Heliyon 7 (2021) e05936

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Power system events classification using genetic algorithm based feature weighting technique for support vector machine

Oyeniyi Akeem Alimi [a,*], Khmaies Ouahada [a], Adnan M. Abu-Mahfouz [a,b], Suvendi Rimer [a]

[a] Department of Electrical & Electronic Engineering Science, University of Johannesburg, Johannesburg 2006, South Africa
[b] Council for Scientific and Industrial Research, Pretoria, South Africa

ARTICLE INFO

Keywords:
Classification
Genetic algorithm
Power system
Support vector machine
Synchrophasors

ABSTRACT

Currently, ensuring that power systems operate efficiently in stable and secure conditions has become a key challenge worldwide. Various unwanted events including injections and faults, especially within the generation and transmission domains are major causes of these instability menaces. The earlier operators can identify and accurately diagnose these unwanted events, the faster they can react and execute timely corrective measures to prevent large-scale blackouts and avoidable loss to lives and equipment. This paper presents a hybrid classification technique using support vector machine (SVM) with the evolutionary genetic algorithm (GA) model to detect and classify power system unwanted events in an accurate yet straightforward manner. In the proposed classification approach, the features of two large dimensional synchrophasor datasets are initially reduced using principal component analysis before they are weighted in their relevance and the dominant weights are heuristically identified using the genetic algorithm to boost classification results. Consequently, the weighted and dominant selected features by the GA are utilized to train the modelled linear SVM and radial basis function kernel SVM in classifying unwanted events. The performance of the proposed GA-SVM model was evaluated and compared with other models using key classification metrics. The high classification results from the proposed model validates the proposed method. The experimental results indicate that the proposed model can achieve an overall improvement in the classification rate of unwanted events in power systems and it showed that the application of the GA as the feature weighting tool offers significant improvement on classification performances.

1. Introduction

Modern power systems rely on advanced technologies, communication networks and other sophisticated tools for efficient monitoring, control and operation of physical equipment. These tools offer varying contributions in terms of performance and reliability enhancement as some of them are designed and equipped with a variety of outstanding features for efficiency and effectiveness. However, the increasing use of these devices and other cyber presence increases the vulnerabilities and channels through which sabotage, terrorism and intrusion can be perpetuated into the network [1, 2, 3]. The consequences of various high-profile incidences such as the Yemen blackout, Ukrainian power grid blackout, Iran nuclear program attack, etc. have shown the devastating effects of intrusion and unwanted events on power infrastructures [1, 3]. Also, as modern power system networks spread across wide geographical landscapes, substations and transmission lines that travel over thousands of miles can be physically assaulted by adversaries. In summary, some of the security challenges modern power systems are facing are presented in Figure 1.

In the literature, several studies have discussed the threats, effects and impacts of these security challenges [4, 5, 6]. Adversaries can access network nodes, alter control commands and inject attacks such as the denial of service (DOS) attack, thereby destabilizing the operation, creating blackouts, financial losses, and in some situations, national security can be put into jeopardy. As power systems plays a major role in today's world, there is a growing need for adequate monitoring of the events at all layers of the network [7]. The introduction of technologies such as the phasor measurement units (PMUs) have significantly improved the possibilities of monitoring and analysing power system dynamics [8]. PMUs provide time synchronized data which include voltage and current phasors, relays, switches, circuit breakers statuses, etc. to control centres thereby enabling the accurate monitoring and identification of events [8, 9, 10].

* Corresponding author.
E-mail address: oalimi@uj.ac.za (O.A. Alimi).

https://doi.org/10.1016/j.heliyon.2021.e05936
Received 5 October 2020; Received in revised form 9 December 2020; Accepted 6 January 2021

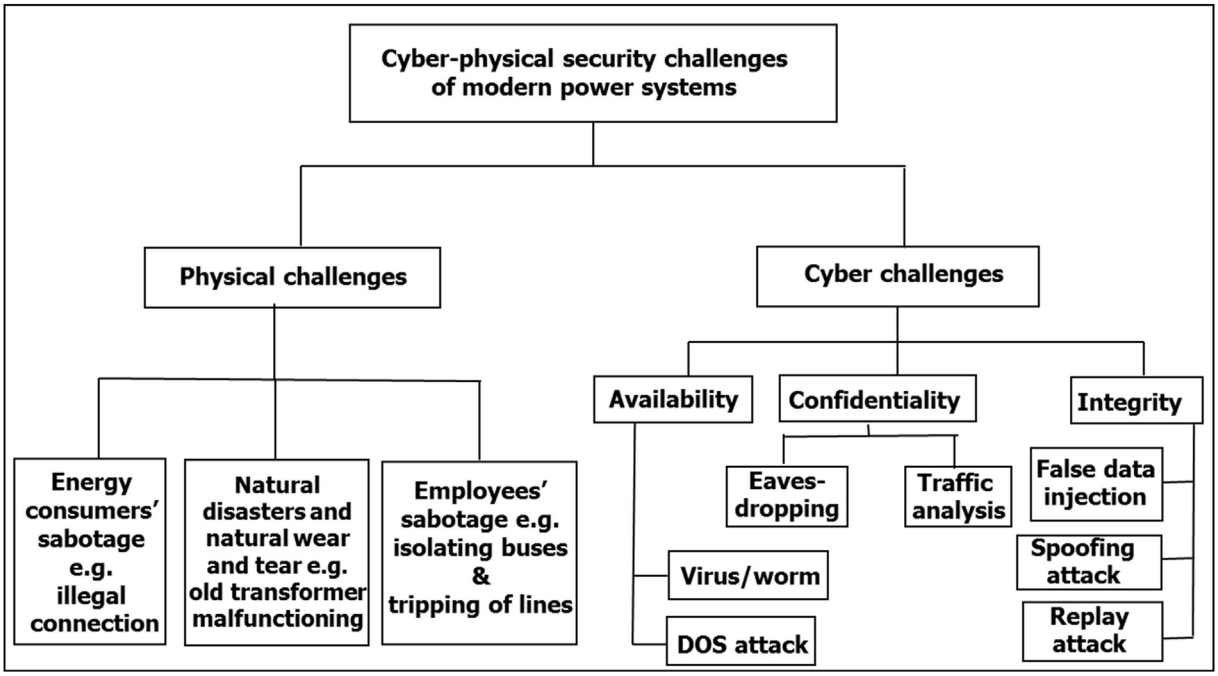

Figure 1. Security challenges of modern power system [1, 2].

However, it is noteworthy that various sophisticated infrastructures including the PMUs are vulnerable. Various experiences and studies in the literature have discussed PMU technology vulnerabilities to numerous intrusions and attacks [10, 11, 12, 13, 14]. Pan *et al.* [10] explained that a sophisticated cyberattack in form of a false measurement injection can mimic a fault thereby triggering a relay trip. Such imitation may not be detected by PMU technology. As intrusions and attacks have become rather inevitable, it is highly important that effective models for identifying and classifying undesirable events into power system are available, so that proper control commands and mitigation responses can be instigated promptly.

In recent years, numerous mathematical/statistical theories and models such as graph theory [15, 16], game theory [17, 18, 19, 20], traditional machine learning (ML) models [7, 21, 22], etc. have been proposed for power system security analysis. However, the proposed methods have shown various forms of shortcomings and poor performances. The poor performances are closely tied to the deployed data preprocessing techniques which include the deployed feature selection and weighting techniques for the large dimension synchrophasor datasets, poor parameter selections for the classifier(s), etc. In some of the existing works, human experiences were used for data preprocessing. Furthermore, several power system security works deployed conventional techniques such as InfoGain [22], relief algorithm [23], correlation-based feature selection, sequential backward selection [24], etc., as preprocessing techniques. However, the results achieved indicated that the proposed techniques offered less classification performance compared to the technique proposed in this study.

In this work, the hybridization of genetic algorithm (GA) and support vector machine (SVM) for the detection and classification of unwanted events into a power system network is presented. The weighted and dominant selected power system features by the GA are utilized to train a linear SVM and radial basis kernel (RBF) SVM model in accurately classifying the malicious control commands from normal one. To validate the efficacy of the developed GA-SVM model, experimentations were done using two power system datasets and the results were compared with various other models.

Specifically, the main contributions of the paper are summarized as:

- Develop an effective classification model that is capable of distinguishing power system events and efficiently classifying the events into their appropriate categories.
- To improve the generalization ability and classification performance, the features of large dimensional power system datasets are initially reduced using principal component analysis (PCA) before they are varied in their relevance and the dominant weights are heuristically identified using a GA.

The remainder of the paper is organized as follows. Section 2 presents related research works. Section 3 presents the methodology, and the simulation results and discussions are presented in Section 4. The conclusions are presented in Section 5.

2. Related works

In the literature, various ML algorithms have been modelled for power system events' detection and classification. Alimi *et al.* [3] presented a review of ML approaches to power systems security solutions. Among the foremost deployed ML tools, SVM has continued to be a dominant model as it gives excellent classification performances [7]. The authors in [21] presented the use of SVM for power system security evaluation. Similarly, Binna *et al.* [25] explored the use of SVM and recurrent neural network for classifying attacks on a power system test system. For performance improvement, Chen *et al.* [26] proposed the use of various ensemble learners as power system's attack detectors. In a similar study, Alimi *et al.* [7] modelled the ensemble of SVM and multilayer perceptron neural network (MLPNN) for cyber-attacks detection on a load flow analysis result of a 24-bus test system. However, as conventional ML algorithms are highly susceptible to errors and misclassifications owing to ineffective parameter selections, feature weighting techniques and complicated design procedures, the proposed models have shown various forms of shortcomings. To this end, various feature selection, feature weighting and optimization techniques have been deployed for ML performance optimization, feature selection and weighting procedures [3]. Ullah *et al.* [22] proposed InfoGain as filter-based selection method for a Bayes Net and J48 power system classification. Ali *et al.* [27] and Shang *et al.* [28] proposed the use of

Particle swarm optimization (PSO) to optimize Extreme Learning Machine classifier and one class SVM respectively. In related work, Abdoos *et al.* [24] deployed Gram–Schmidt orthogonalization (GSO) as feature selection and optimizer for SVM in classifying power quality events. However, some heuristic algorithms have disadvantages when deployed as optimization tools. While GSO are popular with their numerical instability with regards to rounding error, PSO are adjudged to easily fall into local optimum in high dimensional space [3, 29]. Thus, the techniques are not as effective in comparison to the GA that is deployed as feature weighting technique in this paper.

Inspired by the widespread acceptance of GA as a powerful optimization and/or feature enhancing tool that is based on the principle of Darwin's theory of evolution, an attempt has been made in this paper to implement a GA model as the feature weighting tool for SVM model for power system security assessment. The results achieved from the proposed GA-RBF SVM model was compared with those obtained using ordinary and GA-feature weighted linear SVM, Random Forest (RF), MLPNN and existing models in the literature. The proposed GA-SVM model outperformed the results achieved from the other models.

3. Methodology

This section presents the detailed description of the power system architecture, the testbed and the dataset used in the study. Also, the section presents a comprehensive description of the proposed classification model for accurately classifying power system events.

3.1. Power system architectural framework and dataset description

The power system architecture described in [30] was deployed in the study. The architecture presents a typical non-pilot directional over current relay protection scheme. A detailed description of the power system architecture can be accessed in [9, 30]. The framework is a fully operational scaled down 3-bus system that is integrated with four circuit breakers, tagged CB1, CB2, CB3 and CB4, that are controlled by relays. Figure 2 presents the one-line diagram of the 3-bus system.

As shown in Figure 2, Beaver *et al.* [30] explained that via transmission line 1 and line 2, the load is powered by Generators- Gen 1 and Gen 2. The Relays (1, 2, 3 and 4) with incorporated PMU functionality reside at each end of the lines that control the four breakers. As expected, Relay 1 and Relay 2 offer instantaneous over current protection for line 1. Relay 1 provides the same services for line 2 if Relay 3 fails in protecting line 2 against faults. Similarly, at line 2, Relay 3 and Relay 4 protects line 2 against faults and Relay 4 offers line 1 standby if Relay 2 failed a tripping test. All the relays constantly monitor and send the time synchronized data to the control centre via the phasor data concentrator (PDC).

To generate the datasets used in this study, Beaver *et al.* [30] simulated a testbed to implement the power system framework depicted in Figure 2. The testbed includes a real-time power system simulator, making use of commercial PMUs, PDCs and relays. Each relay logs data which includes phase voltage, current phasor, apparent line impedance, etc. Furthermore, breaker events log, Snort alerts log, and control panel alerts logs were captured. In the constructed simulation testbed, a threat model consisting of thirty-seven sets of five major event scenarios was designed to collect the balanced experimental datasets. The 5 major event scenarios include data injection attack, short-circuit fault, line maintenance, remote tripping command injection attack and relay setting attack. Each event aims to achieve some physical effect on the system.

From the simulation, three different classes of datasets were generated, which can be accessed via [31]. Two out of the three datasets were deployed in this study. A two-class data with the event scenarios categorized as either attack (28 events) or normal operation (9 events) while the second dataset deployed in the study is a three-class data with the 37 event scenarios distributed into natural events (8), no event (1) and attack events (28). The datasets generated were randomly sampled at 1%. Each of the two datasets deployed in this study is made up of 78,377 samples with 128 features.

3.2. GA-SVM model description

This subsection presents a detailed description of the proposed classification model which is based on using a GA as a feature weighting tool for the SVM model for accurately classifying power system events. The simplified flowchart of the proposed model is presented in Figure 3. The main steps of the model are the pre-processing steps, the GA-based feature weighting steps and the SVM training and evaluation steps.

3.2.1. The pre-processing steps

Generally, raw datasets require preparation into appropriate form for ease of classification. The pre-processing steps are important as they have significant influence on classification performance [3]. As we are working with voluminous datasets, we used PCA [32] for the reduction of the feature dimension of the two datasets. The PCA helps in decomposing the multivariate quantities into a set of orthogonal components that assist in improving training efficiency, prevention of overfitting and general enhancement of the classifier's accuracy. Each dataset is then partitioned at a ratio of 7:3 representing training and testing set respectively.

3.2.2. GA-based feature weighting steps

Current and future research in relevant fields including power system security studies, constantly work on voluminous data which poses a major challenge for conventional ML algorithms. The most common methods for efficient processing of voluminous datasets are feature weighting and feature selection techniques. While feature selection algorithms refer to algorithms that select the best subset which retain the interpretation of the large original data, feature weighting models operate based on the ideology that data features vary in their importance and each feature's contribution to the classification task should be different [33, 34]. Thus, higher weights are allocated to relevant features and lesser weights are allocated to the less relevant ones [35].

For a linear feature weighting approach for ML classifier's performance improvement, consider a training dataset a, which contains f features and h number of feature instances, defined in (1) [34]:

$$\{a_{i,h}, y_h | h \in H \text{ and } i \in f \} \tag{1}$$

where a_i is the i^{th} feature and y denotes the output labels. The linear feature weighted data $a_i^{'}$ can be defined as (2):

$$a_i^{'} = W_i . a_i \tag{2}$$

where W_i denotes the corresponding weight of the i^{th} feature. It alters the feature space of the classification task by escalating the space of highly weighted features while attenuating the space of less weighted features. The complication of feature weighting tasks relatively intensifies with respect to the volume of the data features.

Various algorithms including heuristic optimization techniques such as GA are widely used to process data feature weights [35, 36]. Inspired by Charles Darwin's theory of genetics, GA are powerful tools that are well known for their strong global search ability in finding solutions to non-deterministic polynomial-time hardness problems. They have been extensively combined with various ML models as a feature weighting tool and parameter optimization technique in several classification tasks [33, 35, 36]. The basic step of a GA process is stated briefly as follows:

1. Creating population chromosome.
2. Evaluating the fitness values of individuals in the population based on the deployed fitness function.
3. Applying the genetic operators (selection, crossover and mutation) to create new populations.

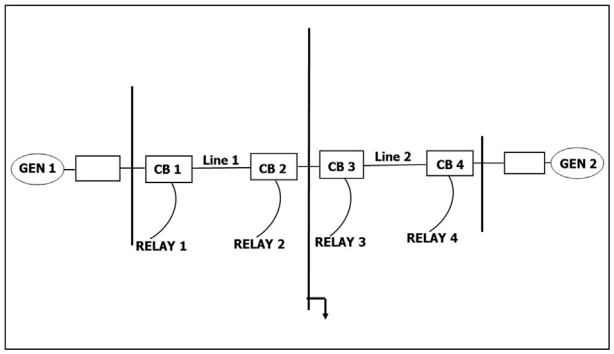

Figure 2. One-line diagram of the deployed test system [9, 30].

4. Step 2 and step 3 are continuously repeated until termination conditions are met. For more details on GA, the authors refer readers to [35, 36, 37].

For the process of using GA as the feature weighting method in our study, it is necessary to define the GA parameters which include the population size (μ), chromosome number (M), crossover rate (P_c), mutation rate (P_m), mutation step size (σ) and iteration number (I). In our work, the values of M and (μ) are considered from [38] as we are dealing with similar dataset structures. We randomly choose values between 0 and 1 for each weight value of the individual feature. For the fitness function, the fitness of each population chromosome is assessed using the quadratic loss function described in [39]. For the parent selection, we use size 2 tournament selection technique [35]. In size 2 tournament selection, 2 individuals are selected out of the population randomly and the better of the two individuals, (based on the fitness function ranking), is chosen as the next parent [35]. For the crossover, we used two-point crossover technique. In the scheme, two crossover points are picked randomly from the parent chromosomes. The bits in between the two points are swapped between the parent organisms to create new offspring. We deployed a simple inversion mutation [40] on each chromosome probabilistically, whereby a point is selected randomly during mutation and it is inverted afterwards. For instance, if the value is X before mutation, the value will be $(1 - X)$ post-mutation. The entire process is repeated until the number of iterations reaches the maximum 100 generations, or there is no improvement of the fitness value recorded. For the choice of the quantitative parameter values, some preliminary experiments were conducted using various parameter settings. During the experimentations, the parameter settings specified in Table 1 produced the best results in terms of performance and overall behaviour of the algorithm. The pseudocode of the GA feature weighting steps is presented in Algorithm 1.

Table 1. Qualitative and quantitative parameter settings for the modelled GA.

Parameter	Setting/Value
Chromosome Number (M)	50
Population size (μ)	100
Iteration number (I)	100
(P_m)	0.1
(P_c)	0.9
(σ)	1%
Parent selection	Tournament selection
Recombination	2-point crossover

3.2.3. SVM training and evaluation steps

A SVM is a unique ML tool that has been consistently deployed in a wide range of classification studies as they are known to produce high accuracy, strong generalization and high stability [41]. SVMs works by using an iterative training algorithm to find an optimal separating hyperplane, which separates sample instances into its categories [7]. For linearly inseparable classification tasks such as power system event classifications, the sample instances are mapped into a high dimensional feature space by means of a nonlinear transformation process.

Consider a SVM classification task having a training dataset D defined in (3) [35]:

$$D = \{(a_1, y_1), ((a_2, y_2), (a_3, y_3), ...(a_n, y_n)\}, \tag{3}$$

where n is defined as the features' number and $i = (1, 2, 3, ...n)$. For an input a_i, the corresponding label is y_i. Assuming the corresponding classes for individual instances are labelled as $y_i \in \{-1, 1\}$, all training samples are expected to meet the qualification equation given in (4) [42]:

$$y_i((\omega, a_i) + b) - 1 + \xi_i \geq 0 \tag{4}$$

where ξ_i is the positive slack variable, introduced to boost the model's fault tolerance and b is the bias. Maximizing the SVM hyperplane distance by using (4) is equivalent to solving the optimization problem in (5) subject to (6) [35,38]:

$$min \frac{1}{2} \| \omega \|^2 + C \sum_{i=1}^{n} \xi_i \tag{5}$$

$$y_i((\omega, a_i) + b) - 1 + \xi_i \geq 0 \tag{6}$$

where C denotes the penalty parameter that is used to control the trade-off between the slack variable penalty and the margin size. Usually, C is defined using the K-fold cross validation method. Similar to most ML algorithms, the classification effectiveness of SVM depends not only on the properties of the data, but also on selected feature processing tools and the appropriate parameter values such as C , the kernel function type, kernel parameter, etc. [7, 35, 42]. Typical Kernel functions $k(a_i, a_j)$ include radial basis function (RBF), linear and polynomial kernel functions which are defined in (7), 8 and (9) respectively [35, 38].

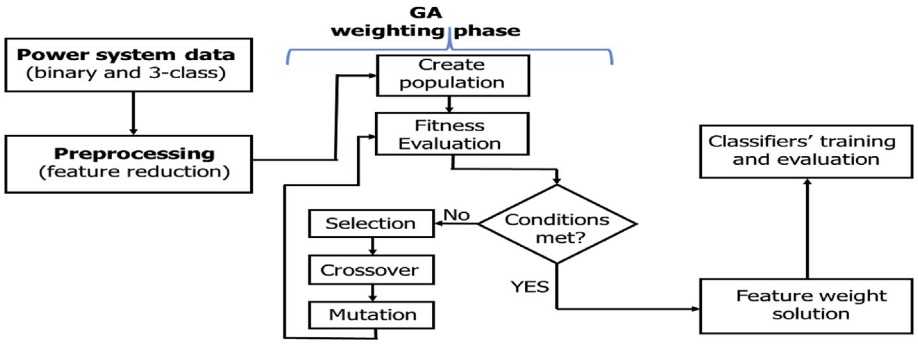

Figure 3. Simplified GA-SVM flowchart.

Algorithm 1
Pseudocode of GA for feature weighting

```
Input: Initial Chromosome Number (M),
       Iteration number (I),
       Size of Population (μ),
       Mutation rate (P_m), Crossover rate (P_c)
Output: Optimal weight solutions
Begin
Initialize the population chromosome with respect to the size (μ)
Evaluate fitness values and sort the population with respect to the fitness values
for (i = 1: I) do
    Select best individuals with respect to the rate (P_c) based on Tournament selection
    Generate new offspring by 2-point crossover
    Perform mutation with regards to the rate (P_m)
    Evaluate fitness of new offspring
    Resort the population
    Prune the worst population individuals
//loop to end once the two conditions are met
End
```

Table 2. SVM, MLPNN and RF parameter settings.

Algorithm	Parameter Value
Linear SVM	$C = 0.9$
RBF SVM	$C = 1.2, \gamma = 1.5$
MLPNN	ReLu activation function, 3 hidden layers of 30 neural nodes each, Solver = Adam [7]
RF	max_depth = 6, n_estimators = 10, max_feature = 1

$$k(a_i, a_j) = \exp\left(-\gamma \left\| a_i - a_j \right\|^2\right) \tag{7}$$

$$k(a_i, a_j) = a_i^T a_j \tag{8}$$

$$k(a_i, a_j) = (1 + a_i \,.\, a_j)^d \tag{9}$$

where $i, j = 1, 2, 3, \ldots\ldots, m.$, d is the degree of the polynomial kernel, $\left\| a_i - a_j \right\|$ is the Euclidean distance and γ is the RBF kernel function parameter. For more details on SVM, the authors refer readers to [42, 43, 44]. In this work, the SVM classifier model is trained using the training datasets with the feature weighted by the GA. We experimented using the linear SVM and RBF kernel SVM. For the modelled RBF and linear SVM, Table 2 presents the parameter settings used. The choice of these parameters is based on numerous initial experiments. During the initial experiments, the parameters were varied to achieve results with reduced generalization errors and overfitting problems. For comparison and validation, we developed a RF and a MLPNN model. The parameters used for the modelled RF is also chosen based on several initial trials. For the modelled MLPNN, the parameter settings are adapted from our previous work [7]. The parameters for the MLPNN and SVM are also presented in Table 2. Note that the parameters in Table 2 were used for both datasets.

- Accuracy: the accuracy can be described as the percentage of correct classifications in relation to the total classification choices [45]. Accuracy can be expressed mathematically as [7]:

$$\%Accuracy = \frac{total\ corrected\ classified\ sample}{total\ dataset\ samples} \tag{10}$$

- Precision: Precision refers to how frequently the ML algorithm is correct. Mathematically, precision can be expressed as [7]:

$$Precision = \frac{TP}{TP + FP} \tag{11}$$

where TP can be described as the number of correctly classified unwanted event instances and FP is the number of normal events instances that are wrongly classified as unwanted events.

- Recall: Recall measures the true positive rates [45]. Recall can be expressed mathematically as [7]:

$$Recall = \frac{TP}{TP + FN} \tag{12}$$

where FN can be described as the rate of false negative observations.

- F-Measure: it is described as the harmonic average of Precision and Recall [7].
- Specificity: Specificity is described as the ratio of actual negatives instances that were predicted as true negative instances.

4. Simulation results and discussions

The performance of the proposed GA-SVM model for classifying the deployed power system datasets were evaluated and compared with those obtained using RF, MLPNN, ordinary linear, RBF SVM and existing models in the literature.

4.1. Evaluation of binary class dataset results

Using the binary class dataset, Table 3 presents the classification results for the RF, MLPNN, linear and RBF kernel SVM with and without the use of the GA feature weighting process. As shown in Table 3, the RBF kernel SVM performed better, compared to the linear SVM and other

Table 3. Comparison of binary class data results using SVM, MLPNN and RF models with and without GA feature weighting.

Classifiers	Accuracy	Precision	Recall	F-Measure
Linear SVM	76.2%	76.2%	75.4%	73.3%
RBF SVM	81.0%	80.1%	82.5%	76.0%
MLPNN	78.8%	78.5%	80.2%	75.3%
RF	81.9%	82.6%	83.9%	77.9%
GA-Linear SVM	87.0%	85.2%	86.6%	80.1%
GA-RBF SVM	91.9%	93.7%	95.0%	87.0%
GA-MLPNN	86.4%	87.2%	85.7%	84.9%
GA-RF	88.2%	87.4%	89.1%	86.1%

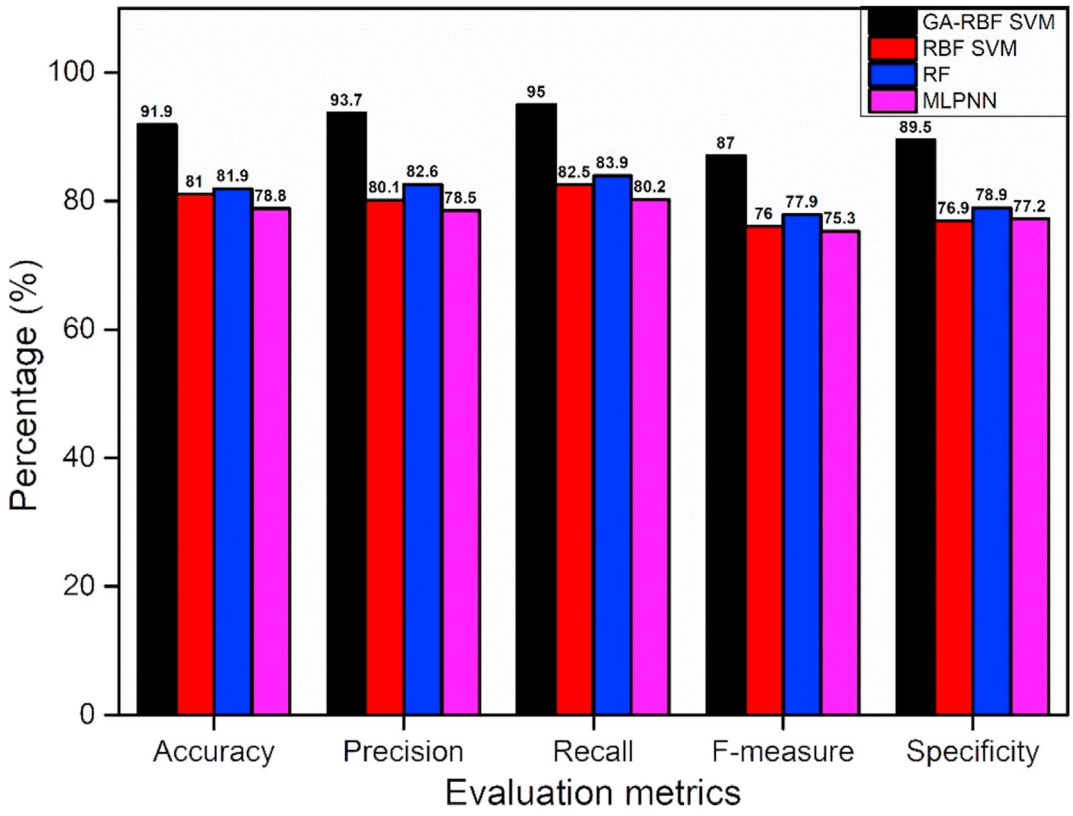

Figure 4. Comparative result of the developed models using binary class dataset.

ordinary conventional models. Unlike linear SVM, the ability of RBF kernel SVM to create nonlinear combinations of features to elevate samples onto an upper dimensional feature space can be attributed to the better performance achieved.

As shown in Table 3, using GA as the feature weighting tool for the binary class data contributed to the performance enhancement of the classification results. Using the GA as the feature weighting tool, the accuracy of MLPNN, RF, Linear SVM and RBF kernel SVM models in correctly classifying the unwanted power system events increased significantly from 78.8%, 81.9%, 76.2% and 81.0%–86.4%, 88.2%, 87.0% and 91.9% respectively. Focusing on key classification errors, Figure 4 presents a presentation of the comparison result of the GA-RBF SVM model with the result of RBF kernel SVM, MLPNN and RF in classifying the 37 event scenarios in the binary class dataset. As depicted in Figure 4, the GA-RBF SVM presented the best result in terms of true negative rates, recall, etc. This is due to the fact that GA-based feature weighting method is capable of effectively choosing the optimal weights of features used for enhancing power system events classification.

Furthermore, using the binary dataset, Table 4 presents the comparison results achieved from the developed model with other related works in the literature. As shown in Table 4, the specificity and recall result achieved from the proposed model is superior to the result achieved using the AdaBoost + JRipper model proposed by Hink *et al.* [45]. The relatively high precision achieved using the AdaBoost + JRipper model by Hink *et al.* [45] can be attributed to the fact that AdaBoost is a meta-algorithm that uses an iterative approach to learn from the mistakes of weak classifiers, and turn them into strong ones [46, 47]. Similarly, Table 4 shows that the adaptive regularized cost-sensitive online gradient descent algorithm (ARCSOGD) model proposed by Li *et al.* [48] presented recall and specificity results which are remarkably lower than the one achieved using the proposed GA-SVM model. Using the same convention for the weighted sum calculation proposed by Li *et al.* [48], the comparison result in Table 4 shows that the weighted sum result from the proposed GA-SVM model result is better than the result achieved using the ARCSOGD model proposed by Li *et al.* [48]. For further validation of our GA feature weighting choice, we compare our algorithm with models that involved the use of other prominent heuristic optimization methods as feature weighting tool for the SVM classifier. As shown in Table 4, the precision, accuracy and recall results achieved from the proposed GA-SVM model are significantly superior to the results achieved using the PSO-SVM model proposed by Huang *et al.* [49] and the SVM-ant colony optimization (SVM-ACO) model proposed by Li *et al.* [50].

Table 4. Comparison of binary class data results with results achieved from existing models using the same binary class data.

Classifiers	Average Precision	Average Accuracy	Average Recall	Average Specificity	Weighted sum
GA-RBF SVM	93.7%	91.9%	95.0%	89.5%	92.3%
JRipper [45]	85.0%	-	70.0%	90.0%	-
AdaBoost + JRipper [45]	94.0%	-	89.0%	78.0%	-
ARCSOGD [48]	-		45.5%	62.9%	54.2%
PSO-SVM [49]	90.2%	89.5%	80.7%	-	-
SVM-ACO [50]	86.0%	84.4%	84.9%	-	-
CFS-RF [51]	96.4%	-	-	-	-

Table 5. Comparison of three-class data results using SVM, MLPNN and RF models with and without GA feature weighting.

Classifiers	Accuracy	Precision	Recall	F-Measure
Linear SVM	76.8%	76.4%	75.6%	72.3%
RBF SVM	80.9%	80.1%	82.5%	75.8%
MLPNN	72.4%	71.9%	71.1%	70.3%
RF	75.2%	75.9%	77.1%	70.9%
GA-Linear SVM	85.7%	84.4%	83.1%	81.7%
GA-RBF SVM	89.8%	90.9%	91.2%	85.9%
GA-MLPNN	80.5%	79.8%	79.2%	78.4%
GA-RF	83.9%	82.7%	84.6%	82.7%

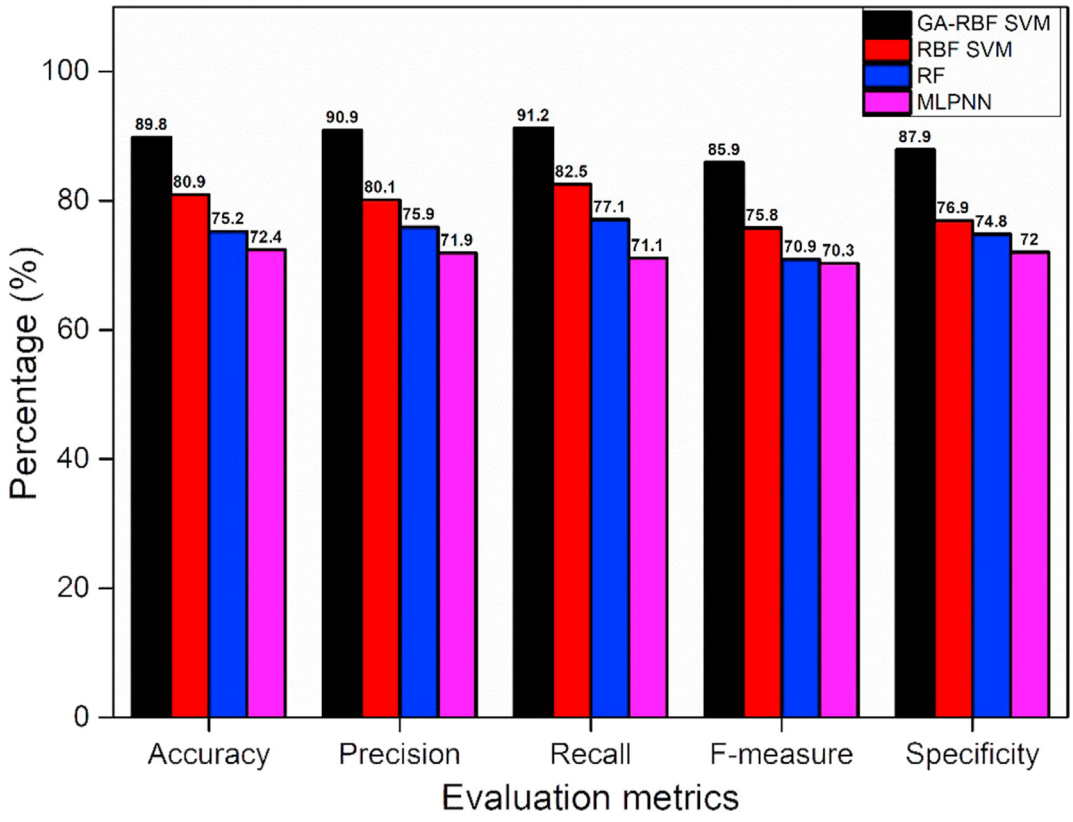

Figure 5. Comparative result of the developed models using three-class dataset.

4.2. Evaluation of three-class dataset results

Using the three-class dataset, Table 5 presents the classification results for the RF, MLPNN, linear and RBF kernel SVM with and without using the GA feature weighting process. Noticeably, using GA as a feature weighting tool boosted the classifiers' performances. The accuracy of the developed RBF kernel SVM, linear SVM, MLPNN and RF models in correctly classifying the unwanted power system events increased significantly from 80.9%, 76.8%, 72.4% and 75.2%–89.8%, 85.7%, 80.5% and 83.9% respectively. Figure 5 presents the comparison result of the GA-RBF SVM model with the result of RBF SVM, MLPNN and RF for the three-class data. As shown in the comparison results in Figure 5, the GA-RBF kernel SVM presented the best classification result.

Table 6 presents the comparison results of the GA-SVM model with the results from related works in the literature. The 90.9% accuracy achieved from the proposed GA-SVM model is slightly superior to the

Table 6. Comparison of three-class data results with results achieved from existing models using the same three-class data.

Classifiers	Average Precision	Average Accuracy	Average Recall	Average Specificity	Weighted sum
GA-RBF SVM	89.8%	90.9%	91.3%	85.9%	92.3%
AdaBoost + JRipper [45]	95.0%	99.0%	100%	95.5%	-
CPM [9]	-	90.4%	-	-	-
PSO-SVM [49]	86.5%	85.7%	83.1%	-	-
SVM-ACO [50]	80.9%	78.0%	77.4%	-	-
ARCSMC [48]	-	-	89.4%	78.4%	83.9%

achieved accuracy using the common path mining (CPM) model proposed by Pan *et al.* [9] and the sequential pattern mining (SPM) model proposed by the authors in [10]. Similarly, the weighted sum results as well as the classification error results in terms of recall and specificity achieved using the proposed GA-SVM model is significantly higher than the results achieved using the adaptive regularized cost-sensitive multi-class online learning (ARCSMC) model by Li *et al.* [48].

For further validation of the GA feature weighting tool we used, Table 6 also presents the comparison results of the proposed GA-SVM model with PSO-SVM model proposed by Huang *et al.* [49] and SVM-ACO model proposed by Li *et al.* [50]. As shown in Table 6, the PSO-SVM [49] performed better than the SVM-ACO [50] method. However, the GA-RBF SVM outperformed the other methods except the AdaBoost + JRipper ensemble [45]. The performance of the AdaBoost + JRipper model [45] can be attributed to the fact that AdaBoost are highly efficient if the weak learner is implemented efficiently [52].

5. Conclusions

In this paper, a hybrid approach for the classification of power system unwanted events based on using a GA as the feature weighting tool for the SVM classifier is presented. In the proposed GA-SVM approach, two voluminous synchrophasor datasets are used as the experimental basis, They were initially reduced using PCA before they are weighted in their relevance and the dominant weights are heuristically identified using the GA to boost classification results. The weighted power system features were used to train the developed SVM classifier model. Key ML performance metrics which include accuracy, precision, recall, F-Measure and specificity are used to evaluate the proposed GA-SVM model and other models considered in the experiment. Results show that using a GA as the feature weighting tool for the SVM modelled contributed to the performance enhancement of the classification results on the binary class dataset and the three-class dataset experimented on. Thus, it is evident that the proposed GA-SVM model can be used for power system security assessment in real world scenarios. Although the proposed GA-SVM approach accomplished outstanding power system classification results, there are some limitations that future works can consider. Future work can consider the application of the GA-SVM model in larger power systems with several other practical event scenarios, disturbances and intrusions. Also, faster feature weighting techniques can be considered in future research.

Declarations

Author contribution statement

Oyeniyi Akeem Alimi, Khmaies Ouahada & Adnan M. Abu-Mahfouz: Conceived and designed the experiments; Performed the experiments; Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper.

Suvendi Rimer: Conceived and designed the experiments; Performed the experiments.

Funding statement

This work was supported by the Council for Scientific and Industrial Research, Pretoria, South Africa, through the Smart Networks collaboration initiative and IoT-Factory Programme (Funded by the Department of Science and Innovation (DSI), South Africa).

Data availability statement

Data included in article/supp. material/referenced in article.

Declaration of interests statement

The authors declare no conflict of interest.

Additional information

No additional information is available for this paper.

Acknowledgements

The authors are highly grateful for the insightful and encouraging comments from the editor and reviewers that have helped to improve the quality of this paper.

References

[1] O.A. Alimi, K. Ouahada, Security assessment of the smart grid: a review focusing on the NAN architecture, in: IEEE International Conference on Adaptive Science & Technology (ICAST), 2018, pp. 1–8.
[2] Y. Xiang, L. Wang, N. Liu, Coordinated attacks on electric power systems in a cyber-physical environment, Elec. Power Syst. Res. 149 (2017) 156–168.
[3] O.A. Alimi, K. Ouahada, A.M. Abu-Mahfouz, A review of machine learning approaches to power system security and stability, IEEE Access 8 (2020) 113512–113531.
[4] C.C. Sun, A. Hahn, C.C. Liu, Cyber security of a power grid: state-of-the-art, Int. J. Electr. Power Energy Syst. 99 (2018) 45–56.
[5] G. Liang, J. Zhao, F. Luo, S.R. Weller, Z.Y. Dong, A review of false data injection attacks against modern power systems, IEEE Transact. Smart Grid 8 (2016) 1630–1638.
[6] R. Deng, G. Xiao, R. Lu, H. Liang, A.V. Vasilakos, False data injection on state estimation in power systems—attacks, impacts, and defense: a survey, IEEE Transact. Indust. Informat. 13 (2016) 411–423.
[7] O.A. Alimi, K. Ouahada, A.M. Abu-Mahfouz, Real time security assessment of the power system using a hybrid support vector machine and multilayer perceptron neural network algorithms, Sustainability 11 (2019) 3586.
[8] J. Ma, Y. Makarov, Z. Dong, Phasor Measurement Unit and its Application in Modern Power Systems, Emerging Techniques in Power System Analysis, Springer, Berlin, Heidelberg, 2010, pp. 147–184.
[9] S. Pan, T. Morris, U. Adhikari, Developing a hybrid intrusion detection system using data mining for power systems, IEEE Transact. Smart Grid 6 (2015) 3104–3113.
[10] S. Pan, T. Morris, U. Adhikari, Classification of disturbances and cyber-attacks in power systems using heterogeneous time-synchronized data, IEEE Transact. Indust. Informat. 11 (2015) 650–662.
[11] D.P. Shepard, T.E. Humphreys, A.A. Fansler, Evaluation of the vulnerability of phasor measurement units to GPS spoofing attacks, Int. J. Crit. Infrastruct. Protect. 5 (2012) 146–153.
[12] C. Beasley, Electric Power Synchrophasor Network Cyber Security Vulnerabilities, 2014.
[13] T. Morris, S. Pan, J. Lewis, J. Moorhead, N. Younan, R. King, M. Freund, V. Madani, Cybersecurity risk testing of substation phasor measurement units and phasor data concentrators, in: Proceedings of the Seventh Annual Workshop on Cyber Security and Information Intelligence Research, 2011, p. 1.
[14] J. Zhang, Z. Chu, L. Sankar, O. Kosut, False data injection attacks on phasor measurements that bypass low-rank decomposition, in: IEEE International Conference on Smart Grid Communications (SmartGridComm), 2017, pp. 96–101.
[15] S. Bi, Y.J. Zhang, Graphical methods for defense against false-data injection attacks on power system state estimation, IEEE Transact. Smart Grid 5 (2014) 1216–1227.
[16] N. Xie, F. Torelli, E. Bompard, A. Vaccaro, A graph theory-based methodology for optimal PMUs placement and multi-area power system state estimation, Elec. Power Syst. Res. 119 (2015) 25–33.
[17] F. Wu, Z. Hu, K. Chen, Game theory based power system security analysis, in: International Conference on Control, Automation and Systems Engineering (CASE), 2011, pp. 1–3.
[18] S. Backhaus, R. Bent, J. Bono, R. Lee, B. Tracey, D. Wolpert, D. Xie, Y. Yildiz, Cyber-physical security: a game theory model of humans interacting over control systems, IEEE Transact. Smart Grid 4 (2013) 2320–2327.
[19] B. Gao, L. Shi, Modeling an attack-mitigation dynamic game-theoretic scheme for security vulnerability analysis in a cyber-physical power system, IEEE Access 8 (2020) 30322–30331.
[20] M. Ni, A.K. Srivastava, R. Bo, J. Yan, Design of a game theory based defense system for power system cyber security, in: IEEE 7th Annual International Conference on CYBER Technology in Automation, Control, and Intelligent Systems, CYBER), 2017, pp. 1049–1054.
[21] L.A. Maglaras, J. Jiang, T.J. Cruz, Combining ensemble methods and social network metrics for improving accuracy of OCSVM on intrusion detection in SCADA systems, J. Informat. Sec. Appl. 30 (2016) 15–26.
[22] I. Ullah, Q.H. Mahmoud, A hybrid model for anomaly-based intrusion detection in SCADA networks, in: IEEE International Conference on Big Data (Big Data), 2017, pp. 2160–2167.

[23] Y. Xu, R. Zhang, J. Zhao, Z.Y. Dong, D. Wang, H. Yang, K.P. Wong, Assessing short-term voltage stability of electric power systems by a hierarchical intelligent system, IEEE Transact. Neural Networks Learn. Syst. 27 (2015) 1686–1696.
[24] A.A. Abdoos, P.K. Mianaei, M.R. Ghadikolaei, Combined VMD-SVM based feature selection method for classification of power quality events, Appl. Soft Comput. 38 (2016) 637–646.
[25] S. Binna, S.R. Kuppannagari, D. Engel, V.K. Prasanna, Subset level detection of false data injection attacks in smart grids, in: IEEE Conference on Technologies for Sustainability (SusTech), 2018, pp. 1–7.
[26] X. Chen, L. Zhang, Y. Liu, C. Tang, Ensemble learning methods for power system cyber-attack detection, in: IEEE 3rd International Conference on Cloud Computing and Big Data Analysis, ICCCBDA), 2018, pp. 613–616.
[27] M.H. Ali, M. Fadlizolkipi, A. Firdaus, N.Z. Khidzir, A hybrid particle swarm optimization-extreme learning machine approach for intrusion detection system, in: IEEE Student Conference on Research and Development (SCOReD), 2018, pp. 1–4.
[28] W. Shang, P. Zeng, M. Wan, L. Li, P. An, Intrusion detection algorithm based on OCSVM in industrial control system, Secur. Commun. Network. 9 (2016) 1040–1049.
[29] M. Li, W. Du, F. Nian, An adaptive particle swarm optimization algorithm based on directed weighted complex network, Math. Probl Eng. (2014).
[30] J.M. Beaver, R.C.B. Hink, M.A. Buckner, An evaluation of machine learning methods to detect malicious SCADA communications, in: 12th International Conference on Machine Learning and Applications, 2013, pp. 54–59.
[31] U. Adhikari, S. Pan, T. Morris, R. Borges, Industrial Control System (ICS) cyber attack datasets, J. Beaver (2014). https://sites.google.com/a/uah.edu/tommy-morris-uah/ics-data-sets. (Accessed 16 March 2020).
[32] S.T. Roweis, EM algorithms for PCA and SPCA, Adv. Neural Inf. Process. Syst. 10 (1997) 626–632.
[33] W. Ali, A.A. Ahmed, Hybrid intelligent phishing website prediction using deep neural networks with genetic algorithm-based feature selection and weighting, IET Inf. Secur. 13 (2019) 659–669.
[34] S. Dalwinder, S. Birmohan, K. Manpreet, Simultaneous feature weighting and parameter determination of Neural Networks using Ant Lion Optimization for the classification of breast cancer, Biocyb. Biomed. Eng. 40 (2020) 337–351.
[35] A.V. Phan, M. Le Nguyen, L.T. Bui, Feature weighting and SVM parameters optimization based on genetic algorithms for classification problems, Appl. Intell. 46 (2017) 455–469.
[36] E. Sevinc, A novel evolutionary algorithm for data classification problem with extreme learning machines, IEEE Access 7 (2019) 122419–122427.
[37] L.L. Lai, Intelligent System Applications in Power Engineering: Evolutionary Programming and Neural Networks, John Wiley & Sons, Inc., 1998.
[38] K. Sajan, V. Kumar, B. Tyagi, Genetic algorithm-based support vector machine for on-line voltage stability monitoring, Int. J. Electr. Power Energy Syst. 73 (2015) 200–208.
[39] Z. Qu, S. Yuan, R. Chi, L. Chang, L. Zhao, Genetic optimization method of pantograph and catenary comprehensive monitor status prediction model based on adadelta deep neural network, IEEE Access 7 (2019) 23210–23221.
[40] S.S. Ray, S. Misra, Genetic algorithm for assigning weights to gene expressions using functional annotations, Comput. Biol. Med. 104 (2019) 149–162.
[41] M. Hao, Y. Li, Y. Wang, S. Zhang, Prediction of P2Y12 antagonists using a novel genetic algorithm-support vector machine coupled approach, Anal. Chim. Acta 690 (2011) 53–63.
[42] H. Xue, Y. Bai, H. Hu, T. Xu, H. Liang, A novel hybrid model based on TVIW-PSO-GSA algorithm and support vector machine for classification problems, IEEE Access 7 (2019) 27789–27801.
[43] V. Kecman, Learning and Soft Computing: Support Vector Machines, Neural Networks, and Fuzzy Logic Models, MIT press, 2001.
[44] A. Dhandhia, V. Pandya, P. Bhatt, Multi-class support vector machines for static security assessment of power system, Ain Shams Eng. J. 11 (2020) 57–65.
[45] R.C.B. Hink, J.M. Beaver, M.A. Buckner, T. Morris, U. Adhikari, S. Pan, Machine learning for power system disturbance and cyber-attack discrimination, in: 7th International Symposium on Resilient Control Systems, ISRCS), 2014, pp. 1–8.
[46] Y. Freund, R.E. Schapire, A decision-theoretic generalization of on-line learning and an application to boosting, J. Comput. Syst. Sci. 55 (1997) 119–139.
[47] R.E. Schapire, Explaining Adaboost, Empirical Inference, Springer, Berlin, Heidelberg, 2013, pp. 37–52.
[48] G. Li, Y. Shen, P. Zhao, X. Lu, J. Liu, Y. Liu, S.C. Hoi, Detecting cyberattacks in industrial control systems using online learning algorithms, Neurocomputing 364 (2019) 338–348.
[49] C.L. Huang, J.F. Dun, A distributed PSO–SVM hybrid system with feature selection and parameter optimization, Appl. Soft Comput. 8 (2008) 1381–1391.
[50] X. Li, X. Zhang, C. Li, L. Zhang, Rolling element bearing fault detection using support vector machine with improved ant colony optimization, Measurement 46 (2013) 2726–2734.
[51] J. Yeckle, S. Abdelwahed, An evaluation of selection method in the classification of scada datasets based on the characteristics of the data and priority of performance, in: Proceedings of the International Conference on Compute and Data Analysis, 2017, pp. 98–103.
[52] X. Wu, Boosting, Advanced Machine Learning, 2020. https://www.eecis.udel.edu/~xwu/class/ELEG867/Lecture8.pdf. (Accessed 7 August 2020).