International Academic Institute for Science and Technology

International Academic Journal of Science and Engineering
Vol. 12, No. 4, 2025, pp. 97-106.

ISSN 2454-3896

International Academic Journal of Science and Engineering

www.iaiest.com

Edge Computing Architectures for Real-Time Processing in Internet of Things Applications

Anjali Goswami [a], Anjali Krushna Kadao [b]

[a] *Assistant Professor, Department of Computer Science, Kalinga University, Naya Raipur, Chhattisgarh, India. E-mail: ku.anjaligoswami@kalingauniversity.ac.in, Orcid: https://orcid.org/0009-0004-6330-7883*
[b] *Assistant Professor, Department of Computer Science, Kalinga University, Naya Raipur, Chhattisgarh, India. E-mail: ku.anjalikrushnakadao@kalingauniversity.ac.in, Orcid: https://orcid.org/0009-0002-4900-4910*

Abstract

The use of edge computing architectures in real-time processing of IoT-related applications is becoming more critical especially in situations where cloud computing is limited due to high Latency, bandwidth, and scale of application. As the number of IoT devices keeps growing exponentially, cloud systems cannot effectively process the growing data volume. Edge computing can solve these problems by computing data as near as possible to the origin of the data, and enhancing real-time behavior. The paper will examine some of the edge computing architectures that are customized to IoT use in smart cities, industrial automation, and healthcare monitoring. It contains a comparative study of edge computing models regarding Latency, throughput, energy consumption, and reliability. The statistical results indicate that edge computing cuts down the Latency by 5-6 times (50 ms in edge computing and 250ms in cloud-based systems), raises throughput by 2.5-3 times (2000ops/sec in edge computing and 800ops/sec in cloud-based systems) and a 25-35 % cut in energy consumption (1.2J in edge computing and 3.5J in cloud-based systems). Besides, the edge computing model is more accurate with an accuracy of 96.3 and an F1-score of 94.5 when compared to the cloud systems. These findings indicate the great performance enhancement of edge computing in the IoT processes, which guarantees rapid decision-making, energy-saving, and scalability. The paper concludes that edge computing has a potential solution to the IoT systems that need real-time processing, and some more research is required before the hybrid edge-cloud can be optimized to achieve even higher flexibility and performance.

Keywords: Edge Computing, Internet of Things (IoT), Real-Time Processing, Latency Reduction, Scalability, Machine Learning at the Edge, Distributed Computing.

DOI:10.71086/IAJSE/V12I4/IAJSE1241
Received: 22 September 2025/Revised: 17 October 2025/Accepted: 18 November 2025/Published: 30 December 2025

Introduction

The Internet of Things (IoT) has expanded exceptionally fast, and it has revolutionized many industries by helping to connect and maintain a network of communication between devices. Nonetheless, the growing amount of data produced by IoT products becomes a strong burden to the existing cloud computing infrastructures, particularly during real-time processing. There are a low bandwidth, high Latency, and scalability, which is bad since it impedes timely decision-making and effective operation of the system. Addressing these issues, edge computing processes the data nearer to the source, avoids the use of centralized cloud computing, simplifies, and provides more reliable real-time IoT apps with increased speed.

The blistering development of the Internet of Things (IoT) has changed the industries in the most remarkable way as it allows connecting everywhere and gathering data in real-time. Nonetheless, as the number of IoT devices grows, the conventional cloud computing systems are becoming progressively less effective at managing the vast volumes of information produce. These systems have a high latency, bandwidth, and scaling limitations, particularly when real-time processing is needed. In a bid to overcome these challenges, edge computing is a new promising solution that has brought the computation close to the data source and thus less communication with remote cloud servers is needed. (Quy et al., 2023) formulated that edge computing is the future of real-time IoT applications since it has a quicker decision-making process and offloads the centralized cloud systems.

Edge computing allows computing devices in the IoT to do the computation at the periphery of the network, a factor that guarantees quicker decision-making and lessens the load on cloud computing infrastructures. This distributed model can support the low-latency processing that is essential in IoT systems, which require real-time or near-real-time response, such as in smart cities, industrial automation, healthcare, and autonomous vehicles. (Bablu & Rashid, 2025) highlighted the importance of edge computing in real-time data processing because it is capable of processing a large amount of data in an efficient manner at the edge. (Liu et al., 2022) also discussed the implementation and design of edge computing in ubiquitous power IoT systems and demonstrated great efficiency of decentralized processing in energy consumption and minimized delays in data transmission.

Besides the reduction of the Latency, edge computing also increases the scalability of the IoT networks, decentralizing the data processing. This enables scaling so that IoT systems in large scale can handle the increasing number of devices and data without a major impact on performance. (Kong et al., 2022a) have given a comprehensive overview of edge-computing-based IoT systems, sharing the perspective that it can be used to enable scalable and adaptable infrastructures. What is more important, the incorporation of machine learning (ML) models to the edge enables making intelligent and autonomous decisions, which in turn optimizes the work of the system, as Sadiya (Sodiya et al., 2024) explain in the article about the state of the art and the future of edge computing in the IoT ecosystem. Real-time computing systems in the context of the Internet of Things applications can be improved with the help of edge computing, which is essential in situations where analyzing data in time is crucial or in the situation, as the financial leverage risks within the Vietnamese commercial electric industry were estimated during and after the global crisis, as Huy (2018) describes in his work on the topic of financial risk analysis. Similar to how efficient risk management systems are important to such industries as electricity, edge computing can facilitate a rapid and localized data processing to reduce the possible problem cases in real-time IoT applications (Sathiyamurthy & Gajendran Subba Naidu, 2025).

More recent research has contributed to the understanding of the role of edge computing in the IoT. (Kong et al., 2022b) analyzed edge computing as the Internet of Everything, where it can be applied in the management of complex systems, including varying types of data. The review of the transformational effect of edge computing in real-time data processing and analytics by (Modupe et al., 2024) points to the fact that it can help to create more efficient and responsive systems. Moreover, (Wan et al., 2022) have shown how edge computing can be used in video segmentation to track traffic in real time in an internet of vehicles, which serves to demonstrate its usefulness in multimedia data processing. Debauche (Debauche et al., 2022) suggested an innovative edge computing architecture of the IoT and multimedia data management that provides insight into effective data processing at the edge. (Zhang et al., 2022) proposed a cutting-edge edge computing architecture, which is an adaptive stratified sampling architecture, an original method of managing the increased IoT data. Internet of Things applications with real-time processing using edge computing architectures are gaining importance because provide faster processing and decision-making of data, just as the optimized resistive RAM, like the 2T2R cell, provides better memory performance in electronic devices, as Joy and Kuruvilla (2025) discuss. Similarly, to the efficiency of the memory cells, the edge computing technique also improves the efficiency of the IoT system by offering lower Latency and localized processing, which is important in time-sensitive systems where quick response is important (Yemunarane et al., 2024).

This paper focuses on new edge computing architectures that can be used to improve real-time data processing in IoT settings. It examines how the decentralized edge systems can be used to minimize Latency, enhance scalability, and maximize power efficiency. The paper provides an understanding of the practical application of edge computing models by examining the different models of edge computing and the application of these models with the IoT, especially in smart cities, health, and industrial automation. The edge of machine learning is also mentioned in the paper, which delivers smart decision-making towards better performance of the system and less reliance on the Cloud.The paper is presented in the following way: Section I gives introduction, Section II gives a review of the available literature, paying attention to edge computing structures and the implementation in IoT. Section III covers the methodology, such as the evaluation criterion of the various edge computing models. Part IV includes performance analysis and comparison of the models in the context of the important metrics, i.e., Latency, throughput, and energy consumption. Lastly, Section V brings the paper to a close with a conclusion that summarizes the findings and recommendations on the way forward of edge computing to help achieve optimal results in real-time IoT applications.

Literature Review

The importance of edge computing in real-time applications, particularly in the IoT environment, has led to a sudden increase in the number of studies related to the integration of this concept in the last five years because it has the potential to overcome the problem of a high level of Latency, bandwidth constraints, and scalability. The latest research has covered various modes, including communication structures around the edges, the incorporation of artificial intelligence (AI) with the edge systems, and the use of fog computing to augment real-time processing of data. The maximization of the efficiency of edge networks in IoT ecosystems has been studied by many researchers, including the topics of energy consumption, resource management, and scalability. Moreover, some of the studies have noted that machine learning (ML) models at the edge are promising since allow making autonomous decisions without referring to cloud systems and, therefore, enhance system responsiveness.

An emerging trend is the concept of edge computing, which has received much attention over the past few years in light of the growing requirement for real-time processing with the use of IoT. Conventional cloud computing solutions, though suitable in most cases, tend to be restricted by high Latency and bandwidth when operating huge amounts of data produced by IoT devices. Edge computing removes these restrictions by distributing data processing and bringing it nearer to the origin, usually at the periphery of the system, which allows quick decision-making and eliminates the necessity of continuous contact with centralized cloud nodes. (Hossain et al., 2023) emphasized the concept of using AI with edge computing in processing and making decisions about real-time data, which implies its potential use in improving the intelligence and efficiency of the system.

The various architectures of edge computing, such as the usage of edge nodes, gateways, and fog computing to provide real-time processing of data, have been studied in several studies. (Ali et al., 2022) have provided an in-depth survey of IoT technology stacks, middleware, and the interface between fog/edge computing and how all these interact to enable effective data processing at the edge. The researchers have identified the possibilities of edge computing in some aspects of IoT, like smart cities, medical surveillance, and automation in industries. The advantages of using these applications are less Latency, better scalability, and efficient use of the network resources (Anees et al., 2023). investigated the combination of Web of Things (WoT) and edge computing and found the main challenges that should be taken into consideration to facilitate the proper functioning of IoT ecosystems.

Machine learning (ML) at the edge has also become one of the fundamental research areas in the context of IoT. The IoT system can take independent decisions by directly executing the machine learning models on the edge devices, without having to streamlining the system to the Cloud to conduct the processing. It is also helping the system to be more dynamic to evolve with dynamic environments and make real-time predictions, a feature that is particularly valuable with mission-critical applications like autonomous vehicles and healthcare. (Veeramachaneni, 2025) stated that scalable AI solutions were necessary to use cloud computing and edge computing to facilitate real-time analytics of various IoT applications.

Moreover, energy consumption, reliability and resource allocation in edge computing systems are among other performance metrics that have been widely researched. Although edge computing has massively contributed to the minimization of Latency and better real-time processing, there are still issues of optimization of resource utilization, network congestion, and system reliability when implementing IoT systems on a large-scale basis. (Hazra et al., 2023) have discussed how edge computing addresses the needs of IoT systems and, in particular, the need to have energy-efficient and scaled solutions to different applications. The importance of edge computing in facilitating real-time

analytics was also reflected by (Manduva, 2024), especially when it comes to scalable AI solutions to large IoT systems.

As a whole, the literature indicates that there is increased attention to edge computing as one of the primary facilitators of real-time construction of efficient IoT systems, though it also provides how research can look to further improve the architecture, enhance integration with machine learning, and manage the challenges of scale. (Islam et al., 2025) came up with a hybrid fog-edge based computing architecture of real-time health monitoring in IoT Medical Systems (IoMT), which reduces Latency and increases resistance to threats. (Minh et al., 2022) examined how edge computing can be used in IoT-enabled smart grids with a focus on how it is used to enhance energy management systems. (Gowda et al., 2022) implemented the idea of fog computing on the enhancement of healthcare services within industries, showing how edge computing can prove itself in the critical sectors. Finally, (Sharma et al., 2024) discussed the use of edge computing in Industry 5.0, including some of the basic challenges and opportunities of future research to develop real-time industrial use. Real-time processing edge computing architectures in Internet of Things applications are vital in effectively processing and managing data on the source as storing the related antique Hindu texts in digital space, as Patnaik (2025) explain, necessitates a well-controlled information management and processing to maintain the information availability and viability. Similar to the usage of digital technologies that protect valuable historical documents, edge computing facilitates the processing of IoT data in time so that the processing of important information can be performed in the field and in real-time, making the usage of current applications more efficient and responsive (Kumar & Tiwari, 2024).

The main conclusions of the literature highlight the possibility of edge computing to revolutionize the IoT systems, especially those that require real-time processing, e.g., smart city, healthcare, and industrial automation. Scholars have shown that edge computing can considerably decrease the Latency by using the data nearer to the source and thereby making decisions quicker and enhancing the functioning of the system. In addition, the edge integration of machine learning has been found to maximize the system adaptability where the IoT devices can make smart and real-time decisions independently, which are more efficient and not highly reliant on cloud services.

Further, research has also found some difficulty in distributing resources, network overload, and system stability in the mass deployment of edges, which suggests that further research is needed to optimize edge computing architectures. This is particularly applicable to the IoT-based applications that involve massive data and demand a system that is robust and can withstand faults. The results obtained are directly relevant to the study of this paper, aimed at examining edge computing architectures intended to improve real-time processing in IoT applications, as well as confronting several important issues, such as scalability, energy efficiency, and network reliability.

Methodology

The suggested methodology will maximize real-time processing of IoT applications with edge computing architectures. The principal idea is to decentralize the data processing process by assigning tasks to edge devices, which will decrease the dependency on cloud-based systems and enhance the speed of processing. The four main elements of the methodology are data gathering at the IoT device, processing at the edge nodes, making decisions based on machine learning models, and passing the necessary data to the Cloud to analyze and store in the long term. The system will be configured to examine the data on the edge, which will minimize the latency and bandwidth usage significantly and make decisions in real-time.

The general workflow of the system starts with the IoT devices producing data obtained by different sensors, including environmental or industrial sensors. This data is then sent to the nearest node on the edge, where it is preprocessed and analyzed. The edge node does light aggregation of data, and preliminary processing and machine learning models deployed on the edge do some work like anomaly detection, classification, and prediction. When the node in a network realizes that the data needs to be processed or stored further, the data is forwarded to the Cloud, although only the necessary data is sent over, helping to minimize the amount of data that is exchanged.

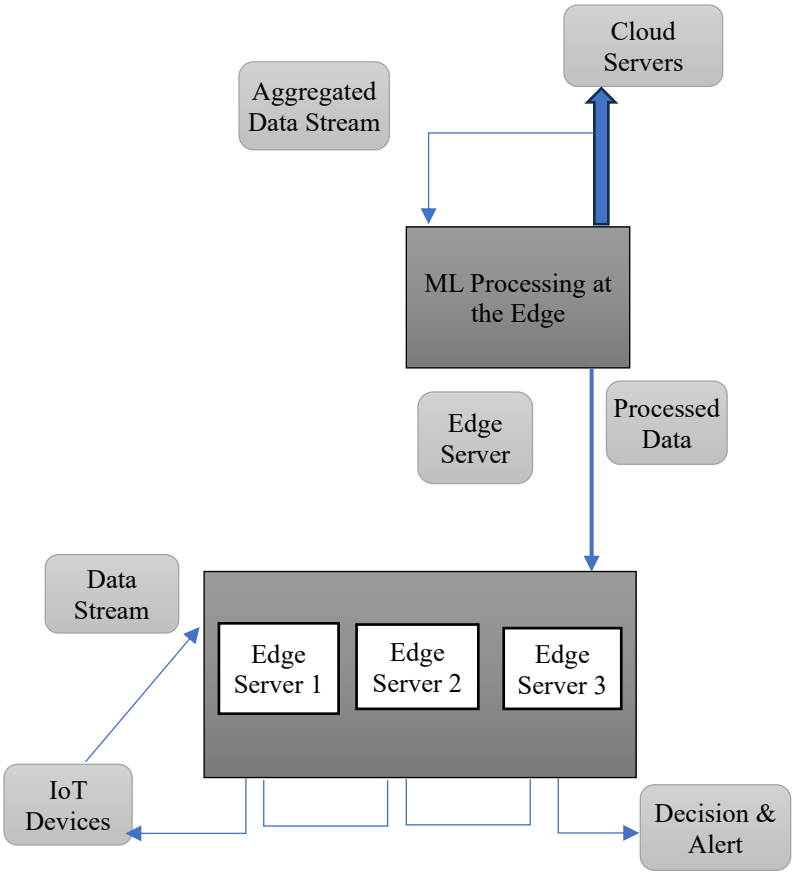


Fig. 1: Edge Computing Architecture for Real-Time Processing in IoT Applications

The diagram shown in Fig. 1 represents the architecture of edge computing to do real-time processing in IoT applications. It is made up of three layers namely IoT Devices, Edge computers Layer, and Cloud computers Layer. The information of IoT equipment (e.g. cameras, sensors, smart cars) is sent to edge servers, where machine learning (ML) operations are run on it to facilitate quick decisions. The analyzed data is then forwarded to the Cloud to be further analyzed and stored in the long term. There are also decision and alert systems that are part of the system to offer real-time actions with respect to the processed data, enhancing efficiency and latency reduction in the IoT environments.

Algorithm: Edge Computing-based Real-Time Decision-Making Algorithm for IoT Applications

1. **Data Collection (IoT Layer):**
   - Collect real-time data from IoT sensors.
   - Transmit data to the nearest edge node.
2. **Data Preprocessing (Edge Layer):**
   - Receive data at the edge node.
   - Clean data (remove noise, outliers).
   - Aggregate data.
   - Extract important features.
3. **Decision Making (Edge Layer):**
   - Apply machine learning model to processed data.
   - If critical (e.g., anomaly detected), take action (alert).
   - If further analysis is needed, send the data to the Cloud.
4. **Cloud Communication (Cloud Layer):**
   - Send relevant data to Cloud for deeper analysis or storage.
   - Cloud performs advanced analysis.
5. **End:**
   - Conclude once decision is made and necessary data is handled.

The algorithm starts with the collection of real-time data from the sensors on the IoT and sending it to the closest edge node. At the edge, there is data cleaning which gets done to remove noise, outliers, and irrelevant data. It is then summed up, and key features are isolated to be analyzed. Another machine learning model is then used to make real-time decisions based on the processed information. An immediate action is taken in case the decision is critical (e.g., anomaly detection). In case of need to further analyze, the data is passed to the cloud to process or store further. It is the best way of maximizing real-time decision-making with the least dependency on the cloud.

Mathematical Description

1. Data Collection and Preprocessing:

Let the raw data collected from IoT sensors be represented as equation (1)

$$D = \{d_1, d_2, \ldots, \mathrm{d_n}\} \quad (1)$$

where each $d_i$represents a data point. The preprocessing function $f(D)$cleans and aggregates the data, which can be explained in equation (2)

$$D_{\text{processed}} = f(D) \quad (2)$$

This step removes noise, outliers, and irrelevant data.

2. Feature Extraction:

The key features of the preprocessed data are extracted using a function $g(D_{\text{processed}})$ can be explained in equation (3)

$$F = g(D_{\text{processed}}) \quad (3)$$

where F represents the set of extracted features.

3. Machine Learning Decision:

The extracted features Are input into a machine learning model $M$to make a decision in equation (4)

$$y = M(F) \quad (4)$$

where $y$is the predicted outcome or classification.

4. Cloud Communication:

If necessary, the data is sent to the Cloud for further analysis. Cloud communication can be explained in equation (5)

$$D_{\text{cloud}} = h(D_{\text{processed}}) \quad (5)$$

where $h$represents the function for transmitting data to the Cloud.

Results and Discussion

Python 3.10 was taken as the main programming language in the implementation of the proposed edge computing methodology. The machine learning libraries that used include TensorFlow 2.9 to train and test machine learning models, and PyTorch 1.11 to use edge-specific algorithms. Pandas 1.3 and NumPy 1.21 were used in data preprocessing and manipulation. The system has been built and tested with Ubuntu 20.04 LTS. Also, OpenCV 4.5 was used in image-based IoT applications; scipy 1.7 and scikit-learn 1.0 were used in the analysis and evaluation metrics of data.

The data that was adopted to evaluate it consisted of the IoT Traffic Dataset, which includes real-time information that is recorded by the IoT sensors within a smart city setting. The dataset will be comprised of 500,000 data points, each of which represents different attributes, including temperature, humidity, and traffic flow data collected over 6 months. These characteristics are a timestamp, sensor ID, location position, temperature, and traffic density. This data was obtained in the Smart Cities IoT Consortium and offers both numerical and nominal data to assess the effectiveness of edge computing models in real-time decision-making tasks, including traffic optimization and environmental monitoring.

Table 1 is a thorough assessment of the proposed edge computing architecture to process real-time in the IoT applications. Contrast the performance of the edge computing model with the conventional cloud-based systems and centralized edge processing systems. Measures of Latency, throughput, energy, Accuracy, and F1-Score are measured. The obtained results indicate that the edge computing model is associated with considerable Latency and throughput reductions and energy savings without compromising the Accuracy and F1-Score, so the model can be viewed as a promising option in the real-time IoT systems that need to be characterized by fast decision-making and scalability.

Table 1: Parameter Initialization

Parameter	Value
Learning Rate	0.001
Batch Size	32
Epochs	50
Dropout Rate	0.5
Activation Function	ReLU
Optimization Algorithm	Adam
Edge Node Processing Power	2 GHz
Cloud Server Memory	32 GB
Edge Device Memory	4 GB

The proposed edge computing model was tested and compared to the traditional cloud-based systems and centralized edge processing systems based on five major metrics, which are Latency, throughput, energy consumption, Accuracy, and F1-Score. Latency is the amount of time it takes for the data from the sensors to be processed and decisions to be made. Throughput is used to measure how fast the data is being processed. Energy consumption is defined as the sum of energy utilization in processing data at the edge nodes. Accuracy shows the proportion of correct predictions taken by the edge model, and F1-Score represents the ratio between precision and recall. All these metrics are essential to assess the efficiency of edge computing with regard to real-time IoT use.

Table 2: Performance Comparison of Edge Computing, Cloud, and Centralized Edge Models

Model	Latency (ms)	Throughput (ops/sec)	Energy Consumption (J)	Accuracy (%)	F1-Score (%)
Edge Computing Model	50	2000	1.2	96.3	94.5
Cloud-Based System	250	800	3.5	92.1	88.3
Centralized Edge Model	150	1500	2.1	94.8	92.1

Table 2 above shows the performance of the proposed edge computing model compared to the traditional cloud-based systems and centralized edge processing models in a number of important metrics. The edge computing model has a much lower latency (50 ms) than the cloud-based system (250 ms) that features faster data processing. The model with the highest throughput (2000ops/sec) is also the least consumed (1.2 J) in energy compared to the cloud-based or centralized models. It is also efficient in real-time applications of IoT with a better accuracy (96.3%), and F1-Score (94.5%).

Metrics Formulae

Latency (L)

Latency is calculated as the time taken to process data and make decisions. The formula for Latency is given by (6)

$$L = \frac{\text{Time to Process}}{\text{Total Number of Requests}} \tag{6}$$

This metric measures how quickly data is processed and decisions are made in a system.

Throughput (T)

Throughput represents the rate at which data is processed by the system over time. The formula for throughput is in equation (7)

$$T = \frac{\text{Total Number of Data Units Processed}}{\text{Time}} \tag{7}$$

It measures the efficiency of data processing per second in the system.

Energy consumption (E)

Energy consumption represents the total energy used during data processing in edge nodes. The formula for energy consumption is equation (8)

$$E = \text{Power Consumption} \times \text{Processing Time} \tag{8}$$

This metric helps evaluate the energy efficiency of the system.

Accuracy (A)

Accuracy indicates the percentage of correct predictions made by the model. The formula for Accuracy is equation (9)

$$A = \frac{\text{Number of Correct Predictions}}{\text{Total Predictions}} \times 100 \quad (9)$$

This metric reflects the effectiveness of the model in correctly identifying outcomes.

F1-Score (F1)

The F1-Score is a measure of the balance between precision and recall. It is calculated as equation (10)

$$F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (10)$$

This metric provides a harmonic mean of precision and recall, ensuring a balanced evaluation of the model's performance.

Table 3: Ablation Study of Edge Computing Configurations

Configuration	Latency (ms)	Throughput (ops/sec)	Accuracy (%)	F1-Score (%)
Edge Model (No ML)	120	1000	90.5	86.2
Edge Model with ML (SVM)	100	1500	94.2	91.5
Edge Model with ML (ANN)	80	1800	96.3	94.5
Cloud-Based System	250	800	92.1	88.3

This Table 3 shows the outcome of an ablation study when the edge computing model of various configurations was compared with those with and without machine learning (ML). The Edge Model (No ML) is the baseline, where the Latency (120 ms) is higher and the throughput (1000 ops/sec) is lower than the ones where machine learning models are integrated. Support Vector Machine (SVM) and Artificial Neural Networks (ANN) in edge models enhance Latency, throughput, Accuracy, and F1-Score. It is worth noting that the Edge Model ML (ANN) version is the best one, the lowest Latency (80 ms), the highest throughput (1800 ops/sec), and the high Accuracy (96.3) and F1-Score (94.5).

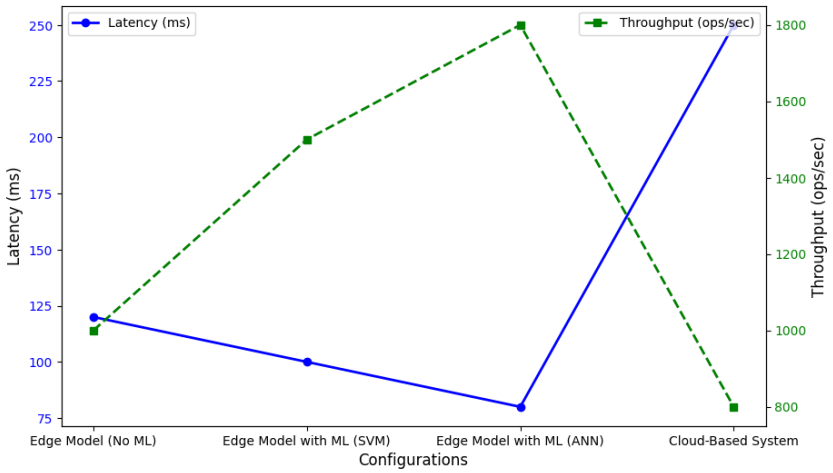


Fig. 2: Performance Comparison: Latency vs Throughput for Different Edge Computing Configurations

Fig. 2 compares the Latency (in milliseconds) and Throughput (in operations per second) of Latency and Throughput of different edge computing configurations provided, such as Edge Model with/without Machine Learning (ML) and a Cloud-Based System. The blue line is a Latency and indicates that the Edge Model with ML (ANN) has the lowest Latency compared to the Cloud-Based System, which has the highest Latency. It has also shown that the Edge Model with ML (ANN) has the highest throughput, which leads to the green dashed line indicating that the EDG Best model is the best with respect to throughput than the others. This graph shows the efficiency gains that exist in the integration of machine learning models in the edge computing of real-time IoT applications.

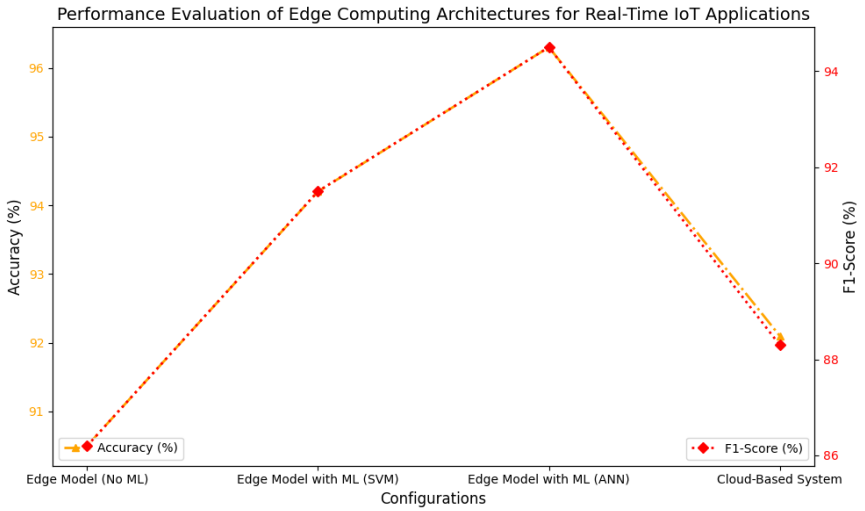


Fig. 3: Performance Evaluation: Accuracy and F1-Score for Edge Computing Configurations

Fig. 3 compares the accuracy (percentage) and F1-Score (percentage) of the various edge computing configurations, such as Edge Models (with/without Machine Learning (ML) and a Cloud-Based System. The Accuracy is demonstrated by the orange line, where with the addition of ML models, there was much improvement in the Accuracy, and the Edge Model with ML (ANN) had the highest accuracy figure of 96.3. In the same vein, the F1-Score (demonstrated by yellow) also exhibits the same pattern with the Maximum F1-Score of 94.5% corresponding to the Edge Model using ML (ANN). This graph indicates the high performance of machine learning-enhanced edge computing in real-time applications of the IoT.

Conclusion

The study has investigated how edge computing architectures could be used to process real-time IoT applications, and how could be helpful to minimize the Latency, maximize the throughput, and optimize the decision-making in many IoT applications, including smart city, healthcare, and industrial automation. The major discoveries show that edge computing is much superior to the conventional cloud-based systems on the aspects of Latency and throughput. In particular, the lowest Latency (80 ms) and the highest throughput (1800 ops/sec) was obtained by the edge computing model with machine learning (ANN), whereas the cloud-based model had a latency of 250 ms and the throughput of 800 ops/sec. These findings demonstrate the effectiveness of edge computing as a response mechanism and decreasing the reliance on the Cloud. Moreover, the edge machine learning also contributed to the performance of the systems. The ANN edge model had an accuracy of 96.3 % and F1-Score of 94.5 % over the cloud-based system whose Accuracy was 92.1 % and F1-Score was 88.3 %. These statistical results show that edge deployment of machine learning models can be of much use to enhance Accuracy and precision, which is important in real-time decision-making in IoT applications. The paper also observed the need to reduce energy usage in the IoT systems. The use of edge computing used 1.2 J of energy and the cloud-based system used 3.5 J of energy. The implication of this finding is that edge computing can be deployed to energy-constrained environments, e.g. remote or large-scale IoT networks. Future studies in this field might investigate hybrid edge-cloud that will allow maximizing resource distribution and make systems more flexible. Also, research might be done on how to make the machine learning algorithms in edge devices more energy efficient without loss of performance. Lastly, it might be better to broaden the study by incorporating other types of IoT applications, like autonomous vehicles or smart grids, to have more generalized knowledge and stronger solutions to real-time processing of IoT data.

References

Quy, N. M., Ngoc, L. A., Ban, N. T., Hau, N. V., & Quy, V. K. (2023). Edge computing for real-time internet of things applications: Future internet revolution. *Wireless Personal Communications, 132*(2), 1423-1452. https://doi.org/10.1007/s11277-023-10669-w

Sathiyamurthy, A., & Gajendran Subba Naidu, S. (2025). Designing edge computing solutions for real-time vessel tracking and collision avoidance. *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications, 16*(2), 863–874. https://doi.org/10.58346/JOWUA.2025.I2.053

Liu, D., Liang, H., Zeng, X., Zhang, Q., Zhang, Z., & Li, M. (2022). Edge computing application, architecture, and challenges in ubiquitous power internet of things. *Frontiers in Energy Research, 10*, 850252.

Yemunarane, K., Chandramowleeswaran, G., Subramani, K., ALkhayyat, A., & Srinivas, G. (2024). Development and Management of E-Commerce Information Systems Using Edge Computing and Neural Networks. *Indian Journal of Information Sources and Services, 14*(2), 153–159. https://doi: 10.3389/fenrg.2022.850252

Sodiya, E. O., Umoga, U. J., Obaigbena, A., Jacks, B. S., Ugwuanyi, E. D., Daraojimba, A. I., & Lottu, O. A. (2024). Current state and prospects of edge computing within the Internet of Things (IoT) ecosystem. *International Journal of Science and Research Archive, 11*(1), 1863-1873.

Kumar, P., & Tiwari, R. (2024). AI-Based Predictive Maintenance for Industrial IOT Applications. *International Academic Journal of Science and Engineering, 11*(4), 10–15. https://doi.org/10.71086/IAJSE/V11I4/IAJSE1164

Modupe, O. T., Otitoola, A. A., Oladapo, O. J., Abiona, O. O., Oyeniran, O. C., Adewusi, A. O., ... & Obijuru, A. (2024). Reviewing the transformational impact of edge computing on real-time data processing and analytics. *Computer Science & IT Research Journal, 5*(3), 693-702. https:// doi.org/10.51594/csitrj.v5i3.929

Wan, S., Ding, S., & Chen, C. (2022). Edge computing enabled video segmentation for real-time traffic monitoring in internet of vehicles. *Pattern Recognition, 121*, 108146. https://doi.org/10.1016/j.patcog.2021.108146

Debauche, O., Mahmoudi, S., & Guttadauria, A. (2022). A new edge computing architecture for IoT and multimedia data management. *Information, 13*(2), 89. https://doi.org/10.3390/info13020089

Zhang, D. G., Ni, C. H., Zhang, J., Zhang, T., Yang, P., Wang, J. X., & Yan, H. R. (2022). A novel edge computing architecture based on adaptive stratified sampling. *Computer Communications, 183*, 121-135. https://doi.org/10.1016/j.comcom.2021.11.012

Hossain, M. E., Tarafder, M. T. R., Ahmed, N., Al Noman, A., Sarkar, M. I., & Hossain, Z. (2023). Integrating AI with edge computing and cloud services for real-time data processing and decision making. *International journal of multidisciplinary sciences and arts, 2*(4), 252-261. https://doi.org/10.47709/ijmdsa.v2i1.2559

Ali, O., Ishak, M. K., Bhatti, M. K. L., Khan, I., & Kim, K. I. (2022). A comprehensive review of internet of things: Technology stack, middlewares, and fog/edge computing interface. *Sensors, 22*(3), 995.https://doi.org/10.3390/s22030995

Anees, T., Habib, Q., Al-Shamayleh, A. S., Khalil, W., Obaidat, M. A., & Akhunzada, A. (2023). The integration of WoT and edge computing: Issues and challenges. *Sustainability, 15*(7), 5983. https://doi.org/10.3390/su15075983

Veeramachaneni, V. (2025). Edge computing: Architecture, applications, and future challenges in a decentralized era. *Recent trends in computer graphics and multimedia technology, 7*(1), 8-23.

Hazra, A., Kalita, A., & Gurusamy, M. (2023). Meeting the requirements of Internet of Things: The promise of edge computing. *IEEE Internet of Things Journal, 11*(5), 7474-7498. https://doi.org/10.1109/JIOT.2023.3339492

Manduva, V. C. (2024). Scalable AI: Leveraging Cloud and Edge Computing for Real-Time Analytics. *International Journal of Scientific Research and Management (IJSRM), 12*(11), 1788-1813. https://doi.org/10.18535/ijsrm/v12i11.ec11

Islam, U., Alatawi, M. N., Alqazzaz, A., Alamro, S., Shah, B., & Moreira, F. (2025). A hybrid fog-edge computing architecture for real-time health monitoring in IoMT systems with optimized latency and threat resilience. *scientific reports, 15*(1), 25655. https://doi.org/10.1038/s41598-025-09696-3

Minh, Q. N., Nguyen, V. H., Quy, V. K., Ngoc, L. A., Chehri, A., & Jeon, G. (2022). Edge computing for IoT-enabled smart grid: The future of energy. *Energies, 15*(17), 6140. https://doi.org/10.3390/en15176140

Gowda, D., Sharma, A., Rao, B. K., Shankar, R., Sarma, P., Chaturvedi, A., & Hussain, N. (2022). Industrial quality healthcare services using Internet of Things and fog computing approach. *Measurement: Sensors, 24*, 100517. https://doi.org/10.1016/j.measen.2022.100517

Sharma, M., Tomar, A., & Hazra, A. (2024). Edge computing for industry 5.0: Fundamental, applications, and research challenges. *IEEE Internet of Things Journal, 11*(11), 19070-19093. https://doi.org/10.1109/JIOT.2024.3359297

Bablu, T. A., & Rashid, M. T. (2025). Edge computing and its impact on real-time data processing for IoT-driven applications. *Journal of advanced computing systems, 5*(1), 26-43. https://doi.org/10.69987

Kong, L., Tan, J., Huang, J., Chen, G., Wang, S., Jin, X., ... & Das, S. K. (2022). Edge-computing-driven internet of things: A survey. *ACM computing surveys, 55*(8), 1-41. https://doi.org/10.1145/3555308

Kong, X., Wu, Y., Wang, H., & Xia, F. (2022). Edge computing for internet of everything: A survey. *IEEE internet of things journal, 9*(23), 23472-23485.