Federated Machine Learning For Augmenting The Safekeeping of Critical Energy Infrastructures

Muzaffar Hussain
C.Abdul Hakeem College of Engineering & Technology
Sumaiya Thaseen
Vellore Institute of Technology
Anbarasu B
Vellore Institute of Technology
Muhammad Rukunuddin Ghalib
Vellore Institute of Technology
Achyut Shankar
Amity University
Pavika Sharma (pavikasharma.ece@gmail.com)
Amity University https://orcid.org/0000-0002-9159-9041
Bharat Bhushan
Sharda University

Research Article

Keywords: Critical Infrastructure, Federated Learning, Machine Learning, Model Training, Random Forest and Wind Turbine.

Posted Date: November 29th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-746263/v1

Abstract

In the recent years, there have been an increase in attacks targeting Supervisory Control and Data Acquisition (SCADA) infrastructures as there are many sensitive data released from peripheral devices. Wind-turbine systems are considered to be the most complex Cyber-Physical infrastructures.A privacy preserving *Federated Machine Learning* solution is proposed in order detect any possible anomalies in such infrastructures. Instead of centralizing the wind-turbine data into a common server, *Federated Machine Learning* allows the data to remain on-premise in the infrastructure. This enables the responsible authorities to consider the advantages of Machine Learning, and simultaneously protect their privacy. Different federated machine learning models namely Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Radial Basis Function (RBF), Multi-Layer Perceptron (MLP) and Random Forest (RF) are deployed to analyze the anomalies in the wind turbines. It is inferred from the experimental results that Random Forest is superior in identifying the anomalies with regard to the performance metrics such asMean Absolute Error (MAE), Root Mean Square Error (RMSE) and Mean Square Error (MSE) are minimal for Random Forest in comparison to other models and Coefficient of Determination (R^2) is higher which is expected for a model that accurately predicts the anomalies. In addition, the time taken by the regression is 42.95 seconds which is minimal in comparison to other classifiers. Thus, a federated Random Forest Learner accurately analyzes the anomalies in critical wind turbine infrastructures.

I. Introduction

The evolution of internet paved way for the tremendous growth in the world on almost every field. Industry 4.0 and process automation is the need of the hour. Industrial automation has revolutionized the world and the technological development of IoT and IIoT devices has been a major factor for enhancing the lifestyle of humans and it is more sophisticated than before. However, standalone systems are directly connected to the internet along with the automated tools. Thus, the isolated environment becomesaccessible and hence the primary concern is security. There are various security concerns which should be considered to make these technologies reliable. Anomaly detection exists in almost all areas and application domains such as military applications, health care, banking, intrusion detection and safety of critical systems [12]

With the rapid expansion of global wind power capacity,anomaly defection by continuous monitoring of Wind Turbines (WT) is of increasing importance [14]. A system for fault recognition of WTs was developed utilizing normal behavior approaches [13].Operational & Maintenance costs of WT are higher so monitoring them is a big challenge. Health condition monitoring of Wind turbines is built based on the specific metrics[2]. Environmental friendly conditions make WT more preferable in almost every country [18]. In comparison with legacy SCADA(i.e. Supervisory Command and Data Acquisition-SCADA) systems, recently developed infrastructures deploy scalable Internet-based technologies for real-time data monitoring and also less expensive [19]. Communication between Industrial Control Systems(ICS) is based on information technology tasks and remote connectivity. There is a possibility of attacks in physical plants. In general, the protocol runs are secured but with the connectivity of ICS it is definitely challenging and susceptible to attacks [20].

Machine leaning and Deep Learning models are superior in detecting the anomalies in industry domain. Spatio-temporal data from the applications are analyzed with the aid of statistical models like regression, clustering etc., to eradicate the degraded performance of the system. SCADA systems are useful in monitoring and controlling the infrastructures system which plays a vital role in the current scenarios. Therefore, in the proposed approach, a robust anomaly detection framework is developed. **Federated learning** (also known as **collaborative learning**) is a machine learning approach that trains decentralized edge devices or servers holding local data instances, without swapping them. This method stands in dissimilarity to outdated centralized machine learning techniques wherever all the local datasets are uploaded to single server, as well as to more traditional distributed methods which often accept that local data samples are identically dispersed.*Federated learning* is able to train a model using data stored at multiple wind-turbine stations without the data leaving the station's premises, as it is illustrated in Figure 1.

The rest of the paper is summarized as follows: Section 2 discusses the review of various fault detection models developed in IIoT domain. The proposed work is described in section 3. Section 4 analyzes the experimental results. The conclusion is given in section 5.

Ii. Literature Review

Du et al [1] deployed a Self-Organizing Map (SOM) for reducing high dimensional data. In this approach, Euclidean distance vector based measure is used with filter for anomaly detection. In addition, data driven method is used for anomaly detection. Zhao et al. [2] developed a Data Driven Anomalous Detection Method with Time Series Analysis. Anomaly Operation Index (AOI) is used to measure the performance in Wind Turbine. The remaining Useful Life (RUL) is also identified in this approach. Sun et al. [3] developed a system using Back Propagation Neural Networks (BPNNs). Genetic Algorithm with Partial Least Squares Regression is also deployed. The model performs with low Prediction errors on Normal Condition and with high Prediction error on Anomaly detection. Rosa et al. [4] built Intrusion and Anomaly Detection System (IADS) for specific detection containing an anomaly detection module. Security challenges for next generation IADS is also built. A unified framework for several heterogeneous component is built into the system.

Rashid et al. [6] utilized a Bagging Regressor ML model to avoid technical issues in Gear Box of Wind Turbines (WT). The Bagging Regressor determines faults in WT GB's before 59 days of failure. In this approach, the efficiency is improved and also computation time is reduced. An accuracy of 99.8 and a MSE of 0.33 is obtained on training the model. Shlomo et al. [7] built a supervised method which interprets the frequent temporal patterns from the SCADA payload of communication protocols and those features are used in the Classification algorithm. A unsupervised algorithm that learns the automaton is incorporated. The unknown states are defined as malicious. A comparative analysis is performed and it is observed that unsupervised algorithm performs better than the supervised. MODBUS-SCADA dataset is used for evaluation in this method. Xiang et al [8] developed a fault detection approach by cascading of CNN+LSTM. Early warning of fault state can be done and early failure is predicted. Sheng et al. [9] developed a Cyber physical model for anomaly detection in SCADA systems. A model is built for characterizing the industrial process through extractions and correlating patterns of ICS nodes. Mokhtari et al. [10] developed a Measurement IDS(MIDS) which is capable of identifying abnormal activities even the intruder tries to conceal systems control layer. The hardware in the Loop

test bed is used to exploit the attack dataset. Random forest performs better than any other classifier algorithm in detecting anomalies with measured dataset testbed. Roelofs et al [11] deployed an Auto encoder to find the anomalous behavior in WT.

Philips et al.[15] developed models using SVM, Decision trees, KNN and k-means methods and the results show that 99.99 accuracy is obtained for Gas pipelines on public datasets. Yang et al. [5] proposed a (Distributed Network Protocol)DNP 3 using Convolutional Neural Network wherein the accuracy obtained is 99.33. Anton et al. [16] applied SVM and Random Forest techniques and analyzed attacks on Layers 2/3/7 of the network packet and application based features for public datasets. Sololov et al.[17] built Recurrent Neural Network (RNN)method and tested with public dataset.

Thus, it is evident from the literature that many machine learning techniques have been deployed but a federated machine learning model has not been developed. Hence, in the proposed approach, the objective is to develop a federated machine learning model for critical energy infrastructures as privacy will be preserved and anomalies can also be detected in a secure manner.

Iii. Proposed Approach

A. Privacy-preserving Federated Learning Approach

Wind turbine multisite federated machine learning analysis is formulated without data distribution. The privacy protection is implemented by a randomized mechanism. The wind turbine privacy-preserving federated learning training is shown step by step.

1. Problem definition

The data detained by the data proprietor location o is represented by the matrix D_O. A deep learning technique at State N locations $\{WT_1, \ldots WT_N\}$, is trained by integrating their own data $\{WD_1, \ldots .W\ D_N\}$. For wind turbine problems, classically, the data scope at every place is incomplete to train a correct deep learning model. All wind facts are collected and deployed $WD = WD_1 \cup \cdots \cup WD_N$ to train a wind model called as WM model.The detail space is denoted as M and the label space as N. The feature space 'M', label 'N' and model IDs 'I' constitute the training dataset. In the proposed approach of wind turbine multisite federated machine learning cataloging scenario: Do represents the federated machine learning data, WTo specifies the institution possessing private federated machine learning data; M is the extracted federated machine learning article and label N signify the analysis or phenotype to be predicted. In this condition, facts groups part the same detail space but are dissimilar in examples. For instance, dissimilar locations have dissimilar data. However, the structures are all wind turbine federated machine learning signs mined from the similar pre-processing. The data distribution is shown in Eqn (1) below:

$$M_i = M_j, N_i = YNj, O\,i \neq O\,j, \forall\ WD\,i, WD\,j, i \neq j, (1) \qquad \text{-----(1)}$$

2. Distributed Training

The two key steps in implementing a wind turbine federated learning approach are: 1) Local update and 2) Global server communication. The detailed training approach is given in pseudocode 1. Figure 2 shows the architecture of the proposed model. The WM model is trained collaboratively to develop a federated learning hierarchy. It is assumed there is a central server for computation. Every individual wind turbine trains the deep learning model and updates the model weights information to a central server. Once the central server has obtained all weights, it analyses and updates the new weights to all the WTs.

B. Gaussian Mechanism

Gaussian mechanism is the structure block of private experimental risk minimization algorithms founded on stochastic gradient descent [40]. Analyzing the privacy of such complex mechanisms turns out to be a delicate and error-prone[41]. In particular, obtaining tight privacy analyses leads to optimal utility which is one of the major challenges in the design of advanced DP mechanisms. An alternative to a-priori analyses is to equip complex mechanisms with algorithmic noise calibration and accounting methods. These methods utilize numerical computations to, e.g. calibrate perturbations and compute cumulative privacy losses at run time, without relying on hand-crafted worst-case bounds. For example, recent works have proposed methods to account for the privacy loss under compositions occurring in complex mechanisms [42][43]

C. Laplace Mechanism

As the provision of the Laplace distribution is infinite, it is mutual for the output of the Laplace mechanism to decrease external the range of Q. Currently, there are two current solutions to overawed this. A truncation is achieved which, comprises a deterministic drawing to the upper/lower limits of the output domain, when the worth falls outside. Alternative method is to bound the support of the reply mechanism, and then sample right from the output domain (e.g. by inverse transform sampling). This can also be attained through denial sampling, by continually redrawing from the unbounded distribution till an output reduction within the domain. This process is called as bounding, as the pure outputs of the mechanism are bounded by plan.

D. Coefficient of determination

The coefficient of determination (R^2) metric measures how good the algorithm to the regression of the calculation by calculating the association between the input constraints and targeted variables. This metric is also called fitting degree of regression as well. Mathematically this is expressed as:

$$R^2 = 1 - \frac{\sum_{m-1}^{K} (Gbt_p(m) - Gbt_a(m))^2}{\sum_{m-1}^{K} (Gbt_a(m) - Gbt'_a(m))^2} \qquad \text{--- (2)}$$

where Gbtp represents gearbox forecast temperature, Gbta specifies gearbox real temperature, K denotes the number of data points which is K=1,2,3.....m and the mean actual gearbox temperature is Gbt'a. The maximum value of R^2 is 1. If the value is nearer to 1 the model is good and appropriate to the regression line to the targeted variable. If the value is less (less than 0.5) the model is not decent for this data.

E. Mean Square Error (MSE)

The MSE indicates the average square difference between the real value and forecast value. The MSE measures the excellence of the regression ML algorithm. The value earlier to zero is better. If the forecast GB temperature is Gbtp(m) and real GB temperature is Gbta(m) and total data samples are K mathematically MSE is expressed as:

$$MSE = \frac{1}{K}\sum_{m-1}^{K}(Gbt_a(m) - Gbt_p(m))^2 \quad \text{----- (3)}$$

F. Mean Absolute Error

The mean square error which denoted as MAE. The MAE calculate the average between the complete real value and the complete forecast value. Mathematically this error is represented as :

$$MAE = \frac{1}{K}\sum_{m=1}^{K}\left|Gbt_p(m) - Gbt_a(m)\right| \text{ --- (4)}$$

where Gbtp(m) denotes the forecast gearbox temperature, Gbta(m) is actual gearbox temperature and K is total data samples.

G. Root Mean Square Error (RMSE)

RMSE score isconsidered as the best for determining the excellence of forecasts. Root mean square error can be expressed as

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}\left\|y(i) - \hat{y}(i)\right\|^2}{N}} \quad \text{---- (5)}$$

Iv. Experiments And Results

A. Data Analysis

SCADA systems model various attributes namely the wind turbine operation method, wind speed, turbine number, angle of the blade, Torque value, Wind Speed Max, and power output. An example of raw SCADA data with a sampling period of 10min is shown in Table 1 and data taken for model and processing is from [8].

Input: 1. WD = {W D 1 , . . . ,W D N } , federated data from N wind turbine ;

2. f_w = { f_{w1} , . . . , f_{wN} } , , where w_i denotes the local model weights;

3. L = { L1 , . . . , L N } , wind turbine labels;

4. €, noise generator that is deployed for preserving privacy

5. O, optimization repetitions;

6. global model updating pace, represents that the global and the private models converse in each optimization iteration;

7. { opt 1 (·) , ... , opt N (·) } , optimizer updating model weights w.r.t. objective function F .

Steps:

1. (wt1,wt2,....wtn)initialize all local model
2. For each round t=1,2,3.....n
3. Local model generation
4. Wd1←wd1+noice(€1)
5. Wd2←wd2+noice(€2)
6. Wdn←wd2+noice(€n)
7. Make aggregate in global model
8. Wdn=∑wd1+wd2+....wdn
9. Write back aggregate to local model
10. for n = 1 to N do
11. wd (n) n ← ˜wd (n) deploy weights to local model
12. end for
13. end for
14. end for Return: global model f wd (N)

Pseudo code 1: Wind turbine Privacy-preserving federated learning analysis

Table 1: Raw SCADA Data

TIME	TURBINE_NUM	WIND_SPEED	KW	WIND_SPEED_SD	WIND_SPEED_MAX	TORQUE_ACTUAL_VALUE	BLADE_1_ACTUAL_ANGL
01-11-2015 00:00	22	0.148473034	0.009655	0.064692982	0.110283159	0.025785188	0.458179171
01-11-2015 00:10	22	0.125081222	0.004962	0.066885965	0.084016393	0.020162854	0.466428095
01-11-2015 00:20	22	0.121182586	0.004913	0.060307018	0.086624441	0.020841411	0.473221326
01-11-2015 00:30	22	0.137751787	0.004454	0.067982456	0.104321908	0.020841411	0.628919755

B. Data Processing and Feature Selection

The first impartial is to contract with the SCADA data of WT and to remove the needless information which is not essential for health valuation of GB(Gear Box) of the WT. The second impartial is to the selection of ML model and train model by the healthy WT GB temperature data. The maximum accurate model is designated based on the results. The third one is to early forecast the GB temperature of the in the unhealthy data of the second WT. The SCADA dataset contains a lot of missing value which are not recorded. After complete analysis, the missing values are removed. The feature selection is implemented using the association statistics. The highly related input variables to the targeted output (Gearbox output temperature) are selected.The value of correlation is represented in the range of 0 to 1. The input variables with a value of less than 0.5 are removed and greater than 0.5 is selected

C. Machine Learning Model Training

After the SCADA data preprocessing and feature selection, the ML model is trained using five different regressors which are additional described in the following sections

(i) Random Forest Classifier

The first algorithm is a Random Forest classifier which uses the average value of sub samples. The RMSE, MAE, MSE and R^2 reported were 0.096, 0.023, 0.026, and 93.6% respectively. The time taken by this classifier was 42.95 seconds. Compare to the others model the prediction accuracy reported for this model is high.

(ii) .K-Nearest Neighbors (KNN) Classifier

KNN classifier predicts the targeted output by exclamation with nearest neighbors in the training data. The RMSE, MAE, MSE and R^2 reported were 0.56, 0.657, 0.782, and 83.05% respectively. The time taken by this regressor was 58.93 seconds. Compared to the previous models the prediction accuracy of this model is less and the error is high. The time taken by this model is high than the previous model.

(iii) Multi-layer Perceptron Classifier

The multi-layer perceptron (MLP) classifier optimizes the squared loss using gradient descent. The RMSE, MAE, MSE and R^2 reported were 0.65, 0.42, 0.687, and 74.45% respectively. The time taken by this regressor was 53.95 seconds. The prediction error for this model is very high as compared to previous models.

(iv) Support Vector Machine (SVM)

SVM belongs to the category of supervised machine learning widely used for two-group classification problems. SVM model sets are deployed for labeling training data for each category. New test data is also categorized in the same manner. The RMSE, MAE, MSE and R^2 reported were 0.77, 0.99, 0.245, and 65.45% respectively. The time taken by this regressor was 59.95 seconds

(v) Radial basis functions

Radial basis functions are used to estimated multivariable (also called multivariate) functions by linear combinations of relations originated on a single univariate function (the radial basis function). This is radialised so that in can be used in extra than one dimension. They are usually functional to approximate purposes or data which are only known at a limited number of points (or too difficult to evaluate otherwise), so that then valuations of the like function can take place often and efficiently. . The RMSE, MAE, MSE and R^2 reported were 0.75, 0.56, 0.134, and 73 % respectively. The time taken by this regressor was 58.95 seconds

V. Discussion

The proposed model is evaluated by considering four performance metrics namely, RMSE, R squared, MAE, and MSE.All the different Machine Learning Models for the Wind Turbine Dataset is evaluated and shown in the table 3.Figure 3, 4, 5 and 6 shows the RMSE, MAE, MSE and R^2 values for all five different

machine learning classifier models. It is observed from all figures the Random Forest is superior in comparison to other models and hence there is a surge in accuracy.

1.Anomalies detection of wind turbine gearbox using best model

Table 2 shows the training and testing of different classifiers on Wind Turbine Dataset [26]. It is observed that the Random Forest performs better in comparison to other classifiers. The WT GB is trained using the best-performed training algorithm; the fault in the GB is detected by comparing the actual and predicted temperature. Figure 7 shows the change in temperature of GB for 2 month period which eventually fails after 80 days. Figure 8 shows the difference between predicted and actual GB temperatures in these 3 months. The first noticeable change from the predicted model occurs at point A label in Figure 8, shows that the GB performance has deviated from normal behavior. The frequency of this deviation is increased further in point B and the temperature difference is 30°C and lasted for a long time. Point B in Figure8 shows that the turbine completely failed at point B due to the overheating of the GB. Thus, it is inferred that the anomalous behavior of WT GB is detected at point A which can be used as an early warning point. Point A in this method can be used as an acute warning point (alarm) as the anomalous behavior if frequency is increased. If the difference in time is calculated between warning at point A and the complete failure of the WT at point B it is 60 days then it means that the fault is detected 60 days earlier and this fault is fxed to avoid the failure at point B. The extremely unusual behavior of the GB model at point A is 20 days earlier than actual failure, wherein it is possible to stop the WT at point A and perform minor repair at this point to avoid the complete failure in point 3.

Table 2: Training and Testing of Different Machine Learning Models for the Wind Turbine Dataset

Data	model	RMSE	MAE	MSE	R^2
Training	**Random forest regressor**	**0.095**	**0.08**	**0.036**	**93.6%**
	KNeighbors regressor	0.56	0.057	0.563	83.05%
	Multi-layer perceptron regressor	0.65	0.102	0.387	74.45%
	SVM(support vector machine)	0.77	0.97	0.456	65.45%
	RBF(Radial Basis Function**)**	0.75	0.36	0.235	73%
Test	**Random forest regressor**	**0.096**	**0.023**	**0.026**	**92.6%**
	KNeighborsregressor	0.56	0.657	0.782	83.05%
	Multi-layer perceptron regressor	0.65	0.402	0.687	74.45%
	SVM(support vector machine)	0.77	0.99	0.245	65.45%
	RBF(Radial Basis Function**)**	0.75	0.56	0.134	73%

V. Conclusion

A federated Machine Learning solution is proposed in order detect any possible anomalies in critical infrastructure such as Wind Turbines. Instead of centralizing the wind-turbine data into a common server, *Federated Machine Learning* allows the data to remain on-premise in the infrastructure. Various federated machine learning models are deployed to analyze the anomalies in the wind turbines. The prediction error is an effective measure for anomaly detection in WTs. It is inferred from the experimental results that Random Forest is superior in identifying the anomalies with regard to the performance metrics namely RMSE, MAE and MSE are minimal for Random Forest in comparison to other models and Coefficient of Determination (R^2) is higher which is expected for a model that accurately predicts the anomalies. In addition, the time taken by the regression is 42.95 seconds which is minimal in comparison to other classifiers. Thus, the proposed method analyzes the anomalies in critical wind turbine infrastructures effectively and it can be used in real time anomaly detection for other critical infrastructures.

Declaration

- Ethical Approval and Consent to participate

Not applicable

- Consent for publication

All the authors of this paper have shown their Participation voluntarily.

- Availability of supporting data

The data will be provided based on data request by the evaluation team.

- Competing interests

The authors of this research article declares that no conflict of interest in preparing this

research article.

- Funding

This Research Received no specific grant from any funding agency in the public,

commercial or not-for-profit sectors.

- Authors' contributions

The authors develop a privacy preserving *Federated Machine Learning* solution is proposed in order detect any possible anomalies in such infrastructures.

References

1. Du M, Ma S, He Q (2016, August) A SCADA data based anomaly detection method for wind turbines. In *2016 China International Conference on Electricity Distribution (CICED)* (pp. 1–6). IEEE
2. Zhao Y, Li D, Dong A, Lin J, Kang D, Shang L (2016, September) Fault prognosis of wind turbine generator using SCADA data. In *2016 North American Power Symposium (NAPS)* (pp. 1–6). IEEE
3. Sun P, Li J, Yan Y, Lei X, Zhang X (2014, September) Wind turbine anomaly detection using normal behavior models based on SCADA data. In *2014 ICHVE International Conference on High Voltage Engineering and Application* (pp. 1–4). IEEE
4. Rosa L, Cruz T, de Freitas MB, Quitério P, Henriques J, Caldeira F, … Simões P (2021) Intrusion and anomaly detection for the next-generation of industrial automation and control systems. Future Generation Computer Systems 119:50–67
5. Yang H, Cheng L, Chuah MC, Deep-learning-based network intrusion detection for SCADA systems, in (2019) IEEE Conference on Communications and Network Security (CNS), 2019, pp. 1–7, http://dx.doi.org/10.1109/cns.2019.8802785
6. Rashid H, Batunlu C (2021) Anomaly Detection of Wind Turbine Gearbox based on SCADA Temperature Data using Machine Learning
7. Shlomo A, Kalech M, Moskovitch R (2021) Temporal pattern-based malicious activity detection in SCADA systems. Computers Security 102:102153
8. Xiang L, Wang P, Yang X, Hu A, Su H (2021) Fault detection of wind turbine based on SCADA data analysis using CNN and LSTM with attention mechanism. Measurement 175:109094
9. Sheng C, Yao Y, Fu Q, Yang W (2021) A cyber-physical model for SCADA system and its intrusion detection. Comput Netw 185:107677
10. Mokhtari S, Abbaspour A, Yen KK, Sargolzaei A (2021) A Machine Learning Approach for Anomaly Detection in Industrial Control Systems Based on Measurement Data. Electronics 10(4):407
11. Roelofs CM, Lutz MA, Faulstich S, Vogt S (2021) Autoencoder-based Anomaly Root Cause Analysis for Wind Turbines. Energy and AI, p 100065
12. Chandola V, Banerjee A, Kumar V (2009) Anomaly detection: A survey. ACM Comput Surveys 41(3):1–58
13. MeikSchlechtingen IF, Santos, SofianeAchiche (2013) Wind turbine condition monitoring based on SCADA data using normal behavior models. Part 1: system descripition. Appl Soft Comput J 13:259–270
14. Zaher A, McArthur SDJ, Infield DG, Patel Y (2009) Online wind turbine fault detection through automated SCADA data analysis. Wind Energy 12:574–593
15. Phillips B, Gamess E, Krishnaprasad S, An evaluation of machine learning based anomaly detection in a scada system using the modbus protocol, in:Proceedings of the 2020 ACM Southeast Conference, 2020, pp. 188–196
16. Anton SDD, Sinha S, Schotten HD, Anomaly-based intrusion detection in industrial data with SVM and random forests, 2019, URL http://arxiv.org/abs/1907.10374, arXiv:1907.10374
17. Sokolov AN, Alabugin SK, Pyatnitsky IA, Traffic modeling by recurrent neural networks for intrusion detection in industrial control systems, in:2019 International Conference on Industrial Engineering, Applications and Manufacturing, ICIEAM 2019, IEEE, 2019, pp. 1–5,http://dx.doi.org/10.1109/ICIEAM.2019.8742961
18. Sun P, Li J, Wang C, Lei X, "A generalized model for wind turbine anomaly identification based on SCADA data,"Applied Energy, vol. 168, pp. 550–567, 2016
19. Xiang L, Wang P, Yang X, Hu A, Su H (2021) Fault detection of wind turbine based on SCADA data analysis using CNN and LSTM with attention mechanism. Measurement 175:109094
20. Slay J, Miller M Lessons learned from the maroochy water breach. In Proceedings of the International Conference on Critical
21. Infrastructure Protection, Hanover NH (2007) USA, 19–21 March 2007. Springer, Berlin/Heidelberg, pp 73–82
22. Lyu L, Han Yu, and Yang Q. "Threats to federated learning: A survey." *arXiv preprint arXiv:2003.02133* (2020)
23. Bassily R, Smith A, and AbhradeepThakurta. "Private empirical risk minimization: Efficient algorithms and tight error bounds." *2014 IEEE 55th Annual Symposium on Foundations of Computer Science*.IEEE, 2014
24. Rogers RM, Roth A, Ullman J, Vadhan S Privacy odometers and filters: Pay-as-you-go composition. In Advances in Neural Information Processing Systems, pp. 1921–1929, 2016

25. Abadi M, Chu A, Goodfellow I, McMahan HB, Mironov I, Talwar K, Zhang L. Deep learning with differential privacy.In Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security, pp. 308–318.ACM, 2016
26. Ramotsoela DT, Hancke GP, Abu-Mahfouz AM, Attack detection in water distribution systems using machine learning, Human-Centric Computing Information Sciences, 9, 2019, http://dx.doi.org/10.1186/s13673-019-0175-8
27. **Note on the SCADA, Status and Warning Data Accessed on: march 23, 2021. [online] Available:** https://wt-fdd.readthedocs.io/

Figures

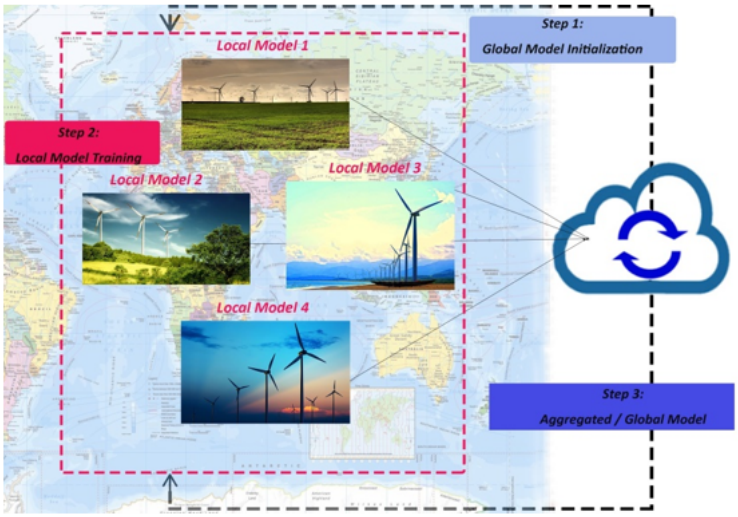

Figure 1

Architecture of federated learning for wind turbine

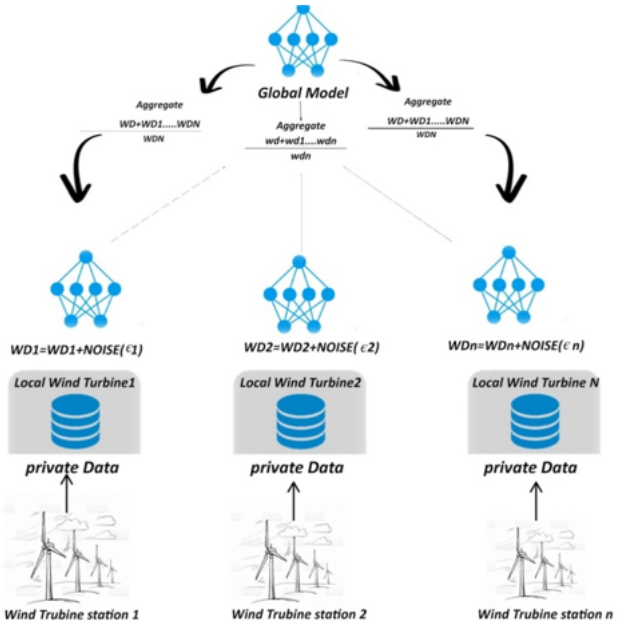

Figure 2

Proposed architecture wind turbine privacy-preserving federated learning analysis

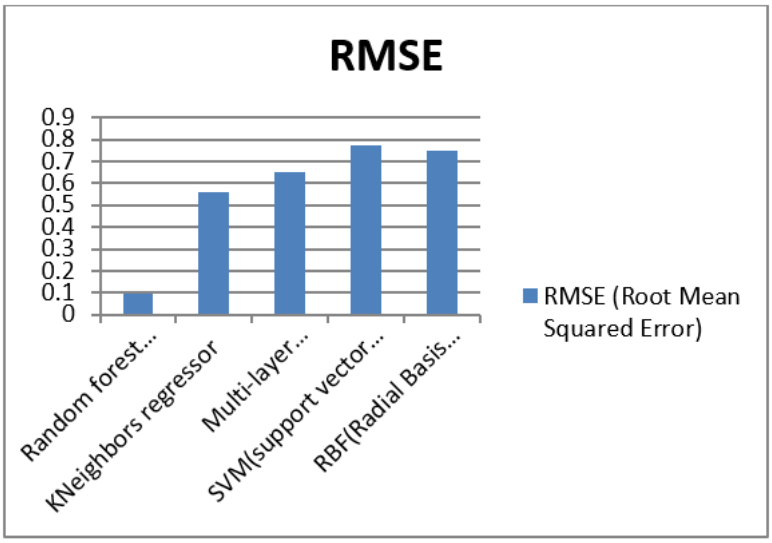

Figure 3

Comparative Analysis of RMSE On Various Federated Classifiers

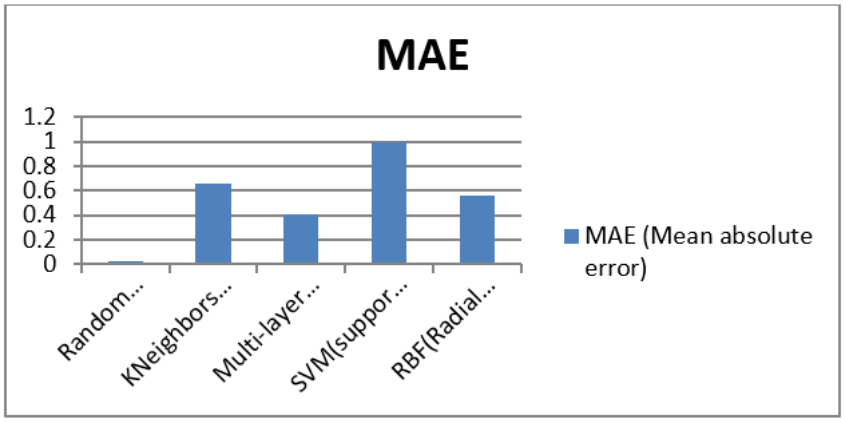

Figure 4

Comparative Analysis of MAE On Various Federated Classifiers

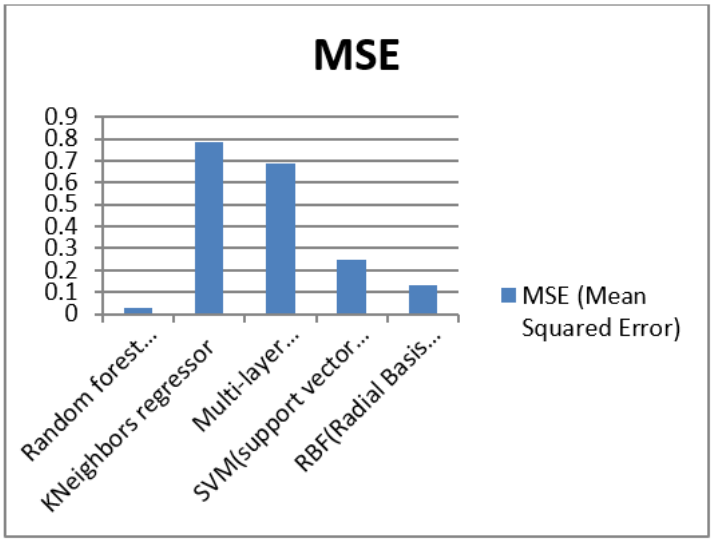

Figure 5

Comparative Analysis of MSE on Various Federated Classifiers

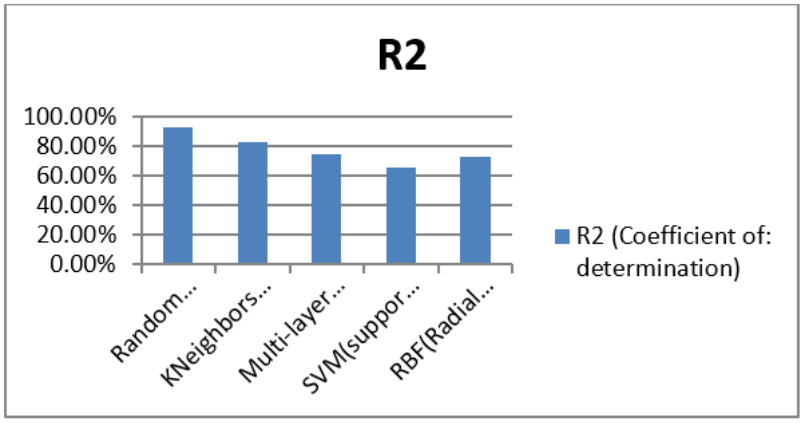

Figure 6

Comparative Analysis of R2 On Various Federated Classifiers

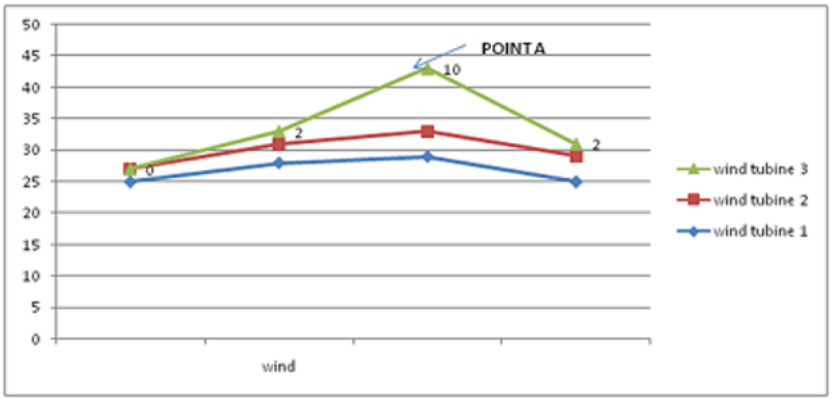

Figure 7

Change in temperature of GB in 2 month

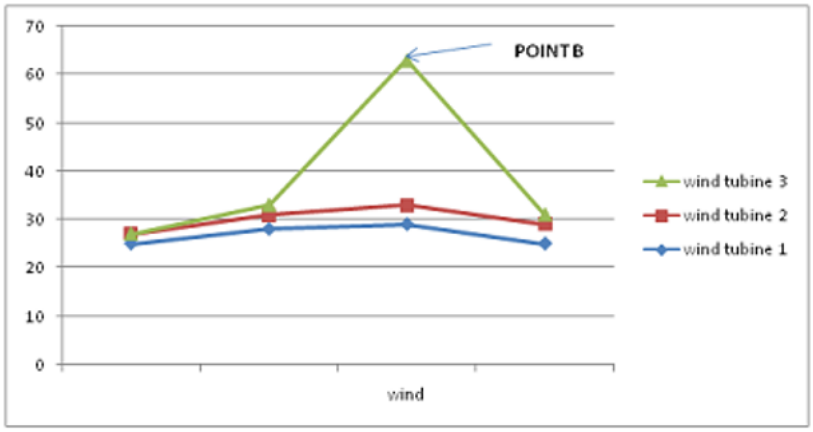

Figure 8

Change in temperature of GB in 80 days