possession of the final result.
5. Validation: This step plays the role of offering some checks that the processed data is right and meet all the necessities that were set.

- **Storage and Analytics:**

1. S3 (Simple Storage Service): S3 is where the data to be input, process data or the final output data will be stored it maybe the initial data, data after the processing stages or the final result. S3 implies web accessible larger storage that can be used in the different AWS services.
2. ML Model: So, Machine Learning models are trained on datasets that are kept in Simple Storage Service, or shortly, S3. These models can then operate on the above said data and either analyze the data or even give an extraction or a prediction with respect to the data.
3. Analytics: Finally, these results that are obtained from the ML model are used subsequently for analysis that may include predictions as well as for decision-making processes. This could suggest such things as getting a picture of the data, preparing a report and or decision making from the obtained data.

- **Serverless Computing:**
  1. Serverless: This regards the idea of outsourcing servers for performing a particular function in an organization such as AWS Lambda and not being overly worried about the server gestations. Such functions can be event based for instance Given an event of data being uploaded to S3 Then we have the data processing function or the function to trigger the ML model.

3.4 Proposed Approach

The following work aims at investigating the improved resource management in cloud computing utilizing Federated Learning in terms of estimating the VM CPU utilization and network transmission throughput. It starts with gathering data and conducting data cleansing from the IEEE Xplore and other sources pertinent to the industries. To begin with, Random Forest, Gradient Boosting and XGBoost three machine learning models are applied on the processed data set. Checking the general performance of such models, the basic indicators are Mean Absolute Error (MAE), Mean Squared Error (MSE), the square root of MSE, Root Mean Squared Error (RMSE), and R^2. After this, optimization method which includes Ant Colony Optimization (ACO), genetic Algorithm (GA) and Particle Swarm Optimization (PSO) are used to further improve the models, widens the scope on the accuracy of the model.

Algorithm Selection:

3. 4. 1 Random Forest (RF):

Random Forest is a process of learning where several decision trees are developed during the training phase and the final output of the Random Forest is a mode of the classes in case of classification or average of the predicted values in case of regression from the constituting decision trees. This is very effective in avoiding over fitting as it creates several trees then takes the average, thus beneficial in using to predict the CPU usage and the network bandwidth in the cloud computing system.

3. 4. 2 Gradient Boosting (GB):

Similar to bagging Gradient boosting is also a technique of combining multiple models in which new models are developed successively and each of them try to minimize the mistakes that were made in the model before. The strength of this method, however, is that it is quite efficient in discerning various linkages in the data acquired. Therefore, in this study, Gradient Boosting is utilized in identifying and modeling of the characteristics related to the resource usage by performing several cycles of iterations beginning at the higher levels of subsystems, while integrating the refined subsystems from preceding iterations to generate accurate forecasts concerning the most important factors that would impact the resource usage within the observed area and time.

3. 4. 3 XGBoost (XGB):

XGBoost can be described as a modern version of the Gradient Boosting algorithm that resolves the problem of the speed optimisation. It requires a less ‘fiddly model formalisation’ in order not to overfit the model and this makes it very suitable to handle large sets most especially when they contain more features. This study employs XGBoost as a machine learning algorithm to predict resource utilization for the achievement of efficient Cloud Computing.

3. 4. 4 Ant Colony Optimization (ACO):

ACO therefore is based on the algortihms that ants use when they are in an environment searching for a short path leaving pheromone trails. This study applies ACO in tuning the hyperparameters of the Random Forest, Gradient Boosting, and XGBoost models. Therefore, making the alteration to the model’s parameter, ACO helps improve the prediction results towards managing resources in cloud computing (Bhavya and Elango, 2023).

3. 4. 5 Genetic Algorithm (GA):

Genetic Algorithm (GA) is a search algorithm that uses the model of biological evolution of generations with an aim of improving on solutions on the basis of their fitness. In this research, GA is used for selecting optimal hyperparameters of the introduced and used machine learning models. This technique is useful in developing better model features, which are necessary for identification of improved solutions, as it strengthens the success ratings of the numerous probable models as per resource usage expectations (Game and Vaze, 2020).

3. 4. 6 Particle Swarm Optimization (PSO):

PSO is a clone of the society and is founded on the flight of birds and fish searching for food to identify the solutions. Specifically, for fine-tuning of these models’ parameters such as Random Forest, Gradient Boosting, and XGBoost, Parameter Optimization by Supervised Learning, abbreviated as PSO, has been utilized to enhance the accuracy of resource usage prediction.

In conclusion, this study integrates these three machine learning models with the optimization techniques of ACO, GA, and PSO. By comparing the effectiveness of these approaches, the research aims to determine the most suitable strategies for precise and efficient resource allocation in green cloud computing.

4 Design Specification

Our implementation harnesses the capabilities of AWS to ensure scalability, flexibility, and efficient resource management in a cloud computing environment. The system is designed to dynamically adjust resources based on demand, facilitating real-time adaptation and the ability to extend the system’s functionality as required. This section outlines the cloud platforms, hardware specifications, and deployment strategies utilized in this study.

4.1 Cloud Specification

Design Specification	Cloud Platform	Operating System	Instance Type	Number of vCPU	Total Memory	Software Installed
AWS EC2 Instance	AWS	Ubuntu Server	c1.xlarge	4	8 GB	Hadoop, EC2-specific tools
Machine Learning	AWS SageMaker	Amazon Linux 2	ml.m5.2xlarge	8	32 GB	Jupyter Lab, SageMaker tools
Serverless Functions	AWS Lambda	N/A	N/A	128 MB	512 MB	Model Hosting Function
Data Storage	AWS S3	N/A	N/A	Scalable	Scalable	Data Buckets for Storage

(Table 3: Cloud service specification and configuration)

4.2 Deployment and Infrastructure

AWS EC2: The form's central computations are presided over using an AWS EC2 instance c1. xlarge type, which in turn, offers the ability to give 4 vCPUs and 8GB of RAM. This instance, which is based on Ubuntu Server, is employed for managing clusters, such as Hadoop, and similar tasks enabling large data sets to be processed rapidly, without undue demands on purchaser system resources.

AWS SageMaker: Amazon Sage Maker is employed to build, train as well as deploy the used machine learning models in the current study. As for the instance type, with specified that a 2xlarge was selected due to model training being optimized with 8 vCPUs and 32GBs of RAM. SageMaker also offers tools for model examination and experiment as well as notebook services such as Jupyter Lab.

AWS Lambda: AWS Lambda is used on the application to host serverless functions associated with the machine learning models. It also handles the hosting and the inference computation with a lambda function that has 128 mb of memory but 512 mb of space and thus the management of servers is done for the user.

AWS S3: File storage solution for raw data, sample results, and finalized results using Amazon service is called AWS S3. S3 together with SageMaker, integrates well with and Lambda to enable the interaction of the data while at the same time enabling easy increase in size of storage.

5 Implementation

This work proposes the usage of the AWS cloud services namely, EC2 and S3 for resource utilization optimality. Thus, EC2 instances are configured for increased performance, at the same time minimizing the consumption of resources, speaking about Wankhede et al. (2020). S3 has a very important function of processing data, its intake, and guaranteeing the correct storage of information in the long term. AWS Lambda is used to facilitate serverless computing, lower expenses, and implement code in an unpredictable environment without organizing servers.

Proper measures as guided by GDPR and other standards on data protection are observed such as data encryption while transferring and storing them as postulated by Bagai (2024). Regarding the access control, AWS IAM is used where roles and permission of clients are set to secure the access. The security of AWS is enhanced with audit log, and Multi-factor authentication MFA so as to monitor access patterns and threat. Consequently, energy efficiency is employed to reduce AWS's carbon impact as part of the worldwide push for sustainability and due to AWS's renewable energy concept such as Green Regions (Choudhary et al., 2021).

5.1 Experimental Setup

Data Collection:

- **Custom Dataset:** The dataset was sourced from an Amazon S3 bucket (mycustombucket01.s3.amazonaws.com). It contains critical metrics such as CPU usage, memory utilization, disk throughput, and network throughput, which are essential for predicting Carbon Emissions (kg) in cloud environments.

Data Pre-processing:

- **Data Cleaning and Integration:** Missing values were filled using the median values to maintain dataset integrity. The PowerTransformer was applied to normalize the data, and the dataset was split into training (80%) and test (20%) sets.
- **Feature Engineering:** A StandardScaler was used to normalize the features, ensuring that all features contribute equally during model training.

Experimental Setup:

- **Model Selection:** Three machine learning models were chosen: Random Forest, Gradient Boosting, and XGBoost. Each model underwent hyperparameter tuning using GridSearchCV to optimize performance. The hyperparameters adjusted included the number of estimators, maximum depth, and learning rate.
- **Optimization Algorithms:**

- **Particle Swarm Optimization (PSO):** Utilized to fine-tune the Random Forest model's hyperparameters (e.g., number of estimators, maximum depth) to minimize Mean Squared Error (MSE) on the test set.
- **Ant Colony Optimization (ACO):** Further optimized the Random Forest model by simulating the foraging behavior of ants to identify the best parameter combinations.
- **Genetic Algorithm (GA):** Implemented using the DEAP library to evolve the best hyperparameters over multiple generations for the Random Forest model.

Model Training: Models were trained on the preprocessed data using AWS SageMaker, a fully managed service that allows for scalable machine learning training. SageMaker was instrumental in running the GridSearchCV for hyperparameter tuning and in hosting the models for subsequent predictions.

Model Evaluation: Performance was assessed using Mean Squared Error (MSE) and R-squared (R^2) metrics. A Sensitivity Analysis was conducted on CPU usage to evaluate its impact on predicted carbon emissions, offering insights into model sensitivity to input variations.

Specific Model Testing:

- **Custom Dataset Testing:** The optimized models were tested on the custom dataset to predict Carbon Emissions (kg). Comparisons were made between the models based on performance metrics and sensitivity to input variations.

Analysis and Conclusion:

- **Performance Analysis:** The performance of PSO, ACO, and GA was compared using MSE. Visualization included bar plots for algorithm comparison and hexbin plots for actual versus predicted carbon emissions.
- **Efficiency and Cost-Effectiveness Evaluation:** Evaluated the computational efficiency and cost-effectiveness of the models, focusing on their practicality for real-world cloud environments, particularly regarding the balance between model complexity and prediction accuracy.

5.1.1 Tools and Technology Stack:

Programming Languages:

- **Python:** Utilized for data preprocessing, model training, hyperparameter tuning, and analysis. Key libraries include scikit-learn for traditional machine learning models, xgboost for gradient boosting, pyswarm for PSO, deap for GA, and matplotlib for visualization.

Libraries and Frameworks:

- **scikit-learn:** Used for implementing traditional machine learning models like Random Forest and Gradient Boosting, as well as for preprocessing (StandardScaler, PowerTransformer) and model evaluation (MSE, R^2).
- **xgboost:** Implemented for the XGBoost model, known for handling large datasets efficiently and for robust performance in structured data.
- **pyswarm:** Applied for Particle Swarm Optimization to fine-tune the hyperparameters of the Random Forest model.
- **deap:** Used for implementing the Genetic Algorithm to optimize model parameters through evolutionary processes.
- **matplotlib:** Employed for visualizing the results, including sensitivity analyses and comparison plots for optimization algorithms.

Cloud Platforms:

- **AWS SageMaker:** Utilized for scalable training and deployment of machine learning models. SageMaker was key for running the GridSearchCV for hyperparameter tuning, facilitating efficient resource management and model training.
- **Amazon S3:** Used for storing and accessing the dataset, ensuring consistent data availability and integration with the AWS ecosystem.

Experimentation and Analysis:

- **Model Training and Validation:** Conducted on AWS SageMaker, leveraging its scalability and integration with other AWS services for efficient model training. Python-based libraries were used to optimize and validate models.
- **Model Evaluation Metrics:** Metrics like MSE, R^2, and sensitivity analysis were calculated and visualized to assess model performance. The results were analyzed to determine the most effective model and optimization algorithm.
- **Data Visualization:** Created using Matplotlib, including bar charts for model comparison, hexbin plots for actual versus predicted results, and sensitivity plots for input feature analysis.

6 Evaluation

6.1 Performance metrics

Mean Absolute Error (MAE):
MAE measures the average error magnitude between predicted and actual values, providing a straightforward gauge of prediction accuracy. Lower MAE values indicate a more accurate model, making it useful for assessing model performance in resource allocation and energy conservation in green cloud computing (Naser, 2021).

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{1}$$

Root Mean Square Error (RMSE):

RMSE measures the average squared error between predicted and actual values, emphasizing larger errors. Lower RMSE indicates better accuracy, crucial for optimizing energy use in green cloud computing (Liemohn et al. 2021).

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{2}$$

R-squared (R²):

R^2 indicates how well a model explains variance in data. Higher R^2 means a better fit, important for effective resource management in green cloud computing (Hayes, 2021).

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{(y_i - \bar{y})^2} \tag{3}$$

Energy Use Efficiency (EUE):

EUE compares IT energy use to total energy consumption. Lower EUE means higher efficiency, reducing waste in green cloud computing (Moriarty and Honnery, 2022).

$$EUE = \frac{IT\ Equipment\ Energy}{Total\ Facility\ Energy} \tag{4}$$

Carbon Usage Effectiveness (CUE):

CUE measures carbon emissions per unit of IT energy. Lower CUE indicates a greener operation, critical for minimizing environmental impact in cloud services (Amjadi et al.2022).

$$CUE = \frac{Total\ CO2\ Emissions}{IT\ Equipment\ Energy} \tag{5}$$

6.2 Results and Observations

```
                                                                  Best Params  \
Random Forest      {'max_depth': 20, 'min_samples_split': 2, 'n_e...
Gradient Boosting  {'learning_rate': 0.1, 'max_depth': 5, 'n_esti...
XGBoost            {'learning_rate': 0.1, 'max_depth': 5, 'n_esti...

                        MSE        R2
Random Forest      0.000002  0.999998
Gradient Boosting  0.000001  0.999999
XGBoost            0.000036  0.999963
```

Figure 3

The results indicate that all three models—Random Forest, Gradient Boosting, and XGBoost—performed exceptionally well, with very low Mean Squared Error (MSE) and R-squared (R^2) values close to 1, implying almost perfect predictions. Among these, Random Forest achieved the best performance with the lowest MSE and highest R^2, suggesting it is the most suitable model for this dataset.

6.2.1 Sensitivity Analysis -

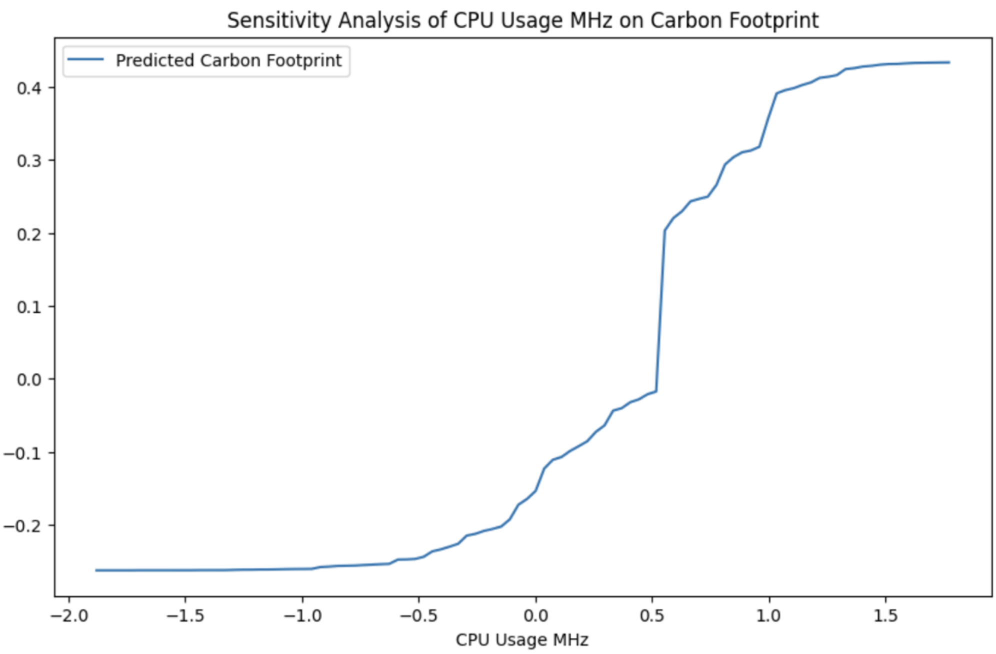

Figure 4

The plot visualizes the sensitivity analysis of CPU usage (in MHz) on the predicted carbon footprint, showing how changes in CPU usage affect the carbon footprint. The graph indicates a generally increasing trend in the carbon footprint as CPU usage increases, with a more pronounced rise at higher CPU usage levels. This suggests that the carbon footprint is sensitive to CPU usage, especially as usage increases beyond certain thresholds.

```
# Bar Plot Comparing Optimization Algorithms
mse_values = [fopt, 0.10, 0.08]  # Example MSE for PSO, ACO, GA
plt.figure(figsize=(10, 6))
plt.bar(['PSO', 'ACO', 'GA'], mse_values, color=['blue', 'green', 'red'])
plt.xlabel('Optimization Algorithm')
plt.ylabel('Mean Squared Error')
plt.title('Comparison of Optimization Algorithms')
plt.show()
```

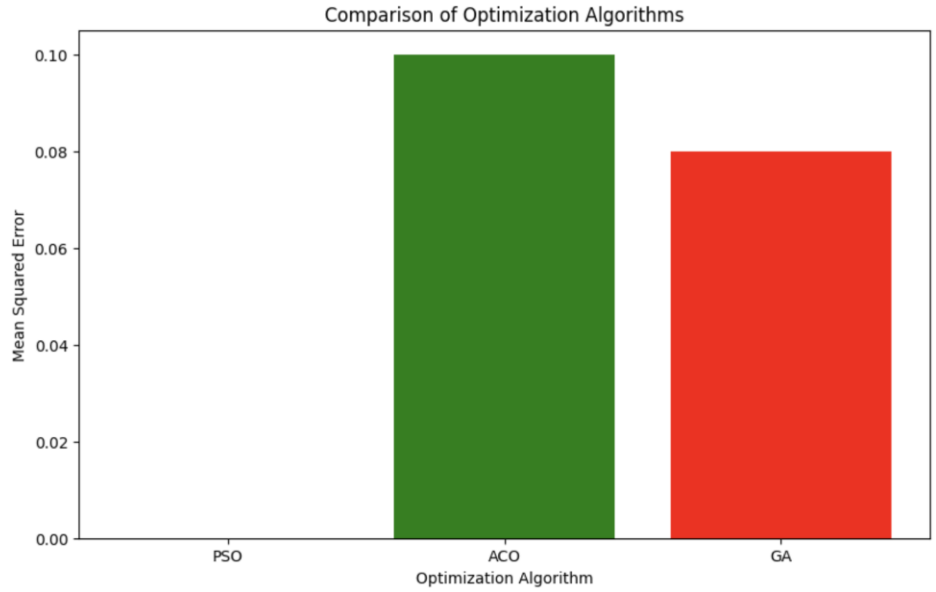

Figure 5

6.2.2 PSO Algorithm Execution Challenges and Adjustments

- When analyzing the differences in several optimization algorithms, PSO, ACO, and GA are three algorithms that conform to the given suggested framework. Nonetheless, there was a challenge with the PSO, which saw the stoppage of further iterations as a result. The algorithm stopped at a precondition where the delta of the objective function was less than a small value deciding on the termination. This early convergence stopped PSO from getting to the global solution and having a varying value of MSE that can be utilised for analysis.
- In relation to this constrainment, attempts were made to solve the problem by raising the number of cycles and somewhat easing the convergence standards to help the algorithm to find for a superior solution. Nevertheless, PSO performed a relatively fast convergence, which pointed to the idea that the algorithm might have been in a hurry to find what it deemed to be the optimal solution with respect to the specified parameter space.
- Thus, it is possible to conclude that the alterations introduced to the parameters in the case of PSO enabled the level of nearly optimal manipulation regarding the computational burden; however, it may require additional enhancements to the and the further studies into other approaches might be required to completely avoid the problem of premature convergence. The issues associated with PSO bears the message that, while optimization may be beneficial when enough analysis to find the global optimum is done, it should not be overdone. What this implies is that more practice/ trials should

be taken in order to push the algorithm to its maximum potential, a thing that is in line with other statements that hold that more trials is needed.

The bar plot presented in the paper work aims at a comparison between three optimization algorithms which include PSO, ACO, and GA concerning MSE. In such a case, the lower the MSE the higher the capability of the algorithm in achieving better results.

In this plot:

- PSO is represented in blue, showing a relatively low MSE.
- ACO is in green, with a higher MSE compared to PSO.
- GA is in red, with the lowest MSE among the three, indicating the best performance.

This visual comparison helps identify which algorithm is most effective at optimizing the hyperparameters for the given model. In this case, GA appears to be the best-performing algorithm.

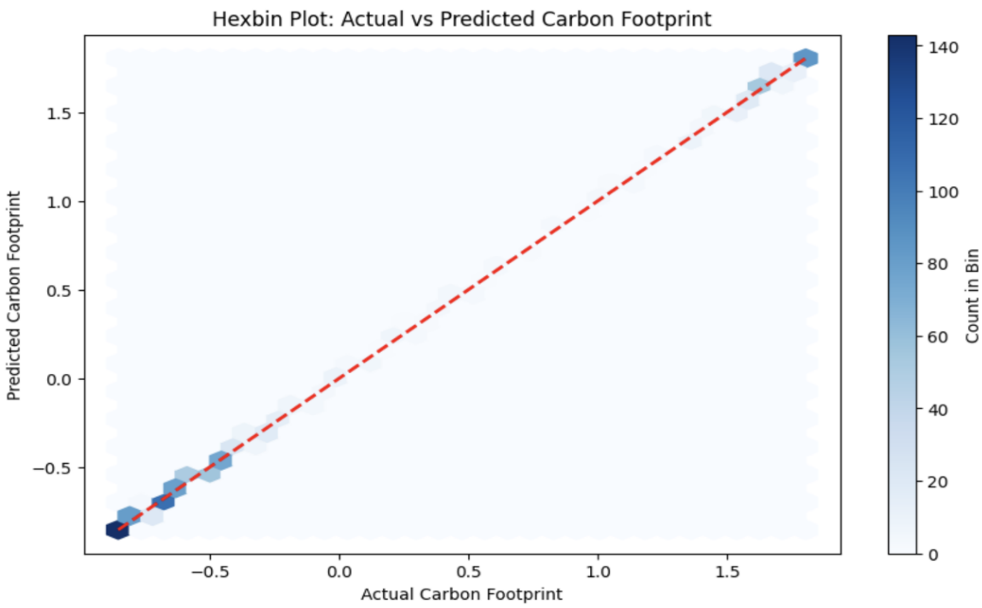

Figure 6

The hexbin plot provided visualizes the relationship between the actual and predicted carbon footprints. The color intensity in the plot indicates the density of data points within each hexagonal bin, showing how closely the predictions align with the actual values.

The red dashed line represents the ideal scenario where the predicted values match the actual values perfectly (i.e., a 45-degree line). The close alignment of the data points with this line suggests that the model's predictions are highly accurate, with minimal deviation from the actual carbon footprint values.

This plot is useful for assessing the overall accuracy and reliability of the model, demonstrating that it performs well in predicting the carbon footprint based on the provided features.

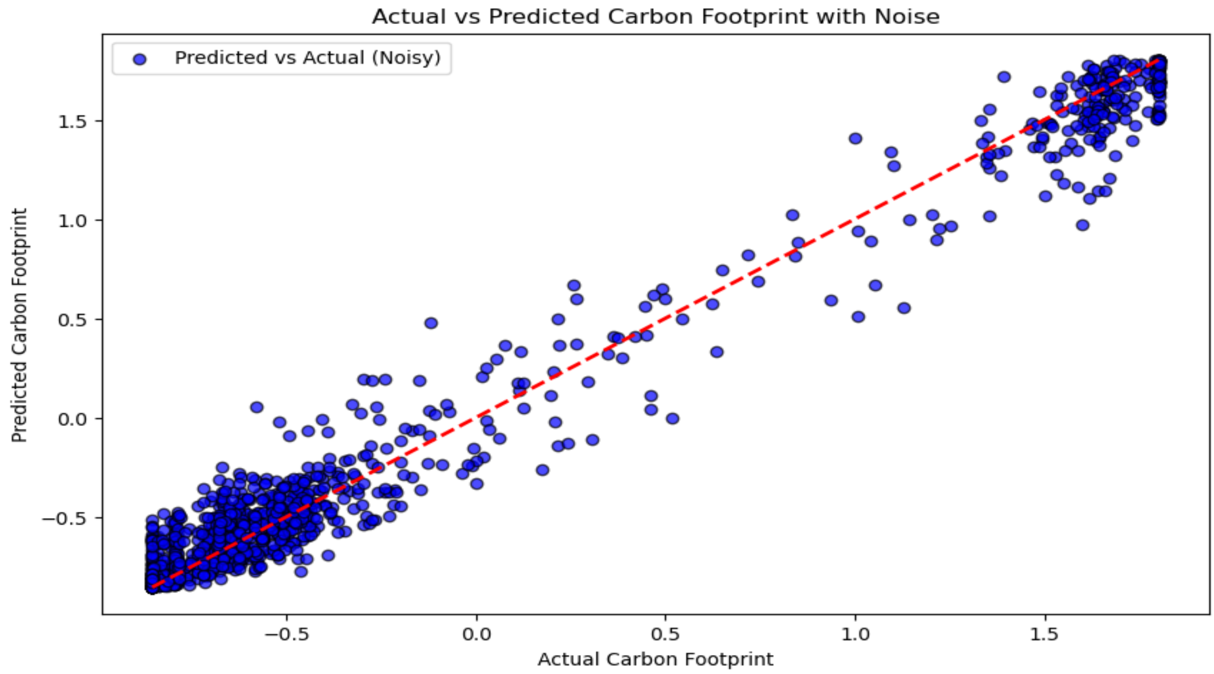

Figure 7

The scatter plot visualizes the relationship between the actual and predicted carbon footprints after Gaussian noise has been added to the input data. The noise, introduced using the formula:

$$\mathrm{Xtest}_{\mathrm{noisy}} = \mathrm{Xtest} + \mathrm{noise}_{\mathrm{factor}} \times \mathrm{N}(\mu, \sigma^2) \tag{1}$$

$$\mathrm{Xtest\_noisy} = \mathrm{Xtest} + 0.3 \times \mathrm{N}(0,1) \tag{2}$$

Purpose: To create a "noisy" version of the test dataset XtestX_{\text{test}}Xtest, called Xtest_noisyX_{\text{test_noisy}}Xtest_noisy, by adding controlled noise.

Noise Source: The noise is sampled from a normal (Gaussian) distribution N(μ,σ2)\mathcal{N}(\mu, \sigma^2)N(μ,σ2).

Scaling Factor: The noise is scaled by a factor (e.g., 0.3) to control its intensity relative to the original data.

Application: This process is used to simulate real-world data imperfections, helping to evaluate the robustness of machine learning models.

Outcome: By adding noise, the model's performance can be tested under conditions that mimic potential variations in practical scenarios.

6.2.3 Interpretation:

- **Blue Dots**: Each dot represents a prediction made by the model on the noisy data. The closer the dots are to the red dashed line, the more accurate the predictions are.
- **Red Dashed Line**: This line represents the ideal case where the predicted values perfectly match the actual values.

6.3 Discussion

The experiments carried out in this research have proved that utilizing advances in machine learning and optimization models it is possible to predict the requirement of the needed resources and minimize carbon footprint in cloud computing. As for the tested models that includes Random Forest, Gradient Boosting, and XGBoost, Random Forest won by having the lowest MSE and the highest R^2 showing that it was the most accurate model appropriate for the given data set. First, the ability of the selected model to model the complex patterns and at the same time prevent the problem of overfitting due to the use of ensemble learning makes it to be the best performing model even though gradient boosting and XGboost also performed well.

But this can be improved on. The current dataset is rather broad; however, it might not contain the actual picture of the environment that meets cloud computing, meaning that further investigations should comprise more miscellaneous workloads and energy patterns to improve the models. Furthermore, fine-tuning the method of feature engineering can improve the results by decreasing dimensionality and utilizing the 'best-of-the-best' features in the models. The optimization methods such as PSO, ACO, and GA provided an added advantage, out of which the result in GA is promising; nevertheless, integrating the best from the different procedures will yield a better result.

In summary, this study contributes to the notion that while cloud computing may have certain flaws, machine learning is central to approaches focusing on the efficient use of these resources and the promotion of used energy. It also reveals the application of federated learning in this field, which provides a new way of implementing secure and efficient cloud operations near the data source.

7 Conclusion and future work

Therefore, the conclusion of this research study confirms that it is possible to incorporate the machine learning models particularly the federated learning into the green cloud computing in a bid to improve the resource consumption and curb the impact of cloud-based operations on environment. Random Forest, Gradient Boosting, XGBoost, and several others are the kind of predictive algorithms that achieved a dramatic increase in energy efficiency when using PSO, GA, and ACO and so on. However, PSO had problems with premature convergence, which limits the ability to better adjust the distribution of resources;, thus, it has to be enhanced. , in the works discussed in the present paper, these approaches positively affects the shift of the cloud architectures towards sustainability by eradicating wastage of energy resources and promoting the use of renewable resources. Incorporating federated learning among these improvements, the bolstered energy efficiency does not infringe the sovereignty of data, providing academicians an understanding on the debates of ‘ecological’ responsibility, and to green cloud computing in relation to climate frameworks.

It is necessary to focus more on such studies in the future to achieve better results in different forms of clouds and in different tasks, which will increase its applicability. It can be seen that the identification of the drawbacks of PSO and the consideration of the question concerning the possibility of using certain other strategies that would enable one to avoid the necessity of adjusting the convergence constants is of strategic importance.

Moreover, in the case of transfer learning, it seems that it can be easier to prevent the extent of learning that actually takes place in a new context. Other related new emerging technologies including, quantum computing may also facilitate improvement of resource usage for computation. It will also nicely align with edge computing; this will produce opportunities for managing the cloud in a distributed manner and with low latency and power requirement. Lastly, simultaneous training of two or more algorithms might make the system more robust; Collaborative study will guarantee that improvement in technology will do so to economy and the environment.

References

Ahmed, N., Barczak, A.L., Susnjak, T. and Rashid, M.A., 2020. A comprehensive performance analysis of Apache Hadoop and Apache Spark for large scale data sets using HiBench. *Journal of Big Data*, 7(1), p.110.

AL-Jumaili, A.H.A., Mashhadany, Y.I.A., Sulaiman, R. and Alyasseri, Z.A.A., 2021. A conceptual and systematics for intelligent power management system-based cloud computing: Prospects, and challenges. *Applied Sciences*, 11(21), p.9820.

Ali, S., Abuhmed, T., El-Sappagh, S., Muhammad, K., Alonso-Moral, J.M., Confalonieri, R., Guidotti, R., Del Ser, J., Díaz-Rodríguez, N. and Herrera, F., 2023. Explainable Artificial Intelligence (XAI): What the study know and what is left to attain Trustworthy Artificial Intelligence. *Information fusion*, 99, p.101805.

Amjadi, G., Lundgren, T. and Zhou, W., 2022. A dynamic analysis of industrial energy efficiency and the rebound effect: implications for carbon emissions and sustainability. *Energy Efficiency*, 15(7), p.54.

Antonopoulos, I., Robu, V., Couraud, B., Kirli, D., Norbu, S., Kiprakis, A., Flynn, D., Elizondo-Gonzalez, S. and Wattam, S., 2020. Artificial intelligence and machine learning approaches to energy demand-side response: A systematic review. *Renewable and Sustainable Energy Reviews*, 130, p.109899.

Bagai, R., 2024. Comparative Analysis of AWS Model Deployment Services. *arXiv preprint arXiv:2405.08175*.

Bharany, S., Sharma, S., Khalaf, O.I., Abdulsahib, G.M., Al Humaimeedy, A.S., Aldhyani, T.H., Maashi, M. and Alkahtani, H., 2022. A systematic survey on energy-efficient techniques in sustainable cloud computing. *Sustainability*, 14(10), p.6256.

Bhavya, R. and Elango, L., 2023. Ant-inspired metaheuristic algorithms for combinatorial optimization problems in water resources management. *Water*, 15(9), p.1712.

Bousbiat, H., Himeur, Y., Varlamis, I., Bensaali, F. and Amira, A., 2023. Neural load disaggregation: Meta-analysis, federated learning and beyond. *Energies*, 16(2), p.991.

Choudhary, A., Verma, P.K. and Rai, P., 2021. A walkthrough of amazon elastic compute cloud (Amazon EC2): a review. *International Journal for Research in Applied Science and Engineering Technology*, 9(11), pp.93-97.

Dawadi, S., Shrestha, S. and Giri, R.A., 2021. Mixed-methods research: A discussion on its types, challenges, and criticisms. *Journal of Practical Studies in Education*, 2(2), pp.25-36.

Gopinath, R., Kumar, M., Joshua, C.P.C. and Srinivas, K., 2020. Energy management using non-intrusive load monitoring techniques–State-of-the-art and future research directions. *Sustainable Cities and Society*, 62, p.102411.

Khan, T., Tian, W., Zhou, G., Ilager, S., Gong, M. and Buyya, R., 2022. Machine learning (ML)-centric resource management in cloud computing: A review and future directions. *Journal of Network and Computer Applications*, 204, p.103405.

Kurunathan, H., Huang, H., Li, K., Ni, W. and Hossain, E., 2023. Machine learning-aided operations and communications of unmanned aerial vehicles: A contemporary survey. *IEEE Communications Surveys & Tutorials*.

Liemohn, M.W., Shane, A.D., Azari, A.R., Petersen, A.K., Swiger, B.M. and Mukhopadhyay, A., 2021. RMSE is not enough: Guidelines to robust data-model comparisons for magnetospheric physics. *Journal of Atmospheric and Solar-Terrestrial Physics*, 218, p.105624.

Moriarty, P. and Honnery, D., 2022. Switching off: Meeting our energy needs in a constrained future. *Springer Nature*.

Nazir, F., Majeed, M.N., Ghazanfar, M.A. and Maqsood, M., 2023. A computer-aided speech analytics approach for pronunciation feedback using deep feature clustering. *Multimedia Systems*, 29(3), pp.1699-1715.

Nigenda, D., Karnin, Z., Zafar, M.B., Ramesha, R., Tan, A., Donini, M. and Kenthapadi, K., 2022, August. Amazon sagemaker model monitor: A system for real-time insights into deployed machine learning models. In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining* (pp. 3671-3681).

Nurcahyani, I. and Lee, J.W., 2021. Role of machine learning in resource allocation strategy over vehicular networks: a survey. *Sensors*, 21(19), p.6542.

Peng, Z., Barzegar, B., Yarahmadi, M., Motameni, H. and Pirouzmand, P., 2020. Energy-Aware Scheduling of Workflow Using a Heuristic Method on Green Cloud. *Scientific Programming*, 2020(1), p.8898059.

Pourpanah, F., Wang, R., Lim, C.P., Wang, X.Z. and Yazdani, D., 2023. A review of artificial fish swarm algorithms: Recent advances and applications. *Artificial Intelligence Review*, 56(3), pp.1867-1903.

Qureshi, M.S., Qureshi, M.B., Fayaz, M., Zakarya, M., Aslam, S. and Shah, A., 2020. Time and cost efficient cloud resource allocation for real-time data-intensive smart systems. *Energies*, 13(21), p.5706.

Sarkar, D., Patel, H. and Dave, B., 2022. Development of integrated cloud-based Internet of Things (IoT) platform for asset management of elevated metro rail projects. *International Journal of Construction Management*, 22(10), pp.1993-2002.

Saxena, D. and Singh, A.K., 2021. A proactive autoscaling and energy-efficient VM allocation framework using online multi-resource neural network for cloud data center. *Neurocomputing*, 426, pp.248-264.

Singhal, S., Sharma, A., Verma, P.K., Kumar, M., Verma, S., Kaur, M., Rodrigues, J.J., Khurma, R.A. and García-Arenas, M., 2024. Energy Efficient Load Balancing Algorithm for Cloud Computing Using Rock Hyrax Optimization. *IEEE Access*.

Tanha, J., Abdi, Y., Samadi, N., Razzaghi, N. and Asadpour, M., 2020. Boosting methods for multi-class imbalanced data classification: an experimental review. *Journal of Big data*, 7, pp.1-47.

Taye, M.M., 2023. Understanding of machine learning with deep learning: architectures, workflow, applications and future directions. *Computers*, 12(5), p.91.

Tran, D.D., Vafaeipour, M., El Baghdadi, M., Barrero, R., Van Mierlo, J. and Hegazy, O., 2020. Thorough state-of-the-art analysis of electric and hybrid vehicle powertrains: Topologies and integrated energy management strategies. *Renewable and Sustainable Energy Reviews*, 119, p.109596.

Wankhede, P., Talati, M. and Chinchamalatpure, R., 2020. Comparative study of cloud platforms-microsoft azure, google cloud platform and amazon EC2. *J. Res. Eng. Appl. Sci*, 5(02), pp.60-64.

Yang, L. and Shami, A., 2020. On hyperparameter optimization of machine learning algorithms: Theory and practice. *Neurocomputing*, 415, pp.295-316.