Machine Learning Algorithms Against Hacking Attack and Detection Success Comparison

1st Levent YAVUZ
Electrical and Computer Engineering Department
Abdullah Gül University
Kayseri, Turkey
levent.yavuz@agu.edu.tr

2nd Ahmet SORAN
Electrical and Computer Engineering Department
Abdullah Gül University
Kayseri, Turkey
ahmet.soran@agu.edu.tr

3rd Ahmet ÖNEN
Electrical and Computer Engineering Department
Abdullah Gül University
Kayseri, Turkey
ahmet.onen@agu.edu.tr

4th SM MUYEEN
Electrical and Computer Engineering Department
Curtin University
Curtin, Australia
sm.muyeen@curtin.edu.au

***Abstract*—Power system protection units has got enormous importance with the growing risk of cyber-attacks. To create sustainable and well protected system, power system data must be healthy. For that purpose, many machine learning applications have been developed and used for bad data detection. However, each method has got different detection and application process. Methods has superiority over other methods. Although, an algorithm can detect some injections easily, same algorithm can be fail when injection type changed. So methods have got different success results when the injection types changed. For that reason, different injection types are applied on power system IEEE 14 bus system via created special hacking algorithm. PSCAD and python linkage has been used for simulation and detection parts. 3 different injection types created and applied on the system and five different most popular algorithms (SVM, k-NN, LDA, NB, LR) tested. Each algorithm's performances are compared and evaluated.**

Keywords— Bad data detection, hacking algorithm, knn, lda, lr, nb, svm

I. INTRODUCTION

The recent inovations of smart grid provide more powerfull, usefull, reliable and sustainable power flows. Transmission and distribution data have got great importance for this reliability. The data on power system seperate different types; circuit breakers and switch positions, power flow, bus voltage, tranmission lines currents and phase angle data. The communication is provided through these data. So, safe network communication is one of the most important part of power grids. For that purpose, every grid is in a special communication network with others and this communication network must be safe and durable against all different kind of hacking attacks. Therefore, the system operator must take measures all possible attacks. For that reason, there are many different machine learning (ML) protection applications have been developed.

Intelligent softwares are generated based on secured grid network and communication against external injections. To provide well-secured power system communication network different types of ML algorithms applied and tested on detection units. Support Vector Machine (SVM), k-Nearest Neighbour (k-NN), Logistic Regression (LR), Naive Bayes (NB) and Linear Discriminant Analysis (LDA) are the most popular ML algorithms which are applied on power system protection units[1]–[6]. Especially recent years, boosting algorithms are also included on detection and protection side. Boosting term refers to weak algorithm converter to strong algorithms, the method uses the way each learner train sequentially and finally obtain an improved new model. Even though this combination method has got many strong points, somehow a single ML algorithm can be more successful when compared boosting algorithm. The injection type seperation is the main reason of this conflict. Attackers develope different injection methods to find a gap of security units and crack the system[7], [8]. Because each ML method has got different mathematical background and different injection types can be overlooked by ML algorithms.

II. GENERATING A SPECIAL HACKING ALGORITHM

The power system faces many different hacking attack such as cracking/authentication failure, worm attack, hybrid attack, etc. The main reasons of these attacks are stealing information, crack the system, malfunctions on devices, false billing, energy theft and blackouts. These results are not only technical disasters but also causing power market negatively. So, improved protection unit must be aware of different kind of attacks and they can catch any kind of injections types even they have never faced[9], [10].

Therefore, there is a special hacking algorithm has been developed to obtain three different injection method and test the all different ML algorithms. The methods are; 1) Random Injection, 2) Continious Injection, 3) Slowly Data Manipulation.

The following sections explain detail of injection and manipulation methods.

A. Random Injection

The attackers can create unwanted manipulated data by using random injection method. The method works as heuristic approach. The hacking algorithm is created by using python coding. The generated hacking algorithm creates random points and try to leak the system by trial and error method.

978-1-7281-6611-7/20/$31.00 ©2020 IEEE

The following equations explain the mathematical methodology of the injection approach; **P**: Phase Angel, **V**: Voltage, **f**: Frequency, **D**: Data, **inf_D**: infected data, **μr**: Random Row, **μc**: Random Column. The total collected data amount is 380 so k starts from 0 to 380. Which means algorithm try to search and find a gap of them.

$$\sum_{k=0}^{380}(\pm D) * \%2 < \mathbf{inf\_D} < \sum_{k=0}^{380}(\pm D) * \%5$$

$$\mu_r\ \mu_c\ [D_{p,f,v}] * \%10 = \text{inf\_}D_{p,f,v}$$

$$\text{inf\_}D_{p,f,v} \rightarrow D_{p,f,v}\ \text{(injected random data)}$$

B. Continious Injection

Although the attack idea is simple, it is really harmful and very hard to identify by system provider due to its natural appearance. The manipulated data replaced with original data but the important part is that; at the beginning of attack there is not much difference between manipulated data and original data. For example, in this project simulation replaced data percentages are ±2% and ±5% of original data. So security units can miss out it. Because most of the anomaly detectors use a threshold, however this type of attack can leak the system. The main problem is this manipulation is continuous which means after first injection the second one comes with same method and go on. Until blackouts happen or occur some device malfunctions.

The following equations explain the mathematical methodology of the injection approach;

$$\sum_{k=0}^{380}(\pm D) * \%2 < \mathbf{inf\_D} < \sum_{k=0}^{380}(\pm D) * \%5$$

$$\sum_{k=0}^{380}(\pm D) = \text{inf\_}D_{p,f,v}$$

$$\text{inf\_}D_{p,f,v} \rightarrow \text{c\_inf\_}D_{p,f,v}\ \text{(Continue for 10 points)}$$

$$\text{c\_inf\_}D_{p,f,v} \rightarrow D_{p,f,v}\ \text{(random data injected)}$$

C. Slowly Data Manipulation

The most difficult atttack to detect by security units. Because of its nature data manipulation grows up in a slowly way. So manipulated data pretend as a part of the system. And it changes their value slowly. Thats why security units can not handle it easily. Unless the data manipulation is detected, operating system would be collapse. Following equations describe the attack type;

First Stage

$$\sum_{k=0}^{380}(\pm D) * \%0 < \mathbf{inf\_D} < \sum_{k=0}^{380}(\pm D) * \%2$$

$$\mu_r\ \mu_c\ [D_{p,f,v}]_{2\%} * \%10 = \text{inf\_}D_{p,f,v}$$

Second Stage

$$\sum_{k=0}^{380}(\pm D) * \%0 < \mathbf{inf\_D} < \sum_{k=0}^{380}(\pm D) * \%4$$

$$\mu_r\ \mu_c\ [D_{p,f,v}]_{4\%} * \%10 = \text{inf\_}D_{p,f,v}$$

III. The Machine Learning Algorithms

An injection attack can occur on power system at any time. The system manager must take precautions against that attacking attempts but lately hacking attacks become more harmful and powerful[11]. Also many different hacking attacks appear nowadays and unfortunately traditional protection units do not satisfy the needs. Therefore, ML algorithms help to protection units. The algorithms can detect many injections attacks in an effective way.

The most well-known and used ML algorithms are explained in this section. Also IEEE 14 bus system simulated on PSCAD and ML algorithms coded on Python to test injections. The algorithms mathematical background and individual simulation results are represented in this section.

A. K-Nearest Neighbour

The algorithm process is simple; a data is classified according to number of neighbors. This calculation is applied by using different coordinate systems such as; Minkowski space, Euclidian space and Manhattan space. To understand the progress of prediction, see fig. 1.

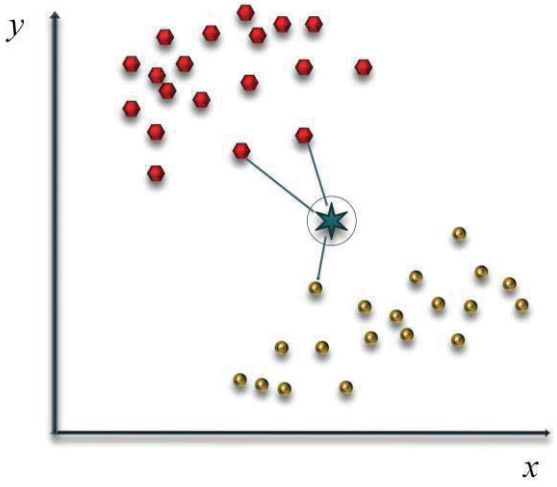

Fig.1 KNN algorithm's decision mechanism

All of the newcomers, and nodes are clustered by using space distances and finally clustering decision has been done via closest distance.

$$d(x,y) = \left(\sum_{i=1}^{n}|(x_i - y_i)^p|\right)^{1/p} \quad (1)$$

B. Support Vector Machine (SVM)

Support Vector Machine is a classification algorithm similar to Logistic Regression. Both try to find the best line that separates the two classes. Algorithm allows the line to be drawn to be adjusted in two classes to pass the elements from the farthest place. It is a nonparametric classifier that

takes no parameters. SVM can also classify linear and nonlinear data, but it usually tries to classify the data linearly [2].

The kernel space method stretches the surface and separate interwoven data by using some special mathematical methods, such as 'linear', 'poly', 'rbf', 'sigmoid' and 'precomputed'[12]. These methods are cutting the space when dataset are stretched enough. See fig 2.

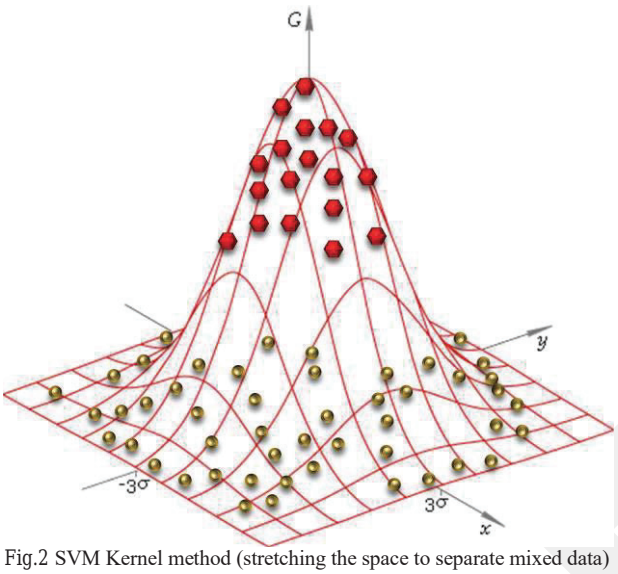

Fig.2 SVM Kernel method (stretching the space to separate mixed data)

C. Linear Discriminant Analysis

Linear Discriminant Analysis (LDA) is used as a size reduction technique in the preprocessing stage for ML applications. The aim is to prevent overfitting and at the same time reduce calculation costs. The idea is maximizing the distance between classes.

In summary, the purpose of LDA is to reduce the size of the data set by maximizing the difference between the classes.

D. Logistic Regression

Logistic Regression (LR) is a regression method for classifying operations. It is used for classification of categorical or numerical data. It works only if the dependent variable can only take 2 different values. (Yes / No, Male / Female, Fat / Weak etc.)

It is widely used in linear classification problems. For this reason, it is very similar to Linear Regression. The regression method uses a function which is called as sigmoid. This function fits all values between 0 and 1.

$$Sigmoid = \frac{1}{1 + e^{-x}}$$

E. Naïve Bayes

The idea is simple probabilistic classifier. It's called as conditional probability in computer and data science. It can work very well with little data and this is one of the most powerful side. Also its decision mechanism is very fast. Terminologically, Bayesian probability is given in (2).

$$p(C_k|x) = \frac{p(C_k)\, p(x|C_k)}{p(x)} \quad (2a)$$

$$posterior = \frac{prior\ x\ likelihood}{evidence} \quad (2b)$$

Where,

- $p(C_k|x)$ is the posterior probability of class
- $p(C_k)$ is the prior probability of class
- $p(x|C_k)$ is the likelihood
- P(x) is evidence

IV. Simulation Results And Comparison of ML Algorithms

After PSCAD simulation has been done and all data are collected, injection process has been starts on it. Hacking attack tries to injection 5 times on same dataset by using different time and data points by developed hacking mechanism. After each injection ML algorithms start to train and test their success on it.

Each simulation cycle obtained data exposed hacking algorithm and finally injected data are procured. After that process ML algorithms have been used for training (%80 of data), training process applied by using labelled data, and the rest of data (20%) is used for test. And finally the recall results are used for comparison with other ML algorithms. Recall results are more important than accuracy of algorithms. Although the system is under attack and there is leaking, it is very dangerous to make the opposite decision. That's why recall value has been chosen as comparison criteria.

A. Setting The Parameter

Tuning parameter are one of the most important effect on accuracy results. Because, as mentioned before each algorithm has got different mathematical background and some parameter must be tuned according to dataset. Wrong tuning decisions may cause failed decision results. For example, there is a one tuning parameter for kNN, k value, this parameter checks each new data point's neighbors.

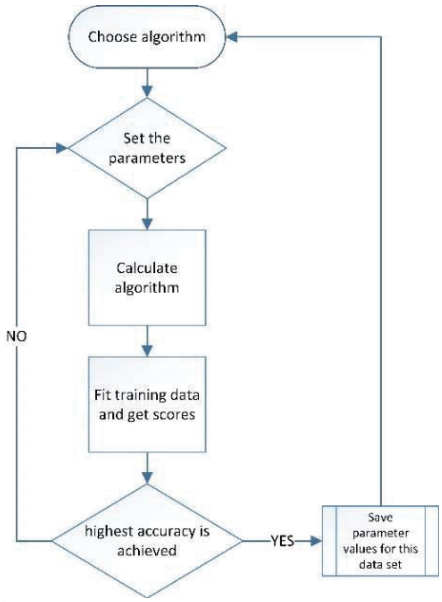

Fig.3 Algorithms parameter tuning process

For that purposes prepared and parameter tuning loop see Fig. 3. Each ML algorithm's parameters are calculated by this code cycle before test part. So the results are their individual highest results.

B. Simulation Model And Case Study

The simulation model is prepared on PSCAD by using IEEE14 bus test system. Voltage, frequency and phase data are collected from each buses. After collection of this clean and healthy data set, injection process has been started. 3 different type of injections are applied on same data set 5 different times and obtained manipulated data.

ML algorithms need to train before testing. So 80% of injected 5 different datasets have been used for training process. Rest of them 20% are used for training part.

In Fig. 4. There are some unintended data and ML algorithms(individually) try to detect them but as it can be clearly seen some of them fail in different cases. While Logistic Regression (LogReg), LDA and NB cannot define the injected data, k-Nearest Neighbor (kNN) and Support Vector Machine (SVM) can catch the point. Which means none of them can reach the perfect solution, alone.

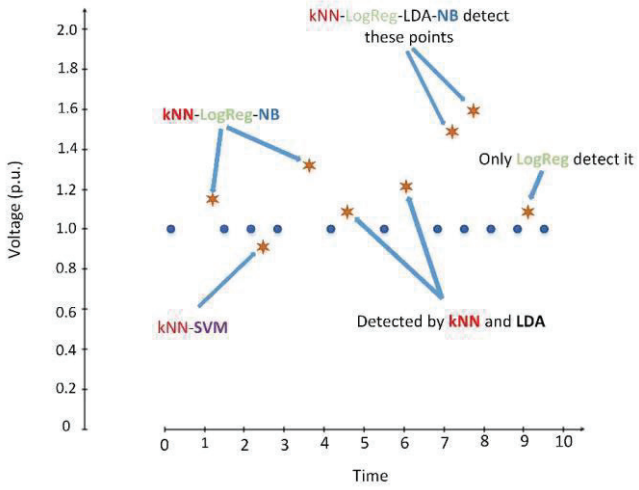

Fig.4 Comparison and evaluation results show that there is no single algorithm can detect all injections

V. Result And Discuss

In table 1. Algorithms individual recall successes are represented and their accuracy results can be different case by case. In random point injection case; for 5% of manipulated data is SVM but when manipulation rate is increased NB will be the winner. Continuous Injection attack results are so different then case 1 because kNN shows best result for 5% of manipulated data however this algorithm's success reduces when manipulated data increases up to 10% but for 15% of injected data kNN become winner.

The scenario is so different for case 3. LogReg shows almost the worst results for 2% of manipulation rate. But as the manipulation rate increases, LogReg become the best option for protection.

Table. 1. Simulation results for each manipulation attack

	Simulation Results (Case 1: Random Injection)		
% of infected data	**5%**	**10%**	**15%**
kNN	89.13	90.17	91.28
NB	90.3	93.31	94.66
LogReg	90.38	91.50	92.05
LDA	85.5	92.37	93.37
SVM	91.35	92.61	92.68
	Simulation Results (Case 2: Continuous Injection)		
% of infected data	**5%**	**10%**	**15%**
kNN	93,11	95,23	97,56
NB	89,23	93,95	92,9
LogReg	90,97	91,20	93,14
LDA	85,59	89,87	93,75
SVM	92,12	95,53	97,15
	Simulation Results (Case 3: Slowly Data Manipulation)		
% of infected data	**2%**	**4%**	**6%**
kNN	80,2	82,85	88,74
NB	31,54	38,60	48,37
LogReg	80,24	83,39	90,28
LDA	81,23	82,99	85,50
SVM	80,54	81,39	87,0

VI. Conclusion

The power system security units must have reliable, fast and unbreakable against all kind of hacking attack. The injection and data manipulation can cause huge damage on it. On the other hand, these units must be adaptive because the system can face different kind of attacking actions.

For that purpose, machine learning algorithms become so crucial. Because many different attacking types are now existing and protection units cannot handle all type of attacks. Some of them uses threshold detectors, some of them uses anomaly detectors via sensors and some protection devices. So ML algorithms must be used as decision maker. However, there are many different type of ML algorithms.

In that project, most popular ML algorithms used as decision maker. Results show that; none of them can handle all type of attacks, alone. While Naïve Bayes algorithm gives best results against random injection attack, k-

Nearest Neighbor algorithm become winner for slowly injection attack. The key point is ML algorithms mathematical background is not convenient against different type of attacks. So a ML algorithm cannot be sufficient alone.

REFERENCES

[1] J. Nordhaug Myhre, K. Øyvind Mikalsen, S. Løkse, And R. Jenssen, "Robust Clustering Using A Knn Mode Seeking Ensemble," *Pattern Recognit.*, Vol. 76, Pp. 491–505, 2018.

[2] M. Esmalifalak, S. Member, L. Liu, And S. Member, "Machine Learning In Smart Grid," Vol. 11, No. 3, Pp. 1644–1652, 2017.

[3] M. Esmalifalak, L. Liu, N. Nguyen, R. Zheng, And Z. Han, "Detecting Stealthy False Data Injection Using Machine Learning In Smart Grid," *Ieee Syst. J.*, Vol. 11, No. 3, Pp. 1644–1652, 2017.

[4] S. Institution Of Engineering And Technology., Y. Li, J. Xiang, And F. Ji, *Iet Power Electronics.*, Vol. 9, No. 10. Institution Of Engineering And Technology, 2016.

[5] L. D. Analysis, "Introduction To Lda Lda," *Cancer Lett.*, 2005.

[6] K. P. Murphy, "Naive Bayes Classifiers," *Bernoulli*, 2006.

[7] H. Yoo And T. Shon, "Novel Approach For Detecting Network Anomalies For Substation Automation Based On Iec 61850," *Multimed. Tools Appl.*, 2014.

[8] G. Kaestle, "Virtual Power Plants As Real Chp-Clusters: A New Approach To Coordinate The Feeding In The Low Voltage Grid."

[9] Y. Gu, T. Liu, D. Wang, X. Guan, And Z. Xu, "Bad Data Detection Method For Smart Grids Based On Distributed State Estimation," *Ieee Int. Conf. Commun.*, Pp. 4483–4487, 2013.

[10] A. S. Musleh, H. M. Khalid, S. M. Muyeen, And A. Al-Durra, "A Prediction Algorithm To Enhance Grid Resilience Toward Cyber Attacks In Wamcs Applications," *Ieee Syst. J.*, Vol. 13, No. 1, Pp. 710–719, 2019.

[11]A. Ashok, M. Govindarasu, And V. Ajjarapu, "Online Detection Of Stealthy False Data Injection Attacks In Power System State Estimation," *Ieee Trans. Smart Grid*, Vol. 9, No. 3, Pp. 1636–1646, 2018.

[12]A. S. Ahmad *Et Al.*, "A Review On Applications Of Ann And Svm For Building Electrical Energy Consumption Forecasting," *Renew. Sustain. Energy Rev.*, Vol. 33, Pp. 102–109, May 2014.