XG-Boost Algorithm Based Anomaly Detection for Industrial Control System.

Niranjan Kundap[1*], Suresh Satpute[2], Trushita Chaware[3]

***Corresponding Author:-** Niranjan Kundap
[1] M.Tech Student, School of ECE, Dr. Vishwanath Karad MIT World Peace University, Pune, India
[2] Automation Solutions, PlantWeb Technologies, Emerson Innovation Center, Pune, India
[3] Professor, School of ECE, Dr. Vishwanath Karad MIT World Peace University, Pune, India
E-mail:[1] kundapniranjan@gmail.com
E-mail: [2] suresh.satpute@emerson.com
E-mail: [3]trushita.chaware@mitwpu.edu.in

Abstract.

Industrial Control Systems (ICSs) are widely applied in critical infra-structures to provide essential utilities to society. As a result, protecting them against terrorist activities, defects and cyber-attacks are crucial. The tremendous rise in cyber-attacks against Industrial Control Systems has triggered the advancement of security systems for accurate and timely detection of ICS cyber-attacks. In this paper, machine learning based ensemble XG-Boost algorithm for anomaly detection in industrial control systems proposed to identify cyber-attacks in ICS environments. Intrusion detection system based on anomalies is more efficient than other forms of intrusion detection system. Proposed system is a build with multiple machine learning algorithms, on Hardware-in-the-Loop (HIL) industrial process dataset (HAI). The HAI dataset was gathered from a realistic industrial control system testbed augmented with a hardware in the loop simulator that simulated pumped storage hydropower generation and steam-turbine power generation. XG-ADICS consist of 4 main steps, starting from data pre-processing then model training followed by model testing and at last classification of anomalous or normal observation. System findings demonstrate that by employing XG-ADICS, we can attain greater ac-curacy 99.96 % with a higher F1 score of 0.9986. The results suggest that the XG-Boost algorithm outperforms other algorithms in finding abnormalities. results demonstrates that the suggested system is appropriate for an ICS environment.

Keywords: Cyber Physical Systems, Cyber Physical Security, Industrial Control System, Machine Learning, Intrusion Detection System.

DOI Number: 10.14704/nq.2022.20.9.NQ44193 **Neuro Quantology 2022; 20(9):1677-1689**

1 Introduction

1.1 Overview

IT cyber-attacks are now routine for everyone; however, people may not be aware of OT (Operational Technology) cyber-attacks. Industrial Control Systems (ICS), Distributed Control Systems (DCS), Programmable Logic Controllers (PLC), Supervisory Control and Data Acquisition (SCADA), and Human Machine Interfaces (HMI) are all ubiquitous components of operational technology. Production, smart grid, rail, electricity generation and distribution, chemical manufacturing, oil and gas, aviation, and so on and so forth utilities all require OT components. Prior to a few years ago, OT systems and IT systems were completely isolated from one another.

Moreover, convergence of OT infrastructure and IT infrastructure helps to connect OT plants from physical layer to cloud layer which extends the coverage, connectivity, and functionality of OT plant also it is vital for Industrial IoT and Industry 4.0, on other hand it also open the way to several cyber-attacks on OT/ICS plants/devices.

Cyber physical systems are built on the real-time interchange of numerous forms of data and sensitive. information between cyber and the 2010 Stuxnet attack targeted on Iran's nuclear power plant, in 2016 cyber-attack on a Ukraine power plant. In 2015, the Black Energy cyberattack caused a power plant in Ukraine to fail, resulting in a severe power outage. A cyber assault on a German steel factory in 2014 took control of industrial machinery, and several blast furnaces were destroyed. In May 2021, American Colonial Pipeline, the East Coast's biggest gasoline, diesel, and natural gas provider, was targeted by a ransomware assault. These are just a few attacks, there are n numbers of large-scale attacks happening on ICS throughout the world [4]. Hence, there is an impending need to secure such ICSs against such attacks. More importantly, the impact of such assaults led to the awareness that new techniques and approaches outside the traditional mechanism are necessary to safeguard ICS. The danger of OT cyberattacks is projected to rise and failure to preserve industrial application's security and privacy can result in massive financial and economic losses.[2]

1.2 Associated Risk and Losses:

Research reported by Ponemon institute state of Industrial Cybersecurity, report says,[5] typical cost of small ICS cyber security attack on OT/ICS to identify, investigate, and remediate, as well as an incident response team and some obligatory replacement of equipment, physical devices, downtime, legal and regulatory fines, is minimum around $3 million to the company. A large-scale OT/ICS cyber assault can cost a firm more than $100 million. In 2017 OT attack happened on Shipping and Logistics company A.P. Mollar Maersk, to resolve entire issue it costed them $300 Million. [5]. According to Gartner, the financial impact of CPS attacks resulting will exceed $50 billion by 2023[6]. Organizations need to realize that OT cyberattacks not only cause financial losses, but there could also be loss of property, reputational damage, permanent and irreversible damage to personnel, the environment, and even national security. It has become crucial for organizations to implement OT Cybersecurity system before it is too late.

1.3 Problem Statement:

As ICS systems are frequently challenged by malicious threats and cyberattacks, a successful attack in an ICS environment may result in the disruption of critical services, the destruction of instruments, and the alteration of operational data, all of which can have catastrophic consequences. Effective and reliable ICS security and protection are necessary to guarantee that everything runs smoothly. The attacks on ICS highlighted the need for new methods and tools beyond traditional mechanism are obligated and must secure ICS. Rather concentrating on network traffic, packets, and signature-based IDS, we proposed the detection mechanism based on cyber physical device readings and values incorporating with machine learning techniques to identify intrusions/attacks. Anomaly detection method based on machine learning is proposed to protect and safeguard the ICSs.

1.4 Existing Solution and Gaps Findings:

The IT-centric solutions are insufficient and incapable of protecting an Industrial Control System. These solutions do not address potential vulnerabilities that may arise when IT systems interact with physical devices or the ICS environment. A firewall breach may go unreported while intruder manipulates operational data, causing PLCs to transmit erroneous orders and commands putting the plant in hazardous circumstances. Now, detecting OT/ICS cyber-attacks, Intrusion Detection System is the well-known solution. Traditional IDS used signature-based approach; this sort of IDS necessitates the use of a predetermined dictionary of attack patterns. if some pattern seen within process operation matches to any one of the predefined attack patterns, it identifies it as an intrusion. Although this method has a low false positive rate, but it does not identify zero-day assaults.

In complicated physical processes it is extremely challenging to develop a comprehensive dictionary of attack's signatures. The most signature-based intrusion detection systems become obsolete after short duration because in modern age, every day a new cyber-attack is developed by someone and somewhere, but that new attack pattern is not available in the pre-determined attack pattern dictionary. The signature-based IDS approach teaches the system to detect certain abnormalities, anomalies which system already knows but it completely fails to detect pattern which system don't know. [7]

1.5 Proposed Solution:

To overcome this problem, Behaviour based IDS can be a great solution. This technique is centered on operational system data from the physical process. based on the obtained data, a model is trained on the process's normal and abnormal behaviour accordingly and used to detect intrusions and abnormalities.

The proposed system focused on operational values from level 0 and level 1 w.r.t Purdue Model shown in figure 2. The aim of anomaly-based IDS is to discover abnormalities in data, Anomaly-based IDS looks for any divergence from expected behavior.

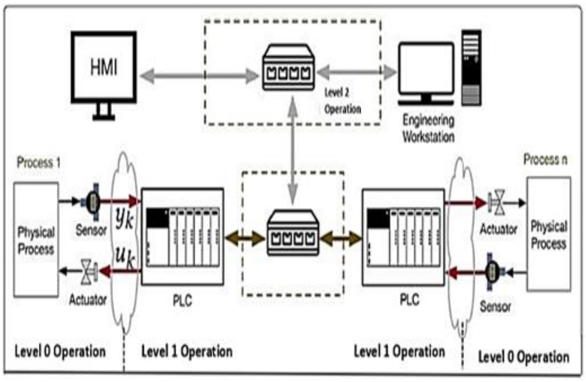

Fig. 2. Level 0 and Level 1 Operational Diagram [9]

Signature based IDS tries to inspect in network traffic, investigates suspicious packets, in contrast, anomaly-based IDS do not inspect or interact with network traffic, nor do they investigate packets; instead, anomaly-based IDS directly evaluate data streams of multiple parameters. It tries to identify any deviation from normal system performance due to malicious behaviours such as altering sensor setpoints or inserting fraudulent limit values and measurements. ICS attacks create the largest danger to an industrial site but detecting them is an extremely critical and complex process. Proposed system helps in the detection of these attacks, it will add a crucial layer of security to cyber-physical systems and is intended to secure ICS operations regardless to the form of an attack.

The remaining part of the paper is organized as follows. Section 2 describes the review of the literature. Section 3 provides explanation about the methodology for the proposed system. Section 4 presents experimental setup of data collection method and dataset information. Section 5 discusses about machine learning model which are used. Section 6 explains the model evaluation metrics for the proposed model, to validate the model. Section 7 reviews the output and result of proposed system and finally we conclude our research and future scope in section 8.

2 Literature Review:

In This [8] Paper, Measurement intrusion detection system (MIDS) has been offered from authors for detecting cyber-attacks and sense unusual activity in the SCADA's measurement data. MIDS SCADA detects any anomalous activities with accuracy of 99.76% incorporating with Random Forest(RF) algorithm. The paper [9] investigates on anomaly detection for ICS using a data-centric approach. Authors discuss the difficulties associated with supervised and unsupervised data sets. By reaching 99.93 percent detection rate by implementing Probabilistic neural network (PNN) with high detection rate and minimal false alarm. Authors of reference [10] discussed MADICS, an anomaly detection paradigm which is semi-supervised approach to predict ICS behaviors. They achieved cutting-edge precision of 0.984, achieved recall of 0.750 and F1-score of 0.851 in terms of detection performance incorporating with Secure Water Treatment (SWaT) dataset. The suggested method of [11], comprises k-means clustering to produce different coaching subgroups. Support Vector Machine (SVM) classification vector is constructed and radial SVM classifier is implemented to detect intrusion. Authors of reference [12] built three supervised machine learning algorithms, Support Vector Machine, Random Forest, and K-nearest neighbor and two unsupervised techniques, One Class-SVM and Autoencoder. SVM and RF provided the greatest overall accuracy and F1-Score results.

Reference [13] examines CPS related applications, technology, and standards. CPS security weaknesses, threats, and assaults are examined. The existing security methods and their challenges are described and examined, along with their key shortcomings. Study of [14] examines types of behavior-Based IDS in ICS, categorize it, find gaps, and recommend future research options. It also investigates signature-based, behavior-based and specification-based IDS. Paper [15] examines the ML-based security approaches for Cyber Physical Security system. This research compares many algorithms, including SVM, Isolation Forest, Deep Belief Network(DBN), Recurrent Neural Network(RNN), and Long Short-Term Memory(LSTM) to examine cyber-physical security in cyberspace. In reference [16] authors cover 150 papers to examine and review all of the machine learning strategies utilized for identifying anomaly detection, different sorts of anomalies as well as their drawbacks. Paper [17] authors suggest time dimension characteristics that are ignored by confusion matrix and its accompanying measures, to address this, authors provide three additional criteria for evaluating classification quality while accounting for temporal component of time-series data.

3 Methodology:

In this section, we discussed the complete process flow of XG-ADICS. Initially, a description of the Cyber Physical System attacks is provided, followed by an explanation of the methodologies and strategy used to identify those attacks. Figure 3 depicts a comprehensive architectural diagram.

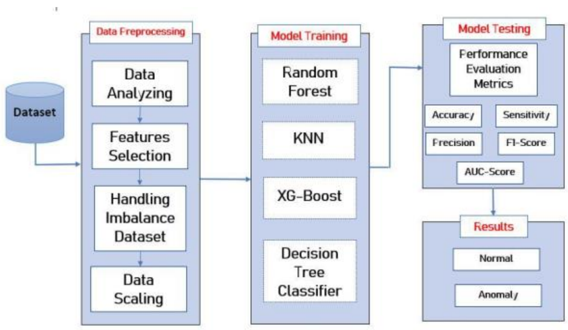

Fig. 3. : System Architecture Diagram

We employed the HAI (HIL-Based Augmented ICS) dataset for the proposed system, which eliminates the necessity for a data collecting procedure. The dataset we are utilizing is an unbalanced dataset, thus we must deal with it. To do so, we employed SMOTE (Synthetic Minority Oversampling Technique) to normalize the data. The scale of features in our dataset are not on the same scale, features if we do not scale features on the same scale, large scale features may dominate small scale features and model output can impacted. Data scaling technique has been used to overcome this issue. Lastly, the Feature Selection Technique is being applied to strive dimensionality reduction while simultaneously minimizing computation complexity, enhancing model accuracy, and making it easier for the algorithm to identify patterns in the data. Next phase is model training, which is a crucial procedure in which we build the model on a dataset using multiple algorithms. For model training, algorithms like Random Forest, KNN, XG-Boost and Decision Tree Classifier are being used to get the best results and on later stage we compare the performance of these algorithms.

Model testing is the next phase, where we test and validate the model, to do so we have used confusion metrics for model performance evaluation. We check accuracy, recall, precision, and F1-score of each method we utilized in final stage. the proposed system determines whether an observation is anomalous or normal on the testing dataset.

3.1 Threats Description:

In today's cybersecurity landscape, anomaly detection encompasses a wide range of disciplines, including fault detection, intrusion detection, event detection, and novelty/outlier identification in sensor networks. In an ICS, any divergence from regular performance may be characterized as an anomaly. It might be caused by several circumstances, including a cyberattack, system component failure, or an insider sabotage. The notion of anomaly detection is discussed in this work in terms of fault and intrusion detection. Anomalies in ICSs can be identified using measured data acquired by the ICS system, which has the benefit of identifying any deviation in normal behavior even though the intrusion is not detected at the network layer even by Network Intrusion Detection System. These assaults happen when an intruder manipulates sensor data or control signal resilience by gaining unobserved access to control networks until the system breaks.

3.2 Data Analyzing:

In IDS modelling, the issue of unbalanced datasets is very critical. Accessibility to the balanced dataset during training phase does have a significant influence on the model performance in machine learning modeling, particularly in classification issues. This issue arises in the model because the number of normal observations in the dataset we are considering is much larger than the number of abnormal observations. A variety of strategies, such as the cost-sensitive learning, one-class learning, and threshold method are introduced to address the problem of unbalanced datasets [18]. All balancing strategies rely on either oversampling or under sampling. In a nutshell, the under-sampling technique seeks to lower the number of observations drawn from the majority class, whereas oversampling tries to increase the number of samples from the minority group. Oversampling exposes the model at risk of overfitting, whereas undersampling risks losing essential information

To resolve this issue, the Synthetic Minority Oversampling Technique (SMOTE) technique [19] approach used to address this issue. The SMOTE technique uses random oversampling to produce new instances from existing data from uncommon classes with the help of KNN algorithm. SMOTE is a common oversampling method proposed by Chawla et al. Its basic objective is to generate new minority class data samples by interpolating and randomly selecting a nearby minority class neighbors. To begin, for each minority class data sample x, one obtains its k-nearest neighbors from other minority class data samples. Second, from the k neighbors, one selects one minority class data sample $\tilde{x}$. Finally, one generates the synthetic data sample x new by interpolating between x and $\tilde{x}$ as follows:

$x$new = x+ rand (0, 1) × (x − $\tilde{x}$),

As a result, by creating new samples (that are not identical to the old samples), this strategy avoids the possibility of overfitting. The SMOTE approach is used in this model to normalize the target column data observations due to the imbalance in the labelled data. Furthermore, the balanced dataset aids in the separation of normal and aberrant data throughout the process of creating the train and test datasets. The Train-Test Split (TTS) technique is used in this work to split the train and test data. 20% of the data has been chosen for testing data.

3.3 Feature Scaling:

Some machine learning algorithms are sensitive to feature scaling, while others are mostly unaffected by it. Having features that are on the same scale and they contribute equally to the result. It has the potential to increase the performance of machine learning algorithms. If we do not scale features, large scale features will dominate small scale features. Whether to use features scaling or not it depends upon the algorithms we are using, and the algorithms used in proposed model gets affected by it. Certain machine learning techniques, such as the distance-based algorithm, the gradient descent algorithm, the matrix factorization algorithm, and the curve-based algorithm are sensitive to feature scaling. Feature scaling is classified into two types, feature normalization and feature standardization. Scaling approach based on feature standardization in which data typically focused it around mean and have a unit standard deviation. As a result, the mean of the property is zero, and the resultant distribution has a standard deviation of one.

Feature Normalization is a scaling technique also termed as Min-Max Scaling that adjusts and rescales data values such that they range between 0 and 1. Equation for normalization is,

$$x' = \frac{x - x_{min}}{x_{max} - x_{min}}$$

Xmax denote the maximum and Xmin denote lowest values of the features. Because we are working with enormous continuous numerical data with various characteristics of varying scales, the Feature Normalization scaling approach is an excellent fit for our system.

3.4 Feature Selection:

The attributes / variables of measured data gathered by the CPS system via embedded sensors are referred to as features in proposed model. In large-scale ICSs, the number of sensors is usually large. This refers to two facts. First, the feature selection strategy has a significant influence on the model's performance in the proposed challenge; second, due to the massive amount of measured data, the model algorithm must be competent of speedy prediction to be viable in real-time applications. Feature selection is just the selection of necessary independent and relevant features that have a stronger relationship with the dependent feature will benefit in the development of a good model, making it more accurate and requiring less computing time. for feature selection Pearson's Correlation Matrix with Heatmap is used. In Heatmap technique, the linear link between two or more factors is measured via correlation. The most correlated features w.r.t the target feature would be picked to build the model in feature selection. Furthermore, features that do not correlate well with the target feature must be avoided and discarded. The Pearson's Correlation Matrix can describe as, The range of the 'r' is in-between -1 to +1. The value '0' indicates that there really is no relationship between the two variables. A positive relationship is shown by a value greater than zero, with an increase in one variable increasing the value of the other. A number less than '0' denotes a negative relationship between two variables, with an increase in one lowering the value of the other.

- r = Pearson Correlation Coefficient
- X_i = X Variable Samples
- Y_i = Y Variable Samples
- $\overline{X}$ = Mean of Values in X Variables
- $\overline{Y}$ = Mean of Values in Y Variables

The calculation simply boils down to dividing the covariance by the product of the standard deviations. The formula can also be written as,

$$\rho_{xy} = Correlation\ (x,y) = \frac{cov(x,y)}{\sqrt{var(x)}\sqrt{var(y)}}.$$

4 Experimental Setup:

A HIL-Based Augmented ICS (HAI) Security Dataset is created using a realistic industrial control system testbed and a hardware-in-the-loop (HIL) simulator. Hyeok-Ki Shin and three other researchers of The Affiliated Institute of ETRI in the Republic of Korea created this testbed and dataset. [20]

$$r_{XY} = \frac{\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \overline{X})^2}\sqrt{\sum_{i=1}^{n}(Y_i - \overline{Y})^2}}$$

A boiler, turbine, water-treatment component, and an HIL simulator were all part of the testbed. Water-to-water heat transfer at low pressure and moderate temperature is used in the boiler technique. The turbine procedure, on the other hand, entailed precisely mimicking the behaviour of an actual rotating machine using a rotor kit testbed. To achieve synchronization with the spinning speed of a steam-power generator, the boiler and turbine operations were linked with the HIL simulator. During the HIL simulation, water was pumped to the upper reservoir and then released into the lower reservoir according to a pumped-storage hydropower generating model. Three distinct controllers were used to regulate the three real-world processes, namely the boiler, turbine, and water treatment processes.[20]

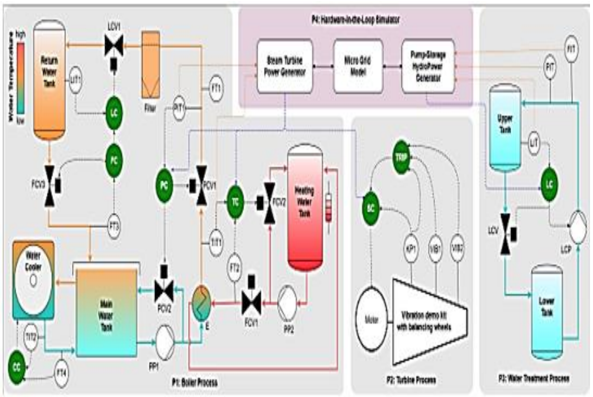

Fig. 4. HAL Testbed Diagram [21]

This ICS testbed dataset is available with open-source access for developing and evaluating models for detecting abnormalities in ICSs. Figure 4 depicts the process architecture of the HAL testbed. [21]. The testbed is divided into three main processes: a boiler (P1), a turbine (P2), a water-treatment process (P3), and an HIL simulator (P4). This process consists of four controllers (level controller (LC), temperature controller (TC), pressure controller (PC), and flow-rate controller (FC)

4.1 Data Set:

The ICS dataset, HAI 21.03, was created in the HAI testbed environment by the researchers at The Affiliated Institute of ETRI, Republic of Korea. The dataset is labelled dataset. The training data were created using a greater range of typical conditions, whereas the test data were generated by applying a greater variety of assault scenarios, allowing for the evaluation of detection performance in different situations [21]. All the assault scenarios were created using the four feedback control loop variables, SP, PV, CV, and control parameter (CP). When some of the parameters were outside of their regular ranges or in unexpected situations as a consequence of attacks, malfunctions, and failures, abnormal behavior ensued. HAI 21.03 has 78 data points in six process controllers. Data collection points were also utilized to generate a dataset of the controllers' inputs (setpoint (SP), process variables (PV), and outputs (Control variable (CV)).

5 Machine Learning models:

Multiple algorithms were utilized to construct the model that predicts the normal and anomalous in the HAI dataset. Due to the HAI dataset, which is basically a labelled dataset, supervised algorithms can be used to discover abnormalities in the dataset. XG-Boost, decision tree classifier, random forest, and KNN these algorithms has been used in the system.

KNN:

The k-Nearest Neighbor (KNN) method is a non-parametric algorithm that is used to solve classification and regression issues. The KNN algorithm assumes that comparable objects exist nearby. In other words, comparable objects are close together. In reality, the categorization is dependent on the features of objects adjacent to the one under consideration, which are measured using certain distance metrics, most prominently the Euclidean distance. The below is a derivation of the Euclidean distance:

$$D = \sqrt{\sum_{i=1}^{n} (x_i - p_i)^2},$$

Where, xi and pi are n-dimensional vector components. After calculating the distance between the considered point and all others, the class of the considered point is allocated as the most common among the first k points with the smallest distance.

Random Forest:

Random Forest is a type of ensemble learning algorithm. It is a method of supervised learning algorithm. It builds a "forest" out of a set of decision trees that are often trained using the "bagging" approach. The bagging strategy is predicated on the idea that merging multiple learning models optimizes final result. To put it simply, random forest builds large number of decision trees and then combines them to produce a more precise and consistent forecast. Random forest does have the key benefit in that it can be used for regression as well as classification problems that are prominent in machine learning systems.

XG-Boost:

XG-Boost which is known as extreme gradient boosting, a famous gradient boosting ensemble technique, which have better speed and performance based on tree-based sequential decision tree machine learning algorithms. Gradient boosting is the technique of strengthening single weak model by merging it with a number of other weak models in order to construct a collectively strong model. Tianqi Chen developed XG-Boost, which has been previously managed by the Distributed (Deep) Machine Learning Community group. XG-Boost is a boosting approach in Ensemble Learning. Ensemble learning consists of a group of predictors, which are several models that collaborate to increase prediction accuracy. By adding weights to the models, the boosting strategy aims to correct flaws introduced by previous models [22].

Decision Tree Classifier:

Decision Trees are a kind of Supervised Machine Learning in which data is continuously separated depending on a specific characteristic. Decision tree having two main nodes, the Decision Node, and the Leaf Node. Decision nodes have numerous branches, and it is used for decision making, while Leaf nodes are the outcome of such decisions and do not have any branches.

6 Model Evaluation Metrics:

Figure 4 depicts a Confusion matrix, which is a N x N matrix used to evaluate the performance of classification model. The matrix compares actual target values to the machine learning model's predictions. This gives us a comprehensive picture of how effectively our classification model performs and what kinds of errors it generates [22]. Furthermore, sensitivity, recall, precision, and F1-score metrics are generated for this matrix, in addition to accuracy, F1, Precision, and Recall being the most often used measures for assessing anomaly detection algorithms. Precision indicates how many of the accurately anticipated occurrences were genuinely positive. Recall reflects how many real positive cases we were able to properly foresee using our model. Because the F1-score is a harmonic mean of Precision and Recall, it presents a unified view of these two metrics. It is finest when Precision matches Recall. The confusion matrix is depicted in Figure 5. In addition to the Confusion matrix, ROC(Receiver Operating Frequency) and AUC(Area Under Curve) parameters used to evaluate the model performance by graphical depiction. The ROC is a graph which illustrates the classification algorithm's performance across all classification parameters compared to baseline model (tpr = fpr) which is 45 degree diagonal line in ROC curve. and includes two metrics: true-positive(tpr) and false-positive rates(fpr). The ROC curve depicts the trade-off among sensitivity (or TPR) versus specificity (1 – FPR). The area under the ROC curve, often termed as the area under the curve (AUC) metric, is used to compare ROC curves.

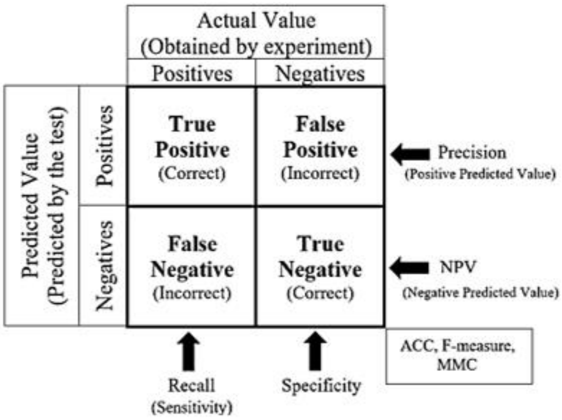

Fig. 5. Confusion Matrix

$$Recall = \frac{TP}{TP + FN} \qquad Precision = \frac{TP}{TP + FP}$$

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \qquad F1 - score = \frac{2}{\frac{1}{Recall} + \frac{1}{Precision}}$$

The formulas for Precision, Recall, Accuracy and F1-Score are shown above.

7 Results and Discussion:

The proposed XG-ADICS system is developed on the Visual Studio Code IDE using the Python programming language. The outcomes and outputs of the model phases are given and discussed in this section. Finally, a comparative analysis of different algorithms used is discussed given in table 1.

First, the technique is to deal with an imbalanced dataset that contains significantly uneven target observations. The SMOTE [19] technique was used to handle the imbalance dataset. The target column can only have two labels since the model will be looking for anomalies in the dataset: 1 for anomaly observations and 0 for normal observations. The orange portion shows the anomalous percentage, whereas the blue portion represents the normal percentage of observations. left pie diagram in figure 6 illustrates a dataset with 97.10 % normal observations and 2.90 % anomalous observations before executing SMOTE. right pie chart in figure 6 shows that after addressing the imbalanced data with the SMOTE technique[19], the dataset now comprises % normal observations and % anomalous observations, shown in figure 6.

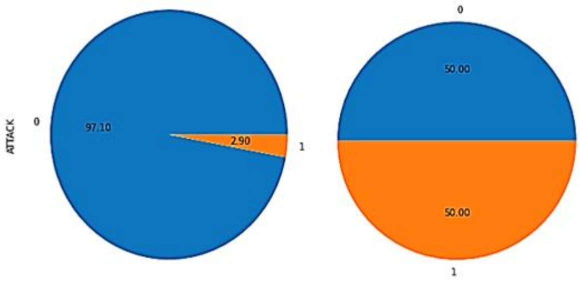

Fig. 6. SMOTE Result

Now that all features in the dataset have different scales, which impacts model performance, feature scaling is the next strategy used to give all features a scale between 0 and 1. To do this, the feature normalization technique is performed. Figures 7.1 depicts features are in different scale and 7.2 depict the feature scale changed to in between 0 to 1,figure 7.1 and 7.2 shows the before and after result of scaling.

	P1_B2004	P1_B2016	P1_B3004	P1_B3005	P1_B4002	P1_B4005	P1_B400B
0	0.09874	1.25036	382.21698	1118.61084	32.0000	99.98731	2835.65454
1	0.09874	1.24944	382.21698	1118.61084	32.0000	100.00000	2832.98731
2	0.09874	1.24746	382.21698	1118.61084	32.0000	99.98115	2842.25244
3	0.09874	1.24624	382.21698	1118.61084	32.0000	100.00000	2833.12744
4	0.09874	1.24364	382.21698	1118.61084	32.0000	100.00000	2839.44482
...	...	...	...	...	...	...	...
118796	0.10121	1.30608	397.63785	1001.99799	33.6555	100.00000	2830.74121
118797	0.10121	1.30226	397.63785	1001.99799	33.6555	100.00000	2839.72559
118798	0.10121	1.29707	397.63785	1001.99799	33.6555	100.00000	2833.96997
118799	0.10121	1.29906	397.63785	1001.99799	33.6555	100.00000	2842.95435
118800	0.10121	1.29784	397.63785	1001.99799	33.6555	100.00000	2847.02539

Fig. 7.1. Before Scaling

	P1_B2004	P1_B2016	P1_B3004	P1_B3005	P1_B4002	P1_B4005
0	0.210843	0.224717	0.000000	1.000000	0.190904	0.999873
1	0.210843	0.223761	0.000000	1.000000	0.190904	1.000000
2	0.210843	0.221705	0.000000	1.000000	0.190904	0.999811
3	0.210843	0.220439	0.000000	1.000000	0.190904	1.000000
4	0.210843	0.217739	0.000000	1.000000	0.190904	1.000000
...	...	...	...	...	...	...
230699	0.198795	0.385779	0.075406	0.721051	0.190904	0.000006
230700	0.954819	0.315866	0.374884	0.059385	1.000000	1.000000
230701	0.734940	0.243244	0.336600	0.804267	0.491474	0.762946
230702	0.198795	0.281282	0.075406	0.721051	0.190904	0.000063
230703	0.734940	0.322269	0.336600	0.804267	0.491474	0.957794

Fig. 7.2. After Scaling

The next strategy, Feature Selection, was performed since the dataset comprises around 84 features, which is a vast number of features, and not every feature has an influence on our target feature, to select right features, reduction of unwanted feature is an important procedure that eliminates unnecessary features and reduces computation time and complexity, improving the overall performance of the model. Pearson coefficients of correlation, The heatmap method used for this operation. Figure 8.1 depicts the correlation heatmap of all features. Figure 8.1 depicts some of the features that must be removed. Figure 8.2, only 20 useful features have been chosen for model development and all other features have been discarded.

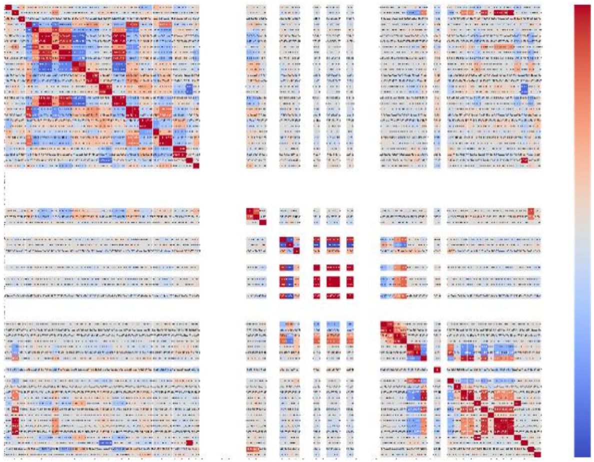

Fig. 8.1. Heatmap for all feature

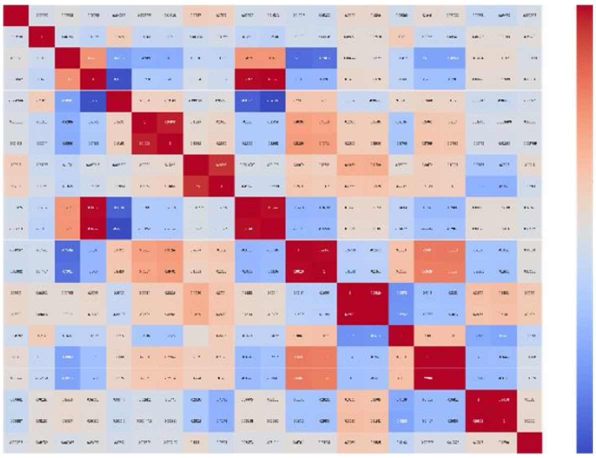

Fig. 8.2. Heatmap after selecting appropriate features

We evaluated several algorithms including Random Forest, K-Nearest Neighbors, XG-Boost, and Decision Tree Classifier. All of these algorithms' performance has been tested using the Confusion Matrix approach, as well as ROC (Receiver Operating Characteristics) and AUC (Area Under Curve). Confusion Matrix describes several metrics, including True Positive, True Negative, False Positive, False Negative, by this we can derive the Accuracy, Recall, Precision, and Sensitivity of the ML model. We can also calculate the ROC and AUC of the ML model using these parameters. Figures 9,10,11,12 depict the confusion matrix score and figures 9.1,10.1,11.1,12.1 depict the ROC curve for all algorithms respectively. Performance of algorithm can be found out with the help of confusion matrix of each algorithm.

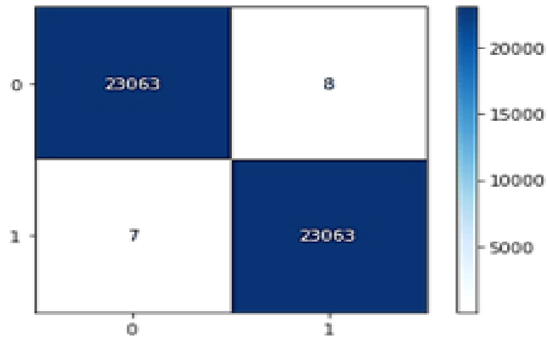

Fig. 9. Confusion Matrix for XG-Boost

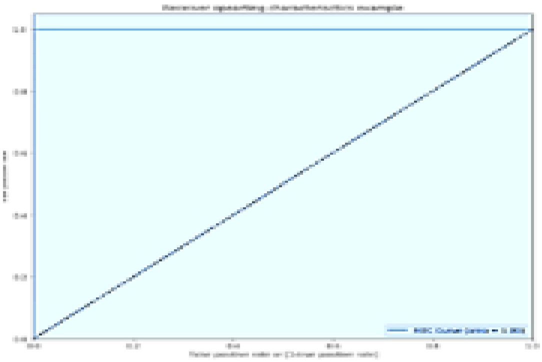

Fig.9.1 ROC Curve

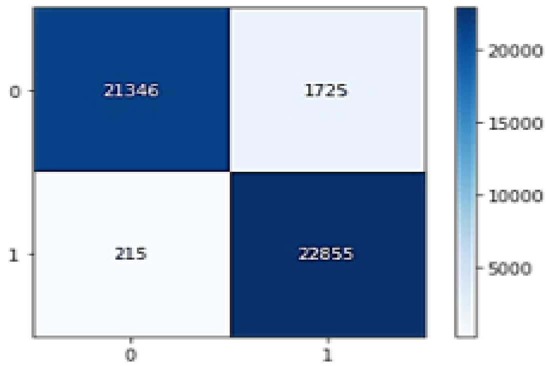

Fig. 10. Confusion Matrix for KNN

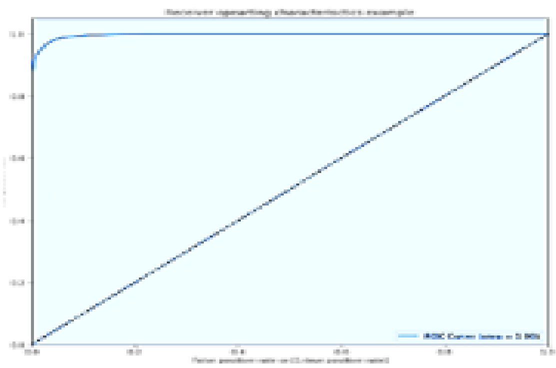

Fig10.1 ROC Curve

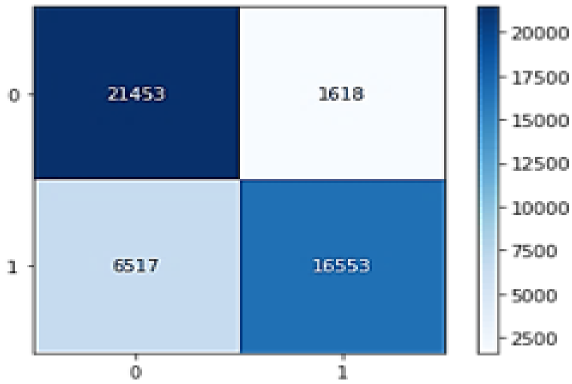

Fig. 11. Confusion Matrix for Random Forest

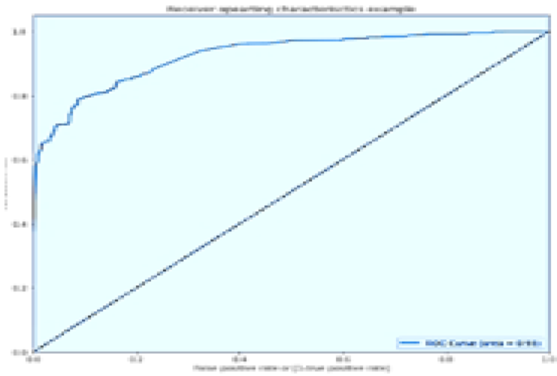

Fig11.1 ROC Curve

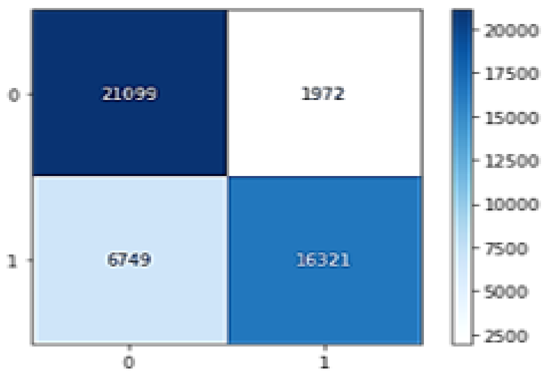

Fig. 12. Confusion Matrix for Decision Tree

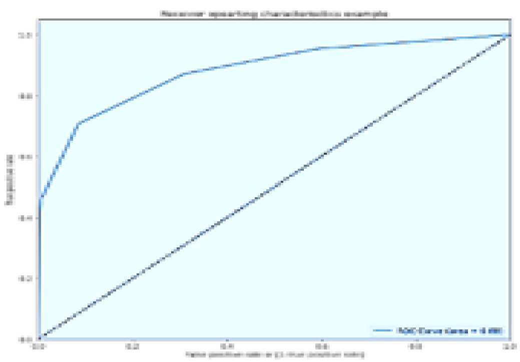

Fig.12.1 ROC Curve

In figure 9, confusion matrix for XG-Boost reveals that it identified 23063 True Positive (TP) and 23063 True Negative (TN) observations, 8 False Positive (FP) and 7 False Negative (FN) observations, which is the best detection score among all detection scores. The KNN algorithm in figure 10, works good with 21453 TP and 16553 TN with 215 FN and little higher FP of 1725. In figure 11 shows, random forest losses the accuracy with just 21453 TP and very low 16552 TN with very high 6517 FN and 1618 FP score. Whereas, by figure 12, Decision Tree failed poorly with score of 21099 TP, 16321 TN with high 6749 FN and 1972 FP.

Table 1 is illustrating the comparison of performance evaluation score for all algorithm has been plotted in tabular format for easy understandability. Figure 13 portrays a comparison between ROC curve and AUC value of all algorithms. AUC score ranges from 0 to 1, greater the AUC value, the better the model performance. When compared to other algorithms, the decision Tree classifier performed badly in detecting anomalies, RF performed mediocrely, and KNN achieved an impressive 0.99 AUC score but did not outperform XG-Boost in other performance measures, while XG-Boost performs exceptionally well in multiple measure, including accuracy, precision, sensitivity, F1-score, and AUC score.
The table 1 demonstrates that the XG-Boost classifier surpasses the other algorithms in terms of performance parameters in every aspect, achieving an accuracy of 99.967%, 99.969% sensitivity and 0.9986 of f1-score, with the AUC score of 0.9999.

Table 1. Model performance score for all algorithms

Metrics / Classifier	Accuracy	Precision	Sensitivity	F1-Score	AUC Score
XG-Boost	0.999675	0.999653	0.999696	0.998675	0.999997
KNN	0.957954	0.925230	0.990680	0.95925	0.999286
Random Forest	0.823953	0.929886	0.718032	0.803093	0.918384
Decision Tree Classifier	0.810992	0.876531	0.707455	0.789159	0.882054

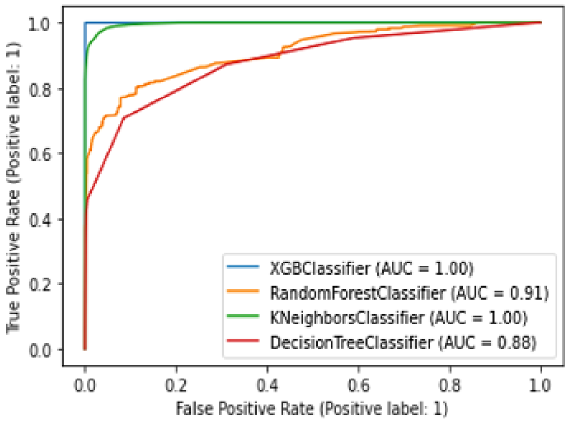

Fig. 13. All ROC and AUC for all algorithm

8 Conclusion and Future Scope:

In critical infrastructure like industrial control systems, Machine learning is increasingly being used to develop anomaly detection. XG-ADICS able to identify abnormalities, but it cannot avoid hostile incursions in the network layer. The network layer intrusion detection system is still needed to protect the ICS because it acts as the first line of defense, preventing malicious packets from entering the system. However, even if the attacker successfully penetrates the network layer security and attempts to manipulate the set points, XG-ADICS will detect it. System will detect the intrusion and will avoid the destruction of any ICS process or plant by intruder. the results demonstrate that with XG-Boost algorithms model is working very well with 99.96 % of accuracy and 0.99 AUC score for detecting anomalies with very minor false detection rate on HAI dataset. We can conclude that, XG-ADICS is capable and efficient for identifying OT attacks, system failures, and insider sabotage by detecting anomalies in industrial control systems.

To make the XG-ADICS even better and more superior, objectives can be met by incorporating with Deep Learning(DL) and Neural Networks techniques such as Recurrent Neural Networks (RNN), Generative Adversarial Networks(GAN), and Autoencoders(AE). ICS process generates tremendous amount of data; as data volume increases, the performance of ML-based models degrades, whereas performance of DL-based models improves. By incorporating the incremental machine learning technique, the system can be continuously taught with the most recent ICS process data, enhancing the system's reliability, and lowering the false detection rate.

Acknowledgment:

Authors would like to thank *Emerson Innovation Center, Pune* and *School of ECE, Dr. Vishwanath Karad MIT World Peace University, Pune, India* for the continuous support and guidance during this work, also authors would like to thank Researchers at *The Affiliated Institute of ETRI, Republic of Korea* for making ICS dataset available in open-source.

References:

What Is the Purdue Model for ICS Security?, https://www.zscaler.com/resources/security-terms-glossary/what-is-purdue-model-ics-security, last accessed 2022/05/15.

McLennan and Zurich Insurance Group: The Global Risks Report 2020. National University of Singapore , Oxford Martin School, University of Oxford (2020).

X-Force Threat Intelligence Index 2022. IBM Corporation, United States of America (2022).

Alladi, T., Chamola, V., Zeadally, S.: Industrial Control Systems: Cyberattack trends and countermeasures. Computer Communications. 155, 1–8 (2020). https://doi.org/10.1016/j.comcom.2020.03.007

The 2021 State of Industrial Cybersecurity. Ponemon Institute© Research Report, Ponemon Institute (2021).

Financial Impact of Cyber-Physical System Attacks, https://www.gartner.com/en/newsroom/press-releases/2020-09-01-gartner-predicts-75--of-ceos-will-be-personally-liabl, (2020).

Khraisat, A., Gondal, I., Vamplew, P., Kamruzzaman, J.: Survey of intrusion detection systems: techniques, datasets and challenges. Cybersecur. 2, 20 (2019). https://doi.org/10.1186/s42400-019-0038-7.

Mokhtari, S., Abbaspour, A., Yen, K.K., Sargolzaei, A.: A Machine Learning Approach for Anomaly Detection in Industrial Control Systems Based on Measurement Data. Electronics. 10, 407 (2021). https://doi.org/10.3390/electronics10040407

M. R., G.R., Ahmed, C.M., Mathur, A.: Machine learning for intrusion detection in industrial control systems: challenges and lessons from experimental evaluation. Cybersecur. 4, 27 (2021). https://doi.org/10.1186/s42400-021-00095-5

Perales Gómez, Á.L., Fernández Maimó, L., Huertas Celdrán, A., García Clemente, F.J.: MADICS: A Methodology for Anomaly Detection in Industrial Control Systems. Symmetry. 12, 1583 (2020). https://doi.org/10.3390/sym12101583

Saikushwanth, V., Rao, G.R.: Intrusion Detection System Using Machine Learning. United International Journal for Research & Technology (UIJRT). Volume 02,.

Bernieri, G., Conti, M., Turrin, F.: Evaluation of Machine Learning Algorithms for Anomaly Detection in Industrial Networks. University of Padua, Italy.

Yaacoub, J.-P.A., Salman, O., Noura, H.N., Kaaniche, N., Chehab, A., Malli, M.: Cyber-physical systems security: Limitations, issues and future trends. Microprocessors and Microsystems. 77, 103201 (2020). https://doi.org/10.1016/j.micpro.2020.103201

Umer, M.A., Junejo, K.N., Jilani, M.T., Mathur, A.P.: Machine learning for intrusion detection in industrial control systems: Applications, challenges, and recommendations. International Journal of Critical Infrastructure Protection. 100516 (2022). https://doi.org/10.1016/j.ijcip.2022.100516

Kim, S., Park, K.-J.: A Survey on Machine-Learning Based Security Design for Cyber-Physical Systems. Applied Sciences. 11, 5458 (2021). https://doi.org/10.3390/app11125458

Nassif, A.B., Talib, M.A., Nasir, Q., Dakalbab, F.M.: Machine Learning for Anomaly Detection: A Systematic Review. IEEE Access. 9, 78658–78700 (2021). https://doi.org/10.1109/ACCESS.2021.3083060

Kovács, G., Sebestyen, G., Hangan, A.: Evaluation metrics for anomaly detection algorithms in time-series. Acta Universitatis Sapientiae, Informatica. 11, 113–130 (2019). https://doi.org/10.2478/ausi-2019-0008

Kotsiantis, S., Kanellopoulos, D., Pintelas , P.: Handling imbalanced datasets: A review. GESTS International Transactions on Computer Science and Engineering. Vol.30,.

Chawla, N.V., Bowyer, K.W., Hall, L.O., Kegelmeyer, W.P.: SMOTE: Synthetic Minority Over-sampling Technique. jair. 16, 321–357 (2002). https://doi.org/10.1613/jair.953

Shin, H.-K., Lee, W., Yun, J.-H., Kim, H.: HAI 1.0: HIL-based Augmented ICS Security Dataset. In: USENIX Security'20 . , Republic of Korea (2020).

Shin, H.-K., Lee, W., Yun, J.-H., Min, B.-G.: Two ICS Security Datasets and Anomaly Detection Contest on the HIL-based Augmented ICS Testbed. In: Cyber Security Experimentation and Test Workshop. pp. 36–40. ACM, Virtual CA USA (2021). https://doi.org/10.1145/3474718.3474719.

Chen, T., Guestrin, C.: XGBoost: A Scalable Tree Boosting System. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. pp. 785–794. ACM, San Francisco California USA (2016). https://doi.org/10.1145/2939672.2939785.

A Survey on Concepts, Applications, and Challenges in Cyber-Physical Systems. KSII TIIS. 8, (2014). https://doi.org/10.3837/tiis.2014.12.001.

Jones, A., Kong, Z., Belta, C.: Anomaly detection in cyber-physical systems: A formal methods approach. In: 53rd IEEE Conference on Decision and Control. pp. 848–853. IEEE, Los Angeles, CA, USA (2014). https://doi.org/10.1109/CDC.2014.7039487

Gyamfi, N., Katsriku, F.K.: Anomaly detection framework for big data. LAMBERT Academic Publishing (2019).