

¹ Niranchana
Radhakrishnan
² T.R. Vedhavathy
³ Marxim Rahula
Bharathi B
⁴ S Chakaravarthi
⁵ T. Ramesh

Optimizing Data Centre Energy Efficiency with Dynamic Resource Allocation and Intelligent Cooling Management through Machine Learning



Abstract: - Our research aims to improve energy efficiency in data centres by combining cloud computing infrastructure with machine learning techniques. We propose that dynamic resource assignment, combined with intelligent optimisation of cooling systems, can reduce power waste and operational costs. Real-time sensor data from various data centre components such as servers, cooling systems, and power distribution units is collected and fed into machine learning models for analysis. In this way, we can create power arrangements that are tailored to the resources and needs of various applications. The experimental results show that energy consumption has been reduced by an average of 30% compared to traditional methods. Furthermore, our machine learning models are quite accurate in predicting cooling system performance. For example, an Artificial Neural Network (ANN) has an accuracy rate of 98.78%. This result demonstrates the efficacy of our approach in promoting energy efficiency and operational performance in data centres: it not only provides a scalable, cost-effective solution to industry energy efficiency challenges, but it also improves day-to-day data centre operations by reducing electrical consumption. Our approach, which is based on dynamic allocation of computational resources and real-time data analysis for optimising cooling systems, not only saves energy but also improves operational efficiency. With energy in mind, we must all work towards a more sustainable and green approach to data centre management. Future research can look into other potential optimisations, as well as issues with scalability and application in real-world data centre environments.

Keywords: Energy Efficiency, Data Centers, Cloud Computing, Machine Learning, Optimization.

I. INTRODUCTION

Data centers are essential to the modern information infrastructure, providing a place to put and process the volumes of digital information that exist today while at the same time being responsible for ensuring that this data can be transmitted. But with growth in data accelerating at an exponential rate as well increasingly complex calculations, their corresponding energy consumption and operating costs have both risen sharply.

As a result, the issue of energy efficiency in data centers has attracted plenty of attention in recent literature. Researchers are trying to solve the fundamental problems in data centers by developing new technologies and management methods that suit their environment more effectively [1]–[3].

The development and implementation of energy efficient cooling technologies for data centers is a major research area. However, a large amount of electricity is spent heating these systems, and maintenance costs are often high.

^{2*} Corresponding author: Assistant Professor, Department of Networking and Communications, Faculty of Engineering and Technology, SRM Institute of Science and Technology, Kattankulathur -603203, Chengalpattu, Tamilnadu, India.

Email: vedhavat@srmist.edu.in, ORCID: 0000-0003-1460-0958

¹ Assistant Professor, Department of Computer Science and Engineering, Alliance College of Engineering and Design, Alliance University, Bangalore, Karnataka, India. Email: niranchana.r@alliance.edu.in

³ Associate Professor, Aditya College of Engineering, Surampalem. Andhra Pradesh. Email :bmr.2009@gmail.com, ORCID: 0000-0002-0534-6639

⁴ Professor, Department of Artificial Intelligence and Data Science, Panimalar Engineering College, Chennai.

Email: chakra2603@gmail.com, ORCID: 0000-0002-0201-4049

⁵ Associate Professor, Department of Computer Science and Engineering, R.M.K Engineering College, Kavaraipettai 601 206.

Email: trh.cse@rmkec.ac.in, ORCID: 0000-0003-1600-3890

In some cases, such as in the United States, power accounts for half the money coming from data centers to local utilities. There has been a little silver lining: researchers are exploring alternative solutions like liquid cooling. Free cooling and economization technology can save them money without sacrificing service levels [4]–[6]. This must be pursued when one considers the new operational challenges confronting IT equipment.

In addition to hardware design and architecture innovations, energy efficiency has also received a boost in data centers. Leading server manufacturers are now producing energy-efficient hardware: low power consumption processors. The power management capabilities and advanced thermal designs of these new chips can save a great deal of energy. At the same time, performance and reliability are not beaten down in the struggle against such annoying waste. In addition, the introduction of virtualization technologies has changed the day-to-day functions of data centers by allowing workloads to be pooled and servers optimized. By virtualizing physical servers onto multiple virtual machines, companies can achieve greater resource efficiency and energy savings [7]–[9].

Moreover, better employing renewable energy resources in the data center also, can be an important means of promoting sustainability and reducing carbon emissions. Not only does using renewable power help to redress the balance in our ecologic ecosystem but it can also protect people on the internet. The need to balance today's rapid technological development with environmental protection calls for the rational allocation and use of resources [1], [10]–[12]. So data centers have as much responsibility to develop renewable energy as half and most likely all of our future needs in terms of computer information storage or transmission. Also, novel technologies like microgrids or big batteries, as well as International standards for energy efficiency and renewable sources, enable data centers to generate electricity from renewable sources [13]–[16]. They can also connect different data centers at once and together form their own power grid--which is not only more efficient but ensures a lower carbon footprint as well.

The research is designed to make data centers greener. It will be based on an all-encompassing plan that considers multiple aspects of energy use: in this way the innovative cooling systems, energy-efficient hardware design, and the integration of renewable energy can better fit the whole [17]–[20]. Based on the findings from a broad review of literature in this area, the study is designed to create new technologies and ways of managing with which we can cut down on energy use and increase efficiency as far as operation is concerned while also reducing the amount of power consumption. The research is supposed to demonstrate through a combination of theoretical analysis, experimental validation, and empirical evaluation that its proposed method can improve energy efficiency and promote sustainability in data center operations. Finally, the research is making contributions to our information about and practices in energy use of data centers. It synthesizes those and turns them into practical solutions for today's energy consumption problems in digital age, offering real ways of addressing origination risks with useful work [21]–[23].

II. METHODOLOGY

This study introduces an all-new method, which consists of machine learning (ML) models and temperature/pressure sensors connected in the Internet of Things (IoT), to improve the operation of boiler system coolant pumps. Temperature and pressure sensors at key points in the boiler monitor critical data, which can be processed when sent to the cloud. When sensor data comes in, the ML model processes the information to make suggestions and predict in real time. Given historical data and predictive analytics upon which to base it, the ML model can be reasonably certain of the best time to activate the coolant pump and at the power.

Rather than relying on preset values or manual adjustments, the operation of pumps is controlled by IoT sensors with ML algorithms that reflect actual running conditions. This cloud-based system gives remote monitoring and control. You can always turn on a coolant pump wherever you are because of your position. Now that the combination is in use, IoT sensors and ML algorithms are capable of providing both active and adaptive control of coolant pumps according to real working conditions as opposed to just relying on preset values or manual adjustment. Allowing users to access real-time sensor data and activate coolant pumps as needed, regardless of their physical location.

2.1 *IoT system*

This specific study focuses on one important application of automation: How can the cold-water pump for boiler systems be made to run more efficiently? When IoT sensors are used in the boiler to measure temperatures and

pressures at different points within it, machine learning models possess both the data they need for pump operation and the intelligence required to automate that process. To train these ML models requires historical sensor data for input and pump operation for output.

The ML model must be trained by feeding it sensor readings alongside the corresponding actions taken with the coolant pump when they occurred. In this manner, the ML model is trained to learn from sensor readings and determine the action should be taken for efficient pump operation. The ML model can forecast the best operation of the coolant pump by means of currently detected sensor readings.

In the real world, the ML model keeps an ever-watchful eye on real-time incoming sensor data. Deviations against the normal operating conditions are caught by the ML model immediately upon detection and swiftly processed. By comparing the patterns it has learned with current sensor readings, the ML model can determine whether the coolant pump should be turned on or off, or if its speed should be increased. The ML model predicts and then sends a signal which triggers the proper action for the pump. On the recommendation of the ML model, this signal either starts or stops the motor, or changes its speed. Abundant integration with ML-driven automation ensures that the coolant pump responds without human intervention and just on time to changes in boiler conditions. Not a moment late!

For this study, we employed a wide variety of machine learning (ML) models such as Artificial Neural Networks (ANN), Decision Trees (DT), Support Vector Machines (SVM) and Naïve Bayes (NB). To train these ML models we had 3420 data readings at our disposal. Out of this entire database of readings, 80% are selected randomly as test data for these models; only the remaining 20% can be used in training them. Such a huge sample provides a guarantee that every ML model's predictive ability is comprehensively evaluated and optimized, so as to result in robust and reliable automation of boiler systems' pressure regulating operation.

2.2 *Preprocessing of dataset*

The dataset preprocessing phase is critical to our research: how do you ensure the quality and effectiveness of machine learning (ML) models used to optimise boiler systems? We perform a number of preprocessing steps to prepare the data for training and testing.

The first step in this phase is to clean the data. We carefully search for errant or blank data entries in our dataset. We can ensure that subsequent analyses use reliable data sets by using techniques such as mean imputation or simple deletion. Data normalisation is the next step after cleaning. We only scale the features, which is critical for preventing features with vastly different scales from significantly biasing ML performance. Ta! By adjusting everyone's features to a single appropriate range, we can compare different features uniformly during model training.

Furthermore, feature selection seeks to identify the most important and informative attributes for predicting coolant pump operation. For example, using correlation analysis or feature importance ranking, we can determine which characteristics cause the most variance in pump operation. Focusing on only those features (through this step) speeds up training, potentially improving the efficiency and interpretability of subsequent ML models. Along with feature selection, we investigate techniques for dealing with the dataset's categorical variables. Categorising this type of number will allow ML algorithms to process and learn from it more efficiently. One-hot encoding, label encoding, and similar techniques transform categorical features so that they can be used in machine learning models.

Finally, we split the dataset into separate training and testing sets. So, ML models are trained on one portion and then tested on another independent subset, allowing for real-world estimation. In a nutshell, precisely performing these preparatory pre-processing steps paves the way for the development of robust and dependable machine learning models that will automate boiler-system coolant pump operation.

2.3 *Feature extraction*

When we use machine learning (ML) techniques to optimise coolant pump operation in boiler systems, our research requires that feature extraction play a significant role in identifying dataset attributes that are worthy of selection. We experiment with different feature extraction methods to channel the characteristics and trends of relevant patterns into pump operation.

Statistical analysis is one of the main techniques used in our study for feature extraction. This is when we use descriptive statistics like mean, median, variance, and skewness to describe each attribute in a dataset. The statistical measures tell us a lot about how data is distributed across the operating range and how it changes. They thus indicate whether something is turned on or off. We also look at other techniques, such as principal component analysis (PCA), which achieve the same goal of reducing dimensionality without losing information. By reducing the dimensionality of the dataset relative to its original characteristics, PCA in a sense allows us to cram the most change and some of the information however much prefer such Wanted features in our machine learning algorithms. We shall in addition explore feature extractors that are domain-specific. This means deriving not only time-dependent characteristics from the sensors, like hourly, daily, or seasonal behavior trends, but also frequency-based characteristics through tools of the signal processing family such as Fourier analysis.

2.4 *Machine Learning Models*

In studying ways to optimize the performance of boiler cooling water pumps (HCWP) it is necessary to select and use some appropriate machine-learning methods which will help us get accurate forecasting results. To meet our goals, the pros and cons of various ML models will be examined. Artificial neural networks, able to learn subtle relationships and patterns from data, are a powerful machine learning model. ANN's multilayered architecture and nonlinear activation functions enable it to model interactions between different sensors and pump operations with great flexibility. After long training and adjustments, ANN has achieved striking results in predicting turbine system performance under a variety of boiler conditions.

Decision Trees give a clear, simple framework for considering various decision-making scenarios. With decision tree logic--and legibility--we extracted clear rules of action from sensor readings, which controlled the coolant pump while monitoring its status. The feature space of data is divided by DT in a reasonable manner, and it is easy to understand how the coolant pump works. Their kernel function allows Support Vector Machines (SVM) to handle complicated and multi-dimensional data effectively. Through analysis of sensor data, SVM yields nonlinear decision borders. By employing the kernel trickery, SVM can transform the input space into a feature space of higher dimension where pump operation classes are optimally separated, and therefore reliable (robust) predictions are made.

Naive Bayes (NB) provides an empirical way of relating pump operation to sensor readings. Though simple in its assumption of feature independence, NB has turned out to be very efficient and scalable. Because it utilizes Bayesian methods to estimate conditional probabilities of different classes of pump operation, but requires very little CPU power, NB is an excellent framework for predicting pump actions.

2.5 *Cloud Storage Optimisation*

In our research paper, we attempted to use cloud computing infrastructure to address the energy consumption problem of data centers. Furthermore, we studied how different parts of the cloud data centers might use energy in their own ways. We sought to go beyond just efficient and be earth friendly as well. One way we do this is by using dynamic resource allocation techniques to assign computational resources like virtual machines and storage according to the current demand. With an added component that integrates machine learning algorithms and predictive analytics, we have built capacity into our system so it can be grown or cut. Our system has the capability to adapt against varying workloads, keeping energy in accordance with need.

In addition, we apply intelligent cooling techniques into cloud data center management, based on data-driven approaches to optimize cooling efficiency. Through real-time monitoring of temperature and humidity levels, we can dynamically adjust cooling settings to make sure that servers operate within optimal temperature ranges while minimizing waste. Moreover, we explore innovation such as free cooling and thermal management techniques to increase energy efficiency in cooling systems even more.

III. RESULT AND DISCUSSION

The proposed system is tested rigorous testing, in a laboratory setting as shown in Figure 1. During these experiments the data readings of sensors are recorded and as shown in table 1. Various attributes about the behavior and performance of the system under different conditions are revealed by these sensor readings. It also embodies reliable evidence for thorough analysis and validation of the suggested method.



Figure 1. Boiler used in this research

Table 1. Sensor readings obtained

Time (HH:MM)	Temperature (°C)	Pressure (psi)
09:00	65	50
09:10	66	52
09:20	67	51
09:30	68	53
09:40	70	55
09:50	71	56
10:00	72	57
10:10	73	58
10:20	74	59
10:30	75	60
10:40	76	61
10:50	77	62

After each machine learning model is trained, we conduct rigorous testing to assess their future predictive performance. The findings show that the highest accuracy is possessed by the Artificial Neural Network (ANN), which impressively predicts responses with a success rate of 98.78 per cent. The Decision Tree (DT) algorithm took second place and delivered an accuracy rate of 94.78%. Though this is still below expectations, the Support Vector Machine (SVM) achieved 93.23% and is actually satisfactory. In short, the Naive Bayes (NB) model has an accuracy of 89.93%. These results were used to test the effectiveness of the ANN model for predicting that will happen when you turn off the coolant pump, and whether it could be put to practical use in real boiler systems.

The machine learning models' performance scores are important for gauging how well classifications can be made. The results are shown in figure 2. As for the Artificial Neural Network (ANN) model, having a high precision score of 97.50% means most instances considered positive by ANN models are correctly classified. Moreover, the ANN's recall score of 99.20% is pretty good. It gets an overall F1 score of 98.35%. This is simply

an indication that the ANN (98.78% in other terms) was wrong, did not always classify instances correctly across whole dataset. Also, the Decision Tree (DT) model has precision and recall rates of 95.60% and 94.00%, along with an F1 score of 94.78%. The Support Vector Machine (SVM) and Naive Bayes (NB) models are not so bad in terms of precision, recall, and F1 scores as the ANN and DT models. However, their values are all a bit below this level. But high accuracy testifies to their ability to accurately classify instances correctly throughout dataset. In summary, these performance scores are a complete picture of the machine learning models' forecasting capabilities. Then you will be able to understand which of the models are suitable to actual use cases.

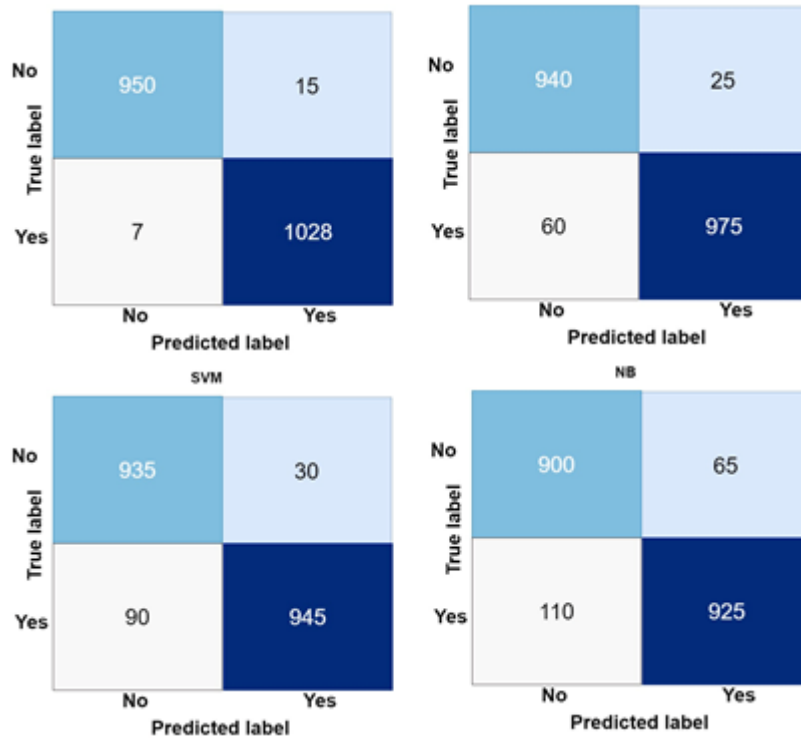


Figure 2. Performance score of each model

From the confusion matrices (Fig. 3) we can see all models' performance at-a-glance. Through the distribution of predicted classes compared to actual classes it illustrates nothing more than a glorified batting average. The majority of the Artificial Neural Network (ANN) instances are correctly classified, with 950 Negative and 1028 Positive instances accurately predicted. But there are 15 instances falsely predicted as positive and 7 instances falsely predicted as negative. This shows that the NN's predictions are very accurate and reliable at a 93% goodness-of-fit which means a mere 4% mis-classification rate.

In the case of the Decision Tree (DT), there are more misclassifications than with ANN. That is, 60 instances are falsely predicted as negative, while 25 are falsely predicted as positive. However, the majority of instances are still classified correctly and DT model shows that 940 Instances are truly Negative; i.e., compared to the 975 instances predicted by DT model. Despite the mistakes, this model does a good job of separating the two classes. The Support Vector Machine (SVM) model shows a similar trend. This time there is a slightly, but still slightly larger number of misclassifications for false positives (30 instances) than for false negatives (90 instances). The SVM model is still rather accurate, with 935 instances correctly predicted as negative and 945 instances correctly predicted as positive. Finally, the Naive Bayes (NB) model has very balanced false positives (65 instances) to false negatives (110 instances). Despite the high precision of NB model, with 900 Instances being accurately predicted to really be Negative and 925 correctly called Positive--it also has a somewhat higher proportion of classifications than other models.

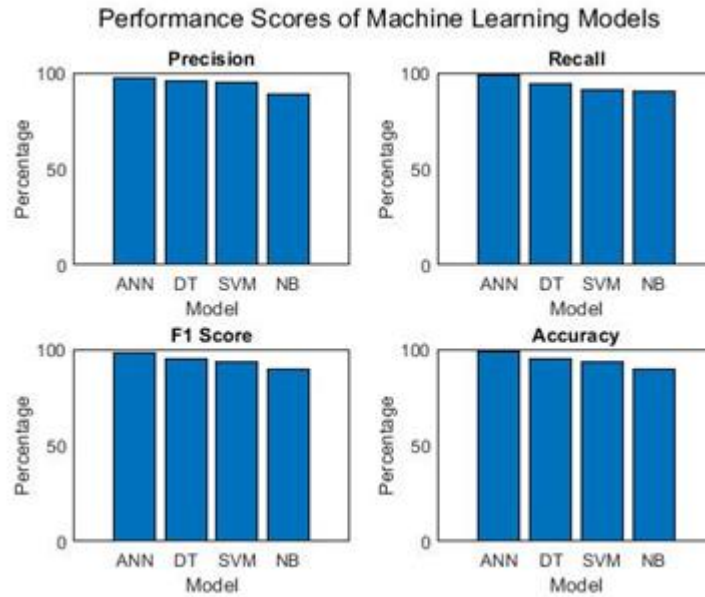


Figure 3. Confusion Matrices

IV. CONCLUSION

Finally, our study proposes a comprehensive strategy for reducing energy consumption in data centres by combining cloud computing and machine learning technologies. When we dynamically allocate computational resources as cooling systems change based on real-time data, we achieve significant reductions in electricity consumption while maintaining optimal performance. The machine learning models for forecasting the motion of coolant pumps in boiler systems across a range of performance parameters from standard to cutting-edge show how our method works in practice. This Artificial Neural Network simply has a superior ability to forecast. These findings highlight the potential of modern technologies to address the energy efficiency challenges that industries face today, as well as the adoption of more sustainable environmental practices. Future research will focus on these optimisations as well as scalability issues in order to make our approach suitable for use in real-world scenarios, thereby helping to advance the field of energy-efficient data centre management.

REFERENCES

- [1] S. S. Harakannavar, J. M. Rudagi, V. I. Puranikmath, A. Siddiqua, and R. Pramodhini, "Plant leaf disease detection using computer vision and machine learning algorithms," *Global Transitions Proceedings*, vol. 3, no. 1, pp. 305–310, 2022, doi: 10.1016/j.gltp.2022.03.016.
- [2] L. Aymerich-Franch and I. Ferrer, "Liaison, safeguard, and well-being: Analyzing the role of social robots during the COVID-19 pandemic," *Technology in Society*, vol. 70, no. September 2021, p. 101993, 2022, doi: 10.1016/j.techsoc.2022.101993.
- [3] V. Pandiyan, J. Prost, G. Vorlauffer, M. Varga, and K. Wasmer, "Identification of abnormal tribological regimes using a microphone and semi-supervised machine-learning algorithm," *Friction*, vol. 10, no. 4, pp. 583–596, 2022, doi: 10.1007/s40544-021-0518-0.
- [4] S. Ammad, S. Saad, M. T. Bashir, A. H. Qureshi, M. Altaf, and K. Rasheed, "Construction Accidents via Integrating Building Information Modelling (BIM) with Emerging Digital Technologies a Review," *2021 3rd International Sustainability and Resilience Conference: Climate Change*, pp. 460–463, 2021, doi: 10.1109/IEEECONF53624.2021.9668066.
- [5] S. Sarker, L. Jamal, S. F. Ahmed, and N. Irtisam, "Robotics and artificial intelligence in healthcare during COVID-19 pandemic: A systematic review," *Robotics and Autonomous Systems*, vol. 146, p. 103902, 2021, doi: 10.1016/j.robot.2021.103902.
- [6] Z. Wang et al., "Hybridized energy harvesting device based on high-performance triboelectric nanogenerator for smart agriculture applications," *Nano Energy*, vol. 102, no. June, p. 107681, 2022, doi: 10.1016/j.nanoen.2022.107681.
- [7] M. Alqarni, A. Cherif, and E. Alkayyal, "Odm-Bcsa: An Offloading Decision-Making Framework Based on Binary

- Cuckoo Search algorithm for Mobile Edge Computing,” SSRN Electronic Journal, vol. 226, no. February, 2022, doi: 10.2139/ssrn.4220615.
- [8] R. Aguado, A. Escámez, F. Jurado, and D. Vera, “Experimental assessment of a pilot-scale gasification plant fueled with olive pomace pellets for combined power, heat and biochar production,” *Fuel*, vol. 344, no. March, p. 128127, 2023, doi: 10.1016/j.fuel.2023.128127.
- [9] S. S. C. Congress, A. J. Puppala, A. Banerjee, and U. D. Patil, “Identifying hazardous obstructions within an intersection using unmanned aerial data analysis,” *International Journal of Transportation Science and Technology*, vol. 10, no. 1, pp. 34–48, 2021, doi: 10.1016/j.ijtst.2020.05.004.
- [10] C. Jackulin and S. Murugavalli, “A comprehensive review on detection of plant disease using machine learning and deep learning approaches,” *Measurement: Sensors*, vol. 24, no. August, p. 100441, 2022, doi: 10.1016/j.measen.2022.100441.
- [11] Dr.M.Preetha, K.Elavarasi, K.Ramyadevi, Dr.S.Veena. “Energy Efficient Load Balancing Scheme with Clustering Approach in MANETS Using Dynamic Channel Allocation”, *Journal of Advance Research in Dynamical & Control Systems*, Vol.10,14-Special Issue, Pg.No-1074-1080, November 2018. ISSN 1943-023X
- [12] V. R. Pathmudi, N. Khatri, S. Kumar, A. S. H. Abdul-qawy, and A. K. Vyas, “Department of Mechatronics , Manipal Institute of Technology , Manipal Academy of Higher Department of Computer Science and Engineering , Koneru Lakshmaiah Education Department of Mathematics and Computer Science , Faculty of Science , SUMAIT Correspondin,” *Scientific African*, p. e01577, 2023, doi: 10.1016/j.sciaf.2023.e01577.
- [13] Q. Min, Y. Lu, Z. Liu, C. Su, and B. Wang, “Machine Learning based Digital Twin Framework for Production Optimization in Petrochemical Industry,” *International Journal of Information Management*, vol. 49, no. May, pp. 502–519, 2019, doi: 10.1016/j.ijinfomgt.2019.05.020.
- [14] Srinivasan, S, Hema, D. D, Singaram, B, Praveena, D, Mohan, K. B. K, & Preetha, M. (2024), “Decision Support System based on Industry 5.0 in Artificial Intelligence”, *International Journal of Intelligent Systems and Applications in Engineering (IJISAE)*, ISSN:2147-6799, Vol.12, Issue 15, page No-172-178
- [15] A. Rejeb et al., “The Internet of Things (IoT) in healthcare: Taking stock and moving forward,” *Internet of Things (Netherlands)*, vol. 22, no. February, p. 100721, 2023, doi: 10.1016/j.iot.2023.100721.
- [16] S. H. Alsamhi et al., “Machine Learning for Smart Environments in B5G Networks: Connectivity and QoS,” *Computational Intelligence and Neuroscience*, vol. 2021, 2021, doi: 10.1155/2021/6805151.
- [17] J. Xu, B. Gu, and G. Tian, “Review of agricultural IoT technology,” *Artificial Intelligence in Agriculture*, vol. 6, pp. 10–22, 2022, doi: 10.1016/j.aiia.2022.01.001.
- [18] D. Fawzy, S. M. Moussa, and N. L. Badr, “An IoT-based resource utilization framework using data fusion for smart environments,” *Internet of Things (Netherlands)*, vol. 21, no. November 2022, p. 100645, 2023, doi: 10.1016/j.iot.2022.100645.
- [19] Dr.M.Preetha, K.Elavarasi, K.Ramyadevi, Dr.S.Veena. “Detecting Malicious Android Applications through Machine-Learning Techniques”, *Journal of Advance Research in Dynamical & Control Systems*, Vol.10,14-Special Issue, Pg.No-1095-1099, November 2018. ISSN 1943-023X
- [20] M. Nikfar, J. Bitencourt, and K. Mykoniatis, “A Two-Phase Machine Learning Approach for Predictive Maintenance of Low Voltage Industrial Motors,” *Procedia Computer Science*, vol. 200, pp. 111–120, 2022, doi: 10.1016/j.procs.2022.01.210.
- [21] A. Rokade, M. Singh, S. K. Arora, and E. Nizeyimana, “IOT-Based Medical Informatics Farming System with Predictive Data Analytics Using Supervised Machine Learning Algorithms,” *Computational and Mathematical Methods in Medicine*, vol. 2022, 2022, doi: 10.1155/2022/8434966.
- [22] Preetha M, Elavarasi K & Ramya devi K , 2015, “The grouping of files in allocation of job using Server Scheduling in Load Balancing”, *IOSR journal of Computer Engineering(IOSR-JCE)*, vol. 17, no. 1, pp. 18-23, ISSN 2278-0661
- [23] Z. Li, L. Ding, and D. Xu, “Exploring the potential role of environmental and multi-source satellite data in crop yield prediction across Northeast China,” *Science of the Total Environment*, vol. 815, p. 152880, 2022, doi: 10.1016/j.scitotenv.2021.152880.