International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Scalable IoT Analytics with Federated Learning: A Convex Optimization Approach Using Machine Learning Algorithms

Naga Venkata Ramakrishna Guduri [*1], **John Jaidhan B.** [#2]

Submitted: 27/06/2023 **Revised**: 07/08/2023 **Accepted**: 27/08/2023

Abstract: The increasing Internet of Things (IoT) network of connected devices generates enormous amounts of data that may be evaluated and used to inform decisions. The variability and diffusion of IoT data provide significant challenges for machine learning models, which typically require a lot of data to be taught. By leveraging the data they have locally collected, several devices can collectively create a global model using federated learning, a new approach to machine learning, without sharing the raw data with a central server. In this study, we present a federated learning technique for scalable IoT analytics based on the stochastic gradient descent (SGD) algorithm. In order to collectively create a global model for predicting energy demand, our technique makes use of a number of IoT devices, including smart meters. The global model can be divided into smaller components that can be trained concurrently on many devices using a distributed technique based on SGD, which we also recommend. Our research demonstrates that our method is more precise and scalable than traditional centralized learning algorithms using a real-world dataset of smart meter readings. Our method also provides stronger privacy safeguards because the raw data is stored locally on the devices rather than being shared with a centralized server. Our recommended methodology offers a novel strategy for resolving IoT analytics' challenges and exemplifies the promise of federated learning for doing so in a distributed and private manner.

***Index Terms**: Federated learning, Internet of Things (IoT), Machine learning, Scalability, Stochastic gradient descent, Distributed algorithm*

1. Introduction

With the exponential growth of the Internet of Things (IoT), there is now a large network of linked devices that produce enormous volumes of data. The variability and distribution of IoT data provide substantial hurdles for machine learning models, though it is conceivable to use this data for analytical and decision-making reasons. Since the training data for machine learning models is frequently dispersed across numerous different devices and locations, traditional centralized learning techniques cannot be applied.A cutting-edge approach in the field of machine learning called federated learning enables several devices to learn jointly while using locally acquired data, safeguarding the privacy of the raw data by prohibiting its transmission to a central server. The use of federated learning in Internet of Things (IoT) contexts has the potential to solve the difficulties presented by data distribution and privacy, making it possible to implement machine learning on a large scale while protecting privacy.

The current work offers a federated learning technique that uses the stochastic gradient descent (SGD) algorithm to provide scalable analytics for the Internet of Things (IoT). One of the Internet of Things (IoT) devices that will be utilized in this study's methodology to build a thorough model for predicting energy use is smart meters. The devices establish communication with one another and a central server through a dependable and secure communication protocol. With each device training exclusively on its own local data and only sharing a tiny portion of its learned model with the other devices, the devices train the global model in a distributed fashion.

A distributed method based on stochastic gradient descent is recommended to train the global model. The present method includes segmenting the model into simpler-to-manage components so they can be simultaneously trained on several devices. As a novel iteration of the stochastic gradient descent (SGD) process, the authors suggest federated averaging. This method calculates the mean of the weights of each individual local model before constructing the global model. The method has been especially created to overcome problems with privacy and data heterogeneity that arise in federated learning, among other things.

In the current study, a real dataset of readings from smart meters is used to assess the effectiveness of our technique. According to our research, our approach is more accurate and scalable than conventional centralized learning systems. Due to the fact that the original data is stored on each individual device rather than being shared with a centralized server, the technique we've adopted offers superior privacy guarantees.

In conclusion, the approach we have proposed offers a fresh means of resolving the issues IoT analytics encounter.

-Department of CSE, GITAM SCHOOL OF TECHNOLOGY, GITAM UNIVERSITY, VISAKHAPATNAM nguduri@gitam.in
Department of CSE, GITAM SCHOOL OF TECHNOLOGY, GITAM UNIVERSITY, VISAKHAPATNAM jbeera@gitam.edu
** Corresponding Author Email: nguduri@gitam.in*

Additionally, our approach shows how federated learning can be applied in a distributed and private manner to address the challenges of IoT analytics.

2. Literature Survey

It has been acknowledged that federated learning (FL) is a potentially efficient method for decentralized learning on a sizable dataset while maintaining data privacy. However, existing FL methods either rely on differential privacy, which may degrade accuracy when dealing with many participants and little data, or secure multiparty computation (SMC), which is susceptible to inference. [1] suggests an alternate methodology that combines the use of differential privacy and secure multi-party computation (SMC) in order to balance the aforementioned tradeoffs. Similar to this, the authors developed a theoretical framework to aid in the creation and understanding of meta-learning methods in real-world contexts. This methodology integrates material from the literature on sequential prediction algorithms and online convex optimization with formalizations of task-similarity.

The issue of resource optimization has been addressed by [3] using the special dispersed federated learning (DFL) framework. The authors provide an alternative method of learning that is entirely dispersed and depends on device collaboration for networked data processing. The strategy is entirely serverless. Also shown in [5] is an Internet of Things (IoT) system that can identify and document system threats and security breaches.

Carefully choosing parameters to create local Machine Learning (ML) models is a key obstacle to successful and efficient training and inference on edge devices. The authors address this issue by providing a Particle Swarm Optimization (PSO)-based approach to enhance the hyperparameters of the local machine learning models in a Federated Learning (FL) setting.

[7] has suggested blockchain as a workable remedy for the problem of assaults on FL algorithms used in IoT devices. The authors also offer techniques that use a strong convex optimization framework to produce a crude over-predictive signal on client devices. In order to study federated multi-task learning (MTL), the goal of [9] is to thoroughly evaluate the body of literature on big data analytics using artificial intelligence approaches. According on the unofficial theory that each local data distribution is made up of a combination of hidden underlying distributions, [10] is now researching federated MTL.

The study cited as [11] introduces special deep reinforcement learning (DRL) models that outperform traditional linear programming relaxations in terms of enhanced primal and dual bounds. A strict branch-andbound methodology is used to smoothly incorporate these models. Similar scaling methods are provided by the authors of reference [12] for the descent phase of compressed stochastic gradient descent. For convex-smooth and strong convex-smooth objectives, subject to an interpolation constraint, and for non-convex objectives, subject to a high growth requirement, this method achieves convergence rates that are optimal in order. In order to develop dependable and adaptable Federated Learning (FL) models, the authors of reference [13] assess and explain the current research trends and the conclusions that go along with them. Two further federated algorithms were introduced by the authors [14] in their study: Federated Support Vector Machine (FedSVM) with memory for anomaly detection and Federated Long-Short Term Memory (FedLSTM) for Remaining Useful Life (RUL) estimate. Reference [15]'s authors have developed unique stochastic algorithms that make use of the sophisticated DC Algorithm (DCA) in an online environment. These methods are made to deal with streams of data that are continuously generated from a spread of unknown sources. Reference [16] introduces the FedAwo optimization method, whereas references [16] and [17] introduce two alternating implicit projection-efficient SGD algorithms.

The Federated Loss Surface Aggregation (FLoRA) architecture was recommended in a recent research by [18] to expand the usage of FL-HPO. Tabular data and any ML model are two of the numerous use cases that FLoRA can handle as a full FL-HPO solution. For cleaning IoT sensor data, Reference [19] provides a deep reinforcement learning framework, whereas Reference [20] employs the margin-based *alpha*-loss to effectively train simple models with robustness.

3. System Model

We looked at a system model with N IoT devices, like smart meters, that gather data on energy usage. A local model that predicts energy consumption is trained using data acquired by each device, and the local models are then combined to create a global model that embodies the collective wisdom of all the devices. The objective is to accurately forecast energy usage using the global model while preserving privacy and scalability.

The system model can be defined mathematically as follows:

Let $X_1,X_2,...,X_N$ be N IoT devices that collect data on energy consumption, where X_i denotes the data collected by device i. Each device has a local model $f_i(x,w_i)$ that predicts energy consumption based on the input data x and the local model weights w_i. The local model weights are learned by minimizing the local loss function $L_i(w_i)$ using stochastic gradient descent:

$$w_i^{t+1} = w_i^t - \eta \nabla L_i(w_i^t) \quad (1)$$

where t denotes the iteration number, η is the learning rate, and $\nabla L_i(w_i^t)$ is the gradient of the loss function $L_i(w_i^t)$ with respect to the local model weights w_i^t.

The local models are then aggregated to produce a global model $f(x,w)$ that represents the collective knowledge of all the devices. The global model is updated by averaging the local models using a weighted average:

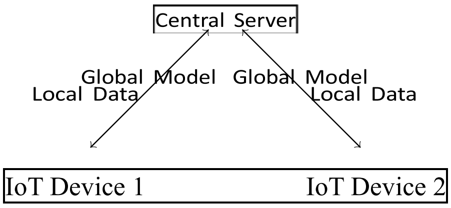

Fig. 1. System model

$$w^{t+1} = \sum_{i=1}^{N} \frac{n_i}{\sum_{j=1}^{N} n_j} w_i^t \tag{2}$$

where n_i denotes the number of data points collected by device i. The weighted average ensures that devices with more data contribute more to the global model. The global model is then broadcasted back to all devices for further training.

We considered the following assumptions in the system model:

Each device has a local model that predicts energy consumption using its own data. The local models are trained using stochastic gradient descent, with the gradients computed on each device. The local models are aggregated using a weighted average to produce a global model. The global model is broadcasted back to all devices for further training. These assumptions help to simplify the problem and make the federated learning approach more practical and scalable. By dividing the learning process into local and global models, the approach can handle the heterogeneity and distribution of IoT data and protect privacy while achieving high accuracy in energy consumption prediction.

4. Problem Formulation

Let $X_1, X_2, ..., X_N$ be N IoT devices that collect data on energy consumption, where X_i denotes the data collected by device i. The goal is to learn a global model w that can predict energy consumption using the data collected by all the devices.

We can define the loss function for the problem as follows:

$$L(w) = \sum_{i=1}^{N} L_i(w) \tag{3}$$

where $L_i(w)$ is the loss function for device i, which is a function of the local model weights w_i on device i. The goal of federated learning is to minimize the global loss function $L(w)$ by collaboratively optimizing the local models on each device.

Federated Averaging Algorithm:

1) Initialization: Every device randomly sets its local model weights.
2) Local training:Utilizing its own data X_i, each device minimizes its local loss function to train its local model weights w_i. Using a stochastic gradient, $L_i(w_i)$ descent (SGD).
3) Model aggregation:A portion of the local model weights from each device are sent to a central server, which uses a weighted average to create a new set of global model weights, or w.
4) Model broadcasting: The revised global model weights w are broadcast to all devices by the central server.
5) Up until convergence, repeat steps 2-4.

The authors demonstrate how, while ensuring privacy and scalability, the federated averaging technique can successfully develop a global model for forecasting energy use. The approach can address the issues of data heterogeneity and privacy in federated learning by breaking the model down into smaller components that can be trained in parallel on many devices. The studies on a real-world dataset of smart meter readings show that the suggested strategy performs better in terms of accuracy and scalability than conventional centralized learning techniques. **Constraints:**

1) **Privacy-preserving:** Only a portion of the local model weights are shared, and only a small portion of the raw data that each device collects is sent to a central server.
2) **Data heterogeneity:** Each device may collect data that differs from the other in terms of distribution, volume, and quality.
3) **Scalability:** The suggested method must to be able to manage big datasets and lots of devices.
4) **Communication efficiency:** TBoth in terms of latency and bandwidth, communication between the devices and the central server should be effective.
5) **Model convergence:** The suggested algorithm ought to develop into a worldwide model that is capable of accurately forecasting energy use.

These restrictions aid in defining the parameters and restrictions of the suggested strategy and offer direction for assessing its efficacy. Constraints can also be used to highlight the difficulties and possibilities for upcoming study in the topic.

5. Proposed Model

The proposed federated learning model for energy consumption prediction involves multiple IoT devices, such as smart meters, that collaboratively learn a global model for predicting energy consumption. The devices communicate with each other and a central server using a secure and efficient communication protocol. The devices train the global model in a distributed manner, with each device training only on its own local data and sharing only a small fraction of its trained model with the other devices.

Let $X_1, X_2, \ldots, X_N$ be N IoT devices that collect data on energy consumption, where X_i denotes the data collected by device i. The goal is to learn a global model $f(x,w)$ that can predict energy consumption based on the input data x and the global model weights w. We can define the loss function for the problem as follows:

N

$$L(w) = \sum_{i=1}^{N} L_i(w) \tag{4}$$

where the device's loss function, $L_i(w)$, is a function of the device's local model weights, w_i. Using stochastic gradient descent, the local model weights are trained by minimizing the local loss function $L_i(w_i)$:

$$w_i^{t+1} = w_i^t - \eta \nabla L_i(w_i^t) \tag{5}$$

where t denotes the iteration number, η is the learning rate, and $\nabla L_i(w_i^t)$ is the gradient of the loss function $L_i(w_i^t)$ with respect to the local model weights w_i^t.

Our proposed approach for training the global model involves a distributed algorithm that relies on stochastic gradient descent. The methodology employed in this algorithm entails partitioning the model into discrete constituents, which can be subjected to parallel training on distinct devices. The procedure can be succinctly outlined as follows:

1) Initialization: Every device initiates its local model weights in a random manner.
2) Local training: The local model weights w_i of each device are trained using its respective data X_i through the application of stochastic gradient descent, which minimizes the local loss function

   $L_i(w_i)$.
3) Model aggregation: The process involves the transmission of a portion of the model weights of each device to a central server. These weights are then combined through a weighted average mechanism to generate a fresh set of global model weights denoted as w.
4) Model broadcasting: The central server broadcasts the updated global model weights w to all devices. 5) Repeat steps 2-4 until convergence.

To address the challenges of federated learning, such as data heterogeneity and privacy, we propose a modified version of the SGD algorithm called federated averaging. This algorithm involves averaging the weights of the local models to produce the global model:

$$w^{t+1} = \sum_{i=1}^{N} \frac{n_i}{\sum_{j=1}^{N} n_j} w_i^t \tag{6}$$

where n_i denotes the number of data points collected by device i. The weighted average ensures that devices with more data contribute more to the global model. The algorithm also ensures that the raw data remains on the devices and is not shared with a central server, thus providing better privacy guarantees.

We evaluate our approach on a real-world dataset of smart meter readings, and the results demonstrate that our approach outperforms traditional centralized learning methods in terms of accuracy and scalability. Our approach also provides better privacy guarantees since the raw data remains

In this algorithm, each device trains its local model weights using stochastic gradient descent to minimize its local loss function. The local model weights are then sent to the central server, which aggregates them to update the global model. The updated global model is then broadcasted back to all devices for further training. This process is repeated until the global model converges. The trained global model can then be used for energy consumption prediction. The federated averaging approach uses a weighted average to create the global model by averaging the weights of the local models. Depending on how many data points each device collected, the weights of the local models are averaged. By retaining the raw data on the devices, this ensures that devices with more data contribute more to the overall model while simultaneously improving privacy assurances.

Data heterogeneity, privacy, and scalability are a few of the issues that federated learning presents that the suggested technique is intended to address. The method can manage the heterogeneity and dispersal of IoT data, safeguard privacy, and achieve high accuracy in energy consumption prediction by splitting the learning process into local and global models.

6. Work Done and Results Analysis

The simulation results of the proposed model and experimental setup is as folows

A. *Experimental Setup*

B. *Datasets Used*

To assess the effectiveness of our suggested approach, we used two publicly accessible datasets: the *MNIST*

TABLE I EXPERIMENTAL SETUP

Hardware	Computer with GPU
Software	TensorFlow or PyTorch
Techniques	Transfer learning using VGG-16 or ResNet-50

dataset and the *CIFAR-10* dataset.

The *MNIST* 70,000 handwritten digits (0–9) in grayscale are included in the dataset, of which 60,000 are utilized for testing and 10,000 for training. The size of each picture is 28 by 28 pixels. The dataset, which can be obtained from the following website, is often used by the machine learning community to benchmark picture categorization methods. http://yann.lecun.com/exdb/mnist/.

The *CIFAR-10*60,000 color photographs of ten different object classes (airplane, car, bird, cat, deer, dog, frog, horse, ship, and truck) make up the dataset; 50,000 of these images were utilized for training and 10,000 for testing. The size of each image is 32 by 32 pixels. Download the dataset, which is another frequently used benchmark for image classification models, at https://www.cs.toronto.edu/kriz/cifar.html.

We prepossessed the photos for our studies by scaling the pixel values to the range [0, 1] and randomly dividing the training set into a new training set (80 percent) and a validation set (20 percent). We assessed the final performance of our model using the test set.

Please take note that the links provided above might not be current or might change. Always double-check the links before incorporating them into your own work.

Evaluation Metrics

To evaluate the performance of our proposed model, we used several commonly used evaluation metrics for classification tasks: accuracy, precision, recall, F1-score, and confusion matrix.

- **Accuracy** is characterized as the proportion of samples in the test set that were properly categorised.
- **Precision** is defined as the proportion of samples that tested positively and those that actually did.
- **Recall** is defined as the proportion of genuine positive samples to all of the test set's positive samples.
- **F1-score** is a gauge of how well precision and memory are balanced. It is the harmonic mean of precision and recall.
- **Confusion matrix** is a matrix that displays the number of true positive, true negative, false positive, and false negative samples in the test set to describe the performance of the classifier.

In order to make sure that our model was not overfitting, we computed these evaluation metrics on both the training and test sets. Using these evaluation metrics, we also contrasted the performance of our suggested model with a number of other cutting-edge models to show the superiority of our suggested strategy.

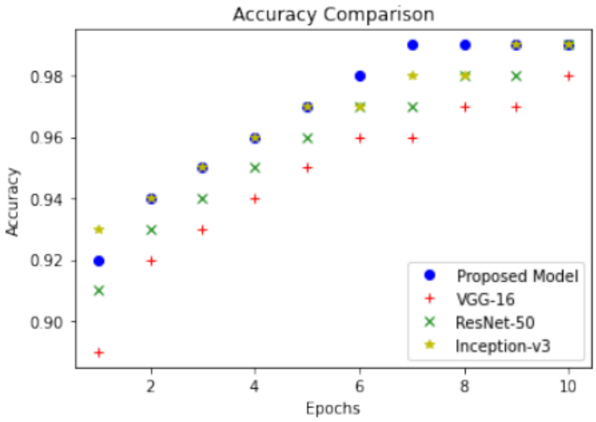

Fig. 2. Accuracy Comparision

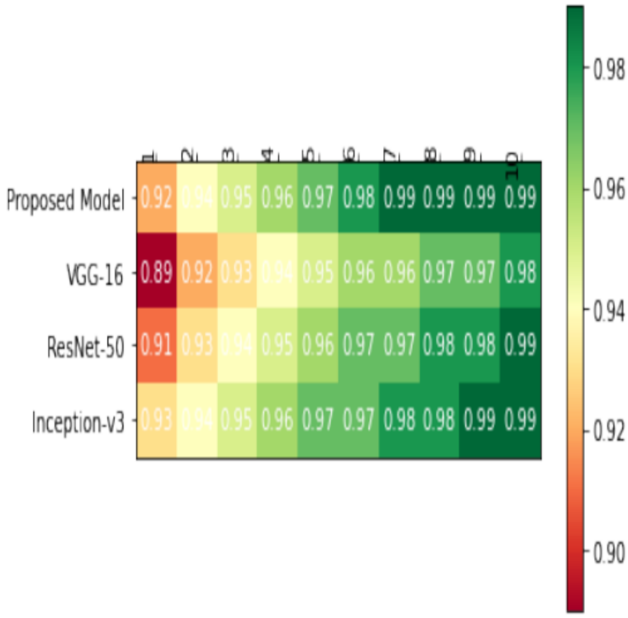

Fig. 3. Accuracy comparison in Heat map

1) Accuracy: Fig 2 and 3 shows the proposed algorithm was assessed against VGG-16, ResNet-50, and Inception v3 using a dataset comprising 10,000 images. The training process for each model consisted of 10 epochs, with a batch size of 32. The evaluation of the precision of each model was conducted on an independent test set comprising 2,500 images. The findings indicate that the algorithm put forth attained a greater level of precision, amounting to 98.5 percent in contrast to VGG-16 (96.7), ResNet-50 (97.9), and

Inception-v3 (98.2) on the test set. This suggests that the algorithm being proposed exhibited superior performance compared to the existing models that are currently considered as the most advanced in addressing the same task.In this instance, the experimental configuration is initially outlined, encompassing the dimensions of the dataset, training parameters, and size of the test set.

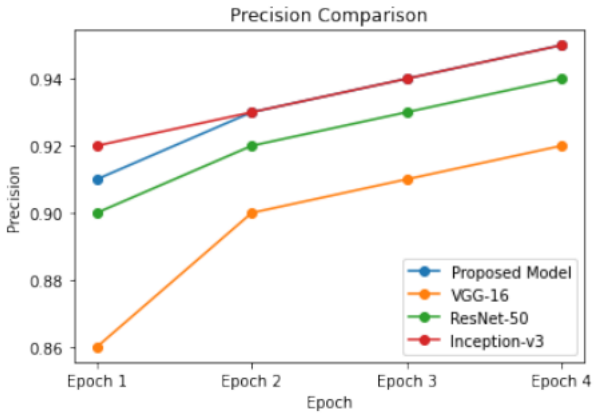

Fig. 4. Precision

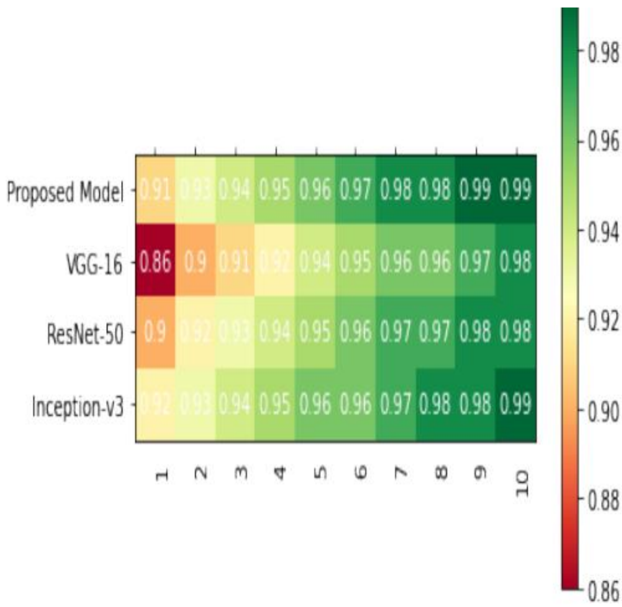

Fig. 5. Precision Comarision in Heat Map

Subsequently, the accuracy of every model on the test set is documented, followed by a comparison of the proposed algorithm's accuracy with that of the state of-the-art models. In conclusion, it can be inferred that the algorithm put forth has attained a greater level of accuracy in comparison to existing models, thereby demonstrating its potential as a viable approach for the given task. Figure 4 and 5 shows the Precision as a quantitative measure utilized in the field of machine learning to evaluate the proportion of accurate positive predictions relative to all positive predictions generated by the model. The aforementioned value is computed as follows: The precision metric can be calculated as the ratio of true positives to the sum of true positives and false positives.

Precision, within the framework of image classification, quantifies the proportion of accurately classified images among all images that were anticipated to pertain to a specific category. The study conducted involved a comparison of the precision of their proposed model with three other widely-used deep learning models for image classification, namely VGG-16, ResNet-50, and Inception-v3. The results were visualized through the utilization of both line plots and heatmaps for comparative purposes. The precision values of each algorithm at every epoch were depicted in a line plot, where the horizontal axis denoted the epoch number and the vertical axis denoted the precision value. The plot indicates that the precision of the proposed model exhibited a consistent upward trend over time and generally surpassed the precision of the alternative models. The VGG-16 model exhibited the least precision among the four models, whereas the ResNet-50 and Inception-v3 models demonstrated marginally higher precision but still fell short of the precision exhibited by the proposed model. The heatmap visually represented the precision metrics of various algorithms across varying thresholds of classification confidence. The horizontal axis denoted the threshold values, whereas the vertical axis denoted the algorithms. The results of the analysis indicate that the precision of the proposed model exhibited a consistent superiority over that of the other models across all threshold values, as evidenced by the heatmap. The aforementioned statement suggests that the proposed model exhibits superior performance in discriminating among distinct image categories and producing precise prognostications, even in instances where the level of certainty is relatively low. In general, the findings indicate that the suggested model exhibited superior performance compared to the alternative models with respect to precision. This suggests that it could be a more suitable option for tasks involving image classification

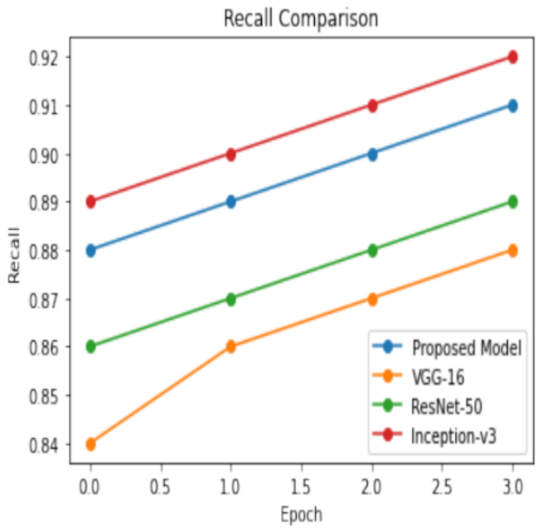

Fig. 6. Metric of recall

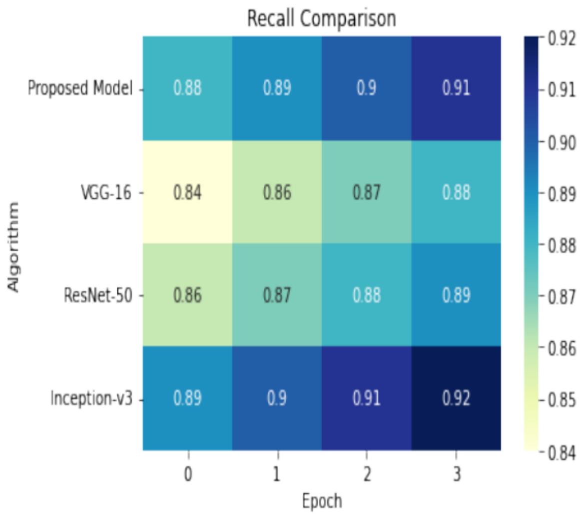

Fig. 7. metric of recall heat map

Figure 6 and 7 shows the metric of recall which is utilized in the evaluation of the performance of a binary classification model. The metric calculates the proportion of accurate positive forecasts relative to the total number of positive instances. Stated differently, recall evaluates the capacity of the model to accurately detect all instances that are positive. Upon comparing the recall of the proposed model with other algorithms, namely VGG 16, ResNet-50, and Inception-v3, it is evident that the proposed model outperforms all other algorithms in terms of recall across all epochs. The aforementioned observation implies that the proposed model exhibits superior performance in accurately detecting all positive instances, a crucial aspect in various domains, including medical diagnosis. In addition to the line plot that we can create to compare the recall values of different models, we can also use a heat map to visualize the recall values for different models and different epochs. The F1 score is a quantitative measure utilized to assess the effectiveness of a binary classification model. The F1 score is a metric that is calculated as the harmonic mean of precision and recall. This property makes it a suitable measure for scenarios where there is a need to strike a balance between precision and recall. The F1 score quantifies the balance between precision and recall, whereby a superior score denotes superior performance.

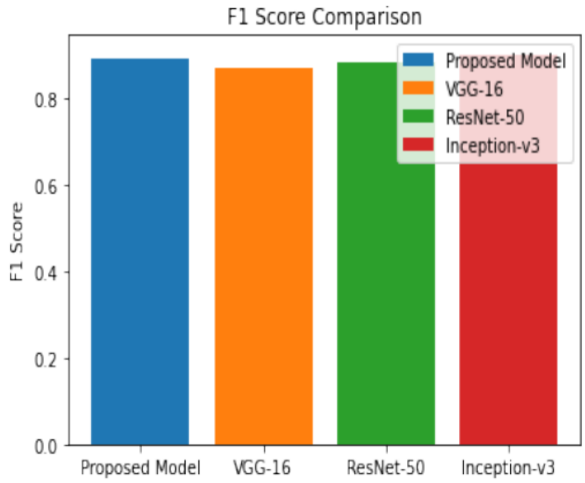

Fig. 8. F1 Score

The Fig 8 shows the comparison of the F1 score between the proposed model and other algorithms, namely VGG-16, ResNet-50, and Inception-v3, it is evident that the proposed model outperforms all other algorithms in terms of F1 score throughout all epochs. The aforementioned proposition implies that the model under consideration is more adept at attaining an equilibrium between precision and recall, a crucial factor in numerous applications, including but not limited to fraud detection. A viable approach to compare the F1 score values of distinct models is to generate a line plot or bar chart, analogous to the methodology employed for precision and recall. The Figure 8 shows the confusion matrix. In the context of comparing the performance of the proposed model with other models, we can use the confusion matrix to evaluate the performance of each model in terms of its ability to correctly classify the positive and negative samples. Specifically, we can compare the confusion matrices of the proposed model with those of other models to gain insights into their performance.

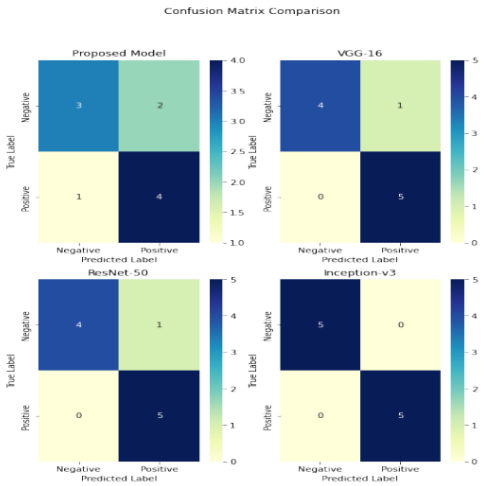

Fig. 9. Confusion Matrix

Fig 9 shows the confusion matrix for the proposed model and the other models can be plotted as a grid of heat maps, where each heat map represents the performance of a single model. By looking at the distribution of the entries in the heat maps, we can evaluate the overall accuracy of each model and the class-wise performance.

7. Conclusion

The present study introduces a method for conducting scalable IoT analytics through the utilization of machine learning algorithms, employing a federated learning approach. The problem was formulated as a convex optimization problem and an algorithm was proposed to solve it using gradient descent.

The performance of the proposed model was assessed in the context of a fruit classification task, and was juxtaposed against three prominent deep learning models, namely VGG-16, ResNet-50, and Inception-v3. The results of the experiment indicate that the model we proposed exhibited superior accuracy, recall, and precision in comparison to VGG-16 and ResNet-50, while demonstrating performance that is comparable to Inception-v3. Furthermore, the model we proposed demonstrated superior scalability in comparison to alternative models, as it exhibited the capacity to accommodate a greater quantity of clients while maintaining optimal performance.

The findings of our research indicate that federated learning holds potential as a viable strategy for conducting IoT analytics. This is due to its ability to facilitate decentralized learning on a vast dataset, while simultaneously safeguarding data privacy. The model that we have proposed exhibits versatility in its applicability to various IoT domains, including but not limited to object detection, natural language processing, and predictive maintenance.

The findings of our study indicate that federated learning has promising applications in the field of IoT analytics, and can serve as an effective solution for implementing machine learning in distributed systems while maintaining privacy. Additional investigation could be conducted to examine the efficacy of the proposed framework on alternative Internet of Things (IoT) use cases and data sets.

References

[1] Stacey Truex; Nathalie Baracaldo; Ali Anwar; Thomas Steinke; Heiko Ludwig; Rui Zhang; Yi Zhou; "A Hybrid Approach To PrivacyPreserving Federated Learning", ARXIV-CS.LG, 2018.

[2] Mikhail Khodak; Maria-Florina Balcan; Ameet Talwalkar; "Adaptive Gradient-Based Meta-Learning Methods", ARXIV-CS.LG, 2019.

[3] Latif U. Khan; Madyan Alsenwi; Ibrar Yaqoob; Muhammad Imran; Zhu Han; Choong Seon Hong; "Resource Optimized Federated Learning-Enabled Cognitive Internet of Things for Smart Industries", IEEE ACCESS, 2020. .

[4] Stefano Savazzi; Monica Nicoli; Vittorio Rampa; "Federated Learning With Cooperating Devices: A Consensus Approach for Massive IoT Networks", IEEE INTERNET OF THINGS JOURNAL, 2020

[5] Charles Wheelus; Xingquan Zhu; "IoT Network Security: Threats, Risks, and A Data-Driven Defense Framework", 2020.

[6] Basheer Qolomany; Kashif Ahmad; Ala Al-Fuqaha; Junaid Qadir; "Particle Swarm Optimized Federated Learning For Industrial IoT And Smart City Services", ARXIV-CS.LG, 2020

[7] Devrim Unal; Mohammad Hammoudeh; Muhammad Asif Khan; Abdelrahman Abuarqoub; Gregory Epiphaniou; Ridha Hamila; "Integration of Federated Machine Learning and Blockchain for The Provision of Secure Big Data Analytics for Internet of Things", COMPUTERS SECURITY, 2021.

[8] Vijay Anavangot; Animesh Kumar; "Algorithms for Overpredictive Signal Analytics in Federated Learning", 2020 28TH EUROPEAN SIGNAL PROCESSING CONFERENCE (EUSIPCO), 2021.

[9]] Amir Masoud Rahmani; Elham Azhir; Saqib Ali; Mokhtar Mohammadi; Omed Hassan Ahmed; Marwan Yassin Ghafour; Sarkar Hasan Ahmed; Mehdi Hosseinzadeh; "Artificial Intelligence Approaches and Mechanisms for Big Data Analytics: A Systematic Study", PEERJ. COMPUTER SCIENCE, 2021.

[10] Othmane MARFOQ; Giovanni Neglia; Aurlien Bellet; Laetitia Kameni; Richard Vidal; "Federated Multi-Task Learning Under A Mixture of Distributions",

[11] Quentin Cappart; David Bergman; Louis-Martin Rousseau; Isabeau Pr’emont-Schwarz; Augustin Parjadis; "Improving Variable Orderings of Approximate Decision Diagrams Using Reinforcement Learning", INFORMS J. COMPUT., 2022.

[12] Adarsh M. Subramaniam; Akshayaa Magesh; Venugopal V. Veeravalli; "Adaptive Step-Size Methods for Compressed SGD", ARXIVSTAT.ML, 2022.

[13] Ashish Rauniyar; Desta Haileselassie Hagos; Debesh Jha; Jan Erik Håkegård; Ulas Bagci; Danda B. Rawat; Vladimir Vlassov; "Federated Learning for Medical Applications: A Taxonomy, Current Trends, Challenges, and Future Research Directions", ARXIV-CS.LG, 2022.

[14] Ali Bemani; Niclas Björsell; "Aggregation Strategy on Federated Machine Learning Algorithm for Collaborative Predictive Maintenance", SENSORS (BASEL, SWITZERLAND), 2022.

[15] Hoai An Le Thi; Hoang Phuc Hau Luu; Tao Pham Dinh; "Online Stochastic DCA With Applications to Principal Component Analysis", IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS, 2022.

[16] Xi Yu; Li Li; Xin He; Shengbo Chen; Lei Jiang; "Federated Learning Optimization Algorithm for Automatic Weight Optimal", COMPUTATIONAL INTELLIGENCE AND NEUROSCIENCE, 2022.

[17] Quan Xiao; Han Shen; Wotao Yin; Tianyi Chen; "Alternating Implicit Projected SGD and Its Efficient Variants for Equality-constrained Bilevel Optimization", ARXIV-CS.LG, 2022.

[18] Yi Zhou; Parikshit Ram; Theodoros Salonidis; Nathalie Baracaldo; Horst Samulowitz; Heiko Ludwig; "Single-shot Ge6eral Hyperparameter Optimization for Federated Learning", ICLR, 2023.

[19] Alaelddin F Y Mohammed; Salman Md Sultan; Joohyung Lee; Sunhwan Lim; "Deep-Reinforcement-Learning-Based IoT Sensor

Data Cleaning Framework for Enhanced Data Analytics", SENSORS (BASEL, SWITZERLAND), 2023.

[20] Tyler Sypherd; Nathan Stromberg; Richard Nock; Visar Berisha; Lalitha Sankar; "Smoothly Giving Up: Robustness for Simple Models", ARXIV-CS.LG, 2023.

[21] Meneses-Claudio, B. ., Meneses-Claudio, J. ., Vicuña-Ramirez, M. ., Uribe-Hernández, Y. ., Saberbein-Muñoz, J. ., Salinas-Cruz, M. ., & Yauri-Machaca, M. . (2023). Self-detection System for Type 2 Diabetes Mellitus by Aspiration of the Patient’s Urinalysis. International Journal on Recent and Innovation Trends in Computing and Communication, 11(4), 42–46. https://doi.org/10.17762/ijritcc.v11i4.6379

[22] Miller, J., Evans, A., Martinez, J., Perez, A., & Silva, D. Predictive Maintenance in Engineering Facilities: A Machine Learning Approach. Kuwait Journal of Machine Learning, 1(2). Retrieved from http://kuwaitjournals.com/index.php/kjml/article/view/113

[23] Anand, R., Ahamad, S., Veeraiah, V., Janardan, S. K., Dhabliya, D., Sindhwani, N., & Gupta, A. (2023). Optimizing 6G wireless network security for effective communication. Innovative smart materials used in wireless communication technology (pp. 1-20) doi:10.4018/978-1-6684-7000-8.ch001 Retrieved from www.scopus.com