Journal of Green Energy Research and Innovation

Illegal Miner Detection based on Dynamic Mode Decomposition and Unsupervised Machine Learning Algorithms

Alireza Simorgh, Khosro Khandani, Maryam Amiri

Highlights

- **Employing a real data set from Markazi Province Distribution Company to detect illegal miners.**
- **Dynamic Mode Decomposition (DMD) has been used to extract new features.**
- **Obtaining high accuracy of classification using unsupervised algorithms.**

Graphical Abstract

Use your device to scan and read the article online

Citation

A. Simorgh, K. Khandani, and M. Amiri," Illegal Miner Detection based on Dynamic Mode Decomposition and Unsupervised Machine Learning Algorithms," *Journal of Green Energy Research and Innovation*, vol. 2, no. 6, pp. 79-88, 2025.

https://doi.org/10.61186/jgeri.2.2.79

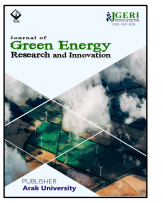
© Author

Journal of Green Energy Research and Innovation 2.2 (2025) 79-88

Online ISSN: 3041-9018

Journal of Green Energy Research and Innovation

Journal Homepage: www.jgeri.araku.ac.ir



Illegal Miner Detection based on Dynamic Mode Decomposition and Unsupervised Machine Learning Algorithms

Alireza Simorgh [1], Khosro Khandani [1,*], Maryam Amiri [2]

[1] *Department of Electrical Engineering, Faculty of Engineering, Arak University, Arak, 38156-8-8349, Iran.*

[2] *Department of Computer Engineering, Faculty of Engineering, Arak University, Arak, 38156-8-8349, Iran.*

A R T I C L E I N F O

Keywords:
Miner detection,
Unsupervised algorithms,
Dynamic mode decomposition.

Article History:
Received: 19 February 2025;
Revised: 24 March 2025;
Accepted: 25 March 2025.

Article type:
Research Article

* ***Corresponding author***
E-mail address
k-khandani@araku.ac.ir (K. Khandani)

A B S T R A C T

Since the most important issue in the production of digital currencies is energy consumption, the use of illegal electricity in mining farms has become very popular. Illegal mining is particularly important in countries such as Iran where the price of electrical energy is extremely low. This issue has caused numerous problems such as frequent blackouts, large losses for industries and even daily power cuts in several large cities. Previous machine learning approaches for miner detection are mostly supervised methods which rely on labeled data. Due to the fact that the number of labeled data is very limited in reality, we propose unsupervised methods in this paper. A real data set from Markazi Province Distribution Company in Iran has been employed to produce the results. The classification process consists of two stages: in the first stage, Dynamic Mode Decomposition (DMD) has been used to extract new features which compose the set of features along with certain factors from the Advanced Metering Infrastructure (AMI). These features are selected for 58 subscribers with positive and negative labels. In the second stage, a number of unsupervised models are built from the results of the first stage. The highest accuracy of classification obtained is 74% from unsupervised algorithms and 85% for supervised algorithms, which is very significant considering the fact that unsupervised algorithms do not need labeled data.

1. Introduction

The energy needed to produce one bitcoin is almost equal to the energy needed by a house for a week. According to the Cambridge Centre for Alternative Finance (CCAF), bitcoin currently consumes about 110 terawatt hours per year – 0.55% of global electricity production, or roughly equivalent to the annual energy harvest of small countries like Malaysia or Sweden [1]. Around 4.5% of all bitcoin mining worldwide took place in Iran between January and April 2021. The price of electricity in Iran is much lower than the average price for households and businesses. Considering the prohibition of the production of digital currencies in some countries such as Iran, Algeria and Vietnam, or legal laws in other countries such as Denmark, the way to identify illegal miners is very important. Today, electricity distribution companies use different approaches to identify illegal miners. Some companies use algorithms based on comparing customers' current energy consumption with the past. Today, various methods are used to identify illegal digital currency mining facilities in electricity distribution companies. One of the most common methods is called AlgExp, which collects the characteristics of consumers' electricity consumption based on their activities, and the collected data are compared with past consumption data [1]. Identifying anomalies in electricity consumption is another approach which was proposed in [2] where an unsupervised method which combines clustering-based and prediction-based methods, has been employed for anomaly detection in power consumption. The authors hypothesized that the same daily consumption behavior appears repeatedly. Based on this assumption, the K-means algorithm has been applied to investigate behavioral scenarios in 24 different groups per day. In [3] Long Short-Term Memory (LSTM) neural network has been used to identify anomalies in electricity consumption data one hour ahead of time. Forecasting energy consumption was also addressed in [4] where LSTM networks have been used to predict future energy consumption. Li et al. in [5] proposed a framework for power consumption anomaly detection based on industrial wireless sensor networks. Their mechanism is based on machine learning and blockchain where the collected data are classified into three categories: outlier class, working day class, and holiday class. According to the types of power consumption anomalies, the extreme class is also divided into three subclasses. If the received data are beyond the normal range, abnormality will be detected.

In the last step, the K-nearest neighbor algorithm was used to detect anomalies that were not detectable in the previous step. Ouyang et al. in [6] have proposed a multidimensional cumulative ensemble model to detect abnormal energy consumption using various IoT sensors in an industrial environment. The authors considered the energy consumption data as a time series and used the hierarchical time series feature extraction method to extract different types of features, including decomposed features, transformed features, summarized features, and transformed features. Also, some machine learning models, such as the ensemble learning method, have been trained to discover classes. Nerurkar et al. proposed a graph-based deep learning model using spectral graphs and transactional features [7]. The proposed model is trained on complex features to classify illegal transactions on Bitcoin. This approach is also very complicated due to the need for a large number of labeled data and numerous features. In [8], a detection method called Bagging-GCN has been proposed based on the combination of the Bagging algorithm and the Graph Convolutional Neural Network (GCN) algorithm. The experiments conducted on the elliptic dataset showed that Bagging-GCN could superiorly identify legal and illegal transactions compared to the traditional learning methods such as logistic regression and support vector machine. In[9], the method of combining layers in Convolutional Neural Networks (CNN) has been used. The paper proposes a model of Bitcoin transaction analysis to identify anomalies related to money laundering. In [10], a machine learning method was introduced to analyze the complete Bitcoin user graph in order to identify suspicious actors potentially involved in illegal activities. Contrary to existing studies, the proposed method in [10] introduced a new set of features which were used to identify potential criminal activities more accurately. Regarding miner detection approaches, Amiri et al. [11] have used the INBORN model, which is one of the pattern-finding methods, to identify unauthorized miners. This research has considered the correlation between different characteristics and has clearly extracted customer behavior patterns.

In most of the previous works, only supervised methods have been used, which has its own difficulty due to the complexities of producing labeled data for this task. In addition, most of the previous papers have been conducted only on the basis of network factors and micro-consumption of subscribers, and no new innovations have been included in this field. For this reason, the lack of an explicit and efficient model without the need for labeled data and using advanced methods of feature creation motivates us to propose an efficient method to solve this problem in this paper. On the other hand, although the introduced methods can identify criminal activities such as illegal mining or electricity theft, they misclassify many subscribers, as well. If the number of false negative labels is large, the subscribers with illegal miners will be mistakenly assigned a negative label. Large number of false positive labels means that, the subscribers who do not possess illegal miners are mistakenly assigned a positive label. Although both of the numbers of false negative and false positive are preferred to be small, any algorithm which results in a fewer false negatives is considered to be more effective. The purpose of this paper is proposing an efficient method to identify unauthorized miners. For this purpose, 58 subscribers from Markazi province distribution company in Iran have been selected. This labeled data set includes 41 subscribers with positive labels and 17 subscribers with negative labels. Some features are extracted from this data set and then Dynamic Mode Decomposition (DMD) is applied to extract some other features. In the next step, various pre-processing methods are applied to the data, and then, different unsupervised models are applied to them. Eventually the models are evaluated based on the accuracy of the miner detection. Figure 1 shows the structure of the proposed method. The innovation of this paper is extracting novel features by employing DMD which helps obtain more effective and reliable unsupervised models. The set of features are composed of DMD features and other features obtained from AMI. This new feature selection approach paves the way for constructing more effective models in similar classification problems. The rest of the paper is organized as follows: Basic concepts and definitions are explained in Section 2. Section 3 introduces the proposed method in detail. Then we evaluate the results in Section 4. Finally, this paper concludes with our future works in Section 5.

2. Basic concepts and definitions

This section introduces the main concepts and definitions including k-means, mean shift, DBSCAN, Agglomerative Clustering and dynamic mode decomposition.

2.1. K-means

The K-means clustering algorithm calculates the center points and repeats this until it finds the optimal center point. This algorithm needs to determine the number of clusters. The number of specified clusters is represented by k. In this algorithm, the data are assigned to a cluster in such a way that the sum of the squares of the distance between the data points and the central point is minimal. It is understandable that the smaller the changes in the clusters, the more similar data we will have within the cluster [12].

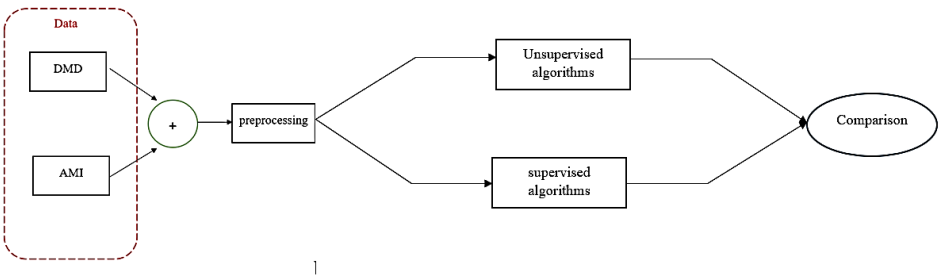


Figure 1. The procedure of the miner detection based on DMD and machine learning.

This algorithm continues to work so that the sum of squared differences from the mean is minimized for each cluster. The objective function of this algorithm is defined as follows. In this regard, μ_m is the center of cluster m, $c^{(1)}$ is the number of the cluster assigned to the data $x^{(i)}$ and $\mu_c(i)$ is the center of the cluster assigned to the data $x^{(i)}$.

$$J\left(c^{(1)}, c^{(2)}, \dots, c^{(k)}, \mu_1, \mu_2, \dots, \mu_m\right) = \frac{1}{2}\sum_{i=1}^{m}\left\|\mathrm{x}^{(i)} - \mu_c(i)\right\|^2 \tag{1}$$

As illustrated in the pseudo-code of Figure 2, first, the number of clusters is determined, then the initial centers of the clusters are determined, and finally, in each step, the distance between the calculated points and the centers of the clusters are updated until the objective function is minimized and there are no more data. In this research, since our goal is to identify miner subscribers from non-miners, the number of clusters is equal to 2. Also, the selection of the center of primary clusters is performed using K-means $++$ method.

2.2. DBSCAN

DBSCAN algorithm is the basis of density-based clustering methods, which has the ability to discover clusters of different sizes and shapes from a large volume of data and is also resistant to noise. Despite these advantages, this algorithm has problems such as the difficulty of determining the exact value of the input parameters, not recognizing clusters with different densities, and not correctly recognizing clusters when the clusters are close to each other [13]. In DBSCAN algorithm, the two main parameters are the minimum sample and the neighborhood radius, where the minimum sample is the minimum number of samples required to form a cluster, and the neighborhood radius is defined as the maximum distance between two points while they still belong to a cluster [12]. Any data point whose distance with an assumed data point is smaller than the radius is considered as a neighbor of that data point. The main idea of DBSCAN is that for each sample of a cluster, a neighborhood with a certain radius eps should contain a minimum number of objects. The neighborhood of a point like p is defined as in Equation (2):

$$N_{eps} = \left\{q \in \frac{D}{dist(p,q)} < eps\right\} \tag{2}$$

where D is the data set, dist(p,q) is the distance between points p and q, and eps is the neighborhood radius. If there is at least one predetermined number of points in the neighborhood radius of eps from a point like p, then this point is called the central point. The central point is defined as in Equation (3):

$$N_{eps}(p) > min\ sampels \tag{3}$$

The objective function in this algorithm is defined in Equation (4):

$$min\ W_1\left(\frac{1}{k}\sum_{i=1}^{k} s(i)\right) + W_2\left(\frac{k}{N_g}\right) \tag{4}$$

$$s.t: 0 < eps < D_{max}$$

$$2 < \min sampels < N_{max}\ ;\ \min sampels \in Z$$

where D_{Max} and N_{Max} are the Euclidean distance between the farthest point in the data set and the maximum number of data points respectively, k is the number of clusters and W_1 and W_2 are the weights to be determined.

2.3. Mean Shift

This method is a non-parametric approach with the aim of finding local maxima or clustering data. In using this method, the number of clusters is not necessary to be determined. It also has no restrictions on the shape of the data distribution or the shape of the generated clusters. This method finds local maxima with an iterative method, in such a way that for each data point in the feature space, the center of a search window is placed at that point, then it is checked whether that data point is located in a local maximum in terms of statistical distribution. If it is not located in a local maximum, the neighborhood center is moved and checked again.

input: K, number of clusters; D, a data set of N points
1. Initialization.
2. **Repeat**
3. **For** each point p in D **do**
4. Find the nearest center and assign p to the
5. Corresponding cluster.
6. **End for**
7. Update clusters by calculating new centers using Mean of the members.
8. **Until** stop- iteration criteria satisfied
9. **Return** clustering result

Figure 2. K-means algorithm [11].

The above comparison continues until reaching convergence and finding data on the local maximum. Data points that are associated with this local maximum are considered as cluster members [14]. The objective function of this algorithm is as follows in Equation (5):

$$f(x) = \frac{1}{nh^d}\sum_{i=1}^{R} K[\frac{x - x_i}{h}] \tag{5}$$

where x_i is a sample in the data set, h is the bandwidth, n is the number of samples, and d is the dimension of the examined space. Also, K is an optimal kernel that is usually considered in the following two ways [15] as in Equations (6) and (7):

$$K(x) = \begin{cases} 1, & -1 \le |x| \le 1 \\ 0, & else \end{cases} \tag{6}$$

$$\text{Gaussian } K(x) = \frac{1}{(2\pi)^{\frac{d}{2}}} e^{\frac{-1}{2}|x|^2} \tag{7}$$

First the appropriate bandwidth and the kernel function are determined. Then, this algorithm chooses the central points which are updated in each step based on the mean shift vector. In this research, the bandwidth is estimated based on the k-nearest neighbor method.

2.4. Agglomerative clustering

In data mining and statistics, hierarchical clustering is a clustering method whose purpose is to build a hierarchy of clusters. Hierarchical clustering methods are divided into two categories. Agglomerative approach: starting from the bottom, at each stage, two clusters are merged together and form a new cluster. Divisive approach: starting from the top, at each stage a cluster is decomposed into smaller clusters that are located at a lower level. New clusters are placed at higher levels and this process is repeated [16]. Each level of the hierarchy represents a category of data that can be viewed as a tree. Each leaf of the tree represents an initial observation, and the root of the tree is the collection of all observations. The results of a hierarchical clustering are generally displayed in the form of a dendrogram [17]. In order to understand which clusters should be grouped together or separated from each other, a measure of difference between clusters must be defined. In most methods, this criterion is achieved by defining a link criterion. The distance criterion determines the distance between two observations and the link criterion defines the distance between two sets of observations by a function of the distance between the observations of each set. Choosing an appropriate distance measure affects the shape of the clusters owing to the fact that for one distance measure, several observations can be close to each other. Some common distance measures for use in hierarchical clustering are given in Table 1 .

The link criterion defines the distance between two sets by a function of the pairwise distance between the observations of each set. Some common link criteria are listed in Table 2. In maximum distance, the distance between two observations a and b of two sets A and B is calculated based on the maximum, in minimum distance, this distance is calculated based on the minimum, and in average distance, it is calculated based on the average.

Table 1. Types of distance measure in hierarchical clustering [18].

Function	Formula
Euclidean	$\sum_{i=1}^{d}\sqrt{(x_i - y_i)^2}$
Manhattan	$\sum_{i=1}^{d}\lvert x_i - y_i\rvert$
Euclidean squared	$\sum_{i=1}^{d}(x_i - y_i)^2$
Maximum	$\sum_{i=1}^{d}\max_i\lvert x_i - y_i\rvert$

Table 2. Types of link criteria in hierarchical clustering [17].

Function	Formula
Maximum distance	$\max\{d(a,b): a \in A, b \in B\}$
Minimum distance	$\min\{d(a,b): a \in A, b \in B\}$
Average distance	$\frac{1}{\lvert A\rvert\lvert B\rvert}\sum_{a\in A}\sum_{b\in B} d(a,b)$

Agglomerative clustering starts with one cluster for each observation and at each step, two clusters that have the least difference from each other are aggregated using the criteria of Tables 1 and 2. This continues until the number of clusters reaches one. Agglomerative algorithm is shown in Figure 3. According to this figure, first each sample is considered as a cluster, then the distance between the clusters is calculated, and the clusters that have the smallest distance from each other are merged together. This process continues until only one cluster remains. In this research since our goal is to distinguish miners from non-miners, the number of clusters will be two in agglomerative algorithm. Also, the distance criterion is based on the Euclidean distance and the link criterion is chosen as the average distance.

2.5. Dynamic Mode Decomposition

Dynamic mode decomposition (DMD) is an equation-free data-driven matrix decomposition method developed in 2008. This method can provide an accurate reconstruction of coherent spatio-temporal structures resulting from nonlinear dynamical systems or an estimate of the short-term future of such systems. DMD approximates the modes of the Koopman operator, which is a linear operator and can represent nonlinear dynamics without linearization. This algorithm is also a dimensionality reduction algorithm which produces a set of modes based on a time series of data, each of which grows with a constant frequency. In particular, for linear systems, these modes and frequencies are similar to the normal modes of the system. The figure below shows the algorithm of this method based on singular value decomposition [19]. DMD algorithm is shown in Figure 4. According to this figure, the decomposition of singular values of the input matrix is first calculated. Then the first r singular values are selected and the $\Sigma_r$matrix is generated. Also $R^{r\times r}$ is a diagonal matrix comprising the first r singular values. Then the matrix $\tilde{A}$ is calculated. The eigenvectors of this matrix are the modes of the dynamic state decomposition, and its eigenvalues are the eigenvalues of the dynamic state decomposition.

3. Proposed Method

As stated earlier, the previous methods in the field of illegal miner detection have employed supervised learning methods. Since such algorithms need labeled data, and the generation of these labeled data is complicated, time consuming and expensive, It is preferable to choose unsupervised methods to detect miners. Additionally, supervised models are often trained based on grid factors obtained from AMI, and innovation in this field has been very limited. In this paper, our approach is to use unsupervised algorithms which do not require labeled data. In addition, the features created by DMD have been used in addition to the AMI features to build models. This has brought significant improvements. In this research, the features from AMI that are used are: power factor (phases 1, 2, 3), active power and reactive power. Furthermore, from the dynamic mode decomposition method, new features i.e. the modes and eigenvalues are produced. Then, various pre-processing methods including normalization, filling missing data, checking the correlation between features, summarizing data and removing outliers are applied to the data. Eventually, unsupervised algorithms are applied to the standardized data and the data are classified into two clusters. Then the best models were selected and the produced results were compared with supervised methods. The stages of this research are outlined as follows:

3.1. Data extraction

A real data set from Markazi Province Distribution Company including 58 labeled subscribers, of which 41 subscribers with positive labels and 17 subscribers with negative labels, has been used in this research. The sampling interval for all subscribers is 15 minutes.

3.2. Feature selection

In this research, the features are: power factor (phase 1, 2, 3), active power and reactive power. Additionally, DMD has been applied to produce new features. From DMD method, new features such as mode and eigenvalues are produced, which indicate the inherent behavior of the data, such as stability or instability .

3.3. Preprocessing

In this step, we prepare the extracted data for applying machine learning algorithms, which includes several parts:

3.3.1. Missing data

In the data set used for this research, there are missing data that are replaced by the two methods of feature average and previous moment sample, and the obtained results are checked.

Input: number of clusters, data
1. $C_{|X|} \coloneqq \{\{x\}|x \in X\}$
2. ***for*** i =|X|-1,….1 **do**
3. $find\ distinct\ clusters\ A, B \in C_{i+1} minimizing\ diam(A \cup B)$
4. $C_i \coloneqq (C_{i+1}\{A, B\}) \cup \{A \cup B\}$
5. **End for**
6. **Return** $C_1 \ldots \ldots C_{|X|}$

Figure 3. Agglomerative Algorithm [17].

input: $Matrices\ X, Y \in R^{m*n}, rank\ r(\leq \min(m,n))$
1. $Compute\ the\ truncated\ SVD\ of\ X(\approx U_r\Sigma_r V_r^T)$: $U_r \in R^{m\times r}$ $comprised\ the\ first\ r\ left\ singular\ vectors, \Sigma_r \in R^{r\times r}$ $is\ a\ diagonal\ matrix\ comprising\ the\ first\ r\ singular\ values,$ $V_r \in R^{n\times r}$ $comprices\ the\ first\ r\ right\ \ singular\ vectors$
2. $\tilde{A} = U_r^T Y V_r \Sigma_r^{-1}$
3. $compute\ the\ eigenpairs\ of \tilde{A}\ \tilde{w} = \lambda\tilde{w}\ , and\ let\ w = U_r\tilde{w}$

Figure 4. Dynamic mode decomposition Algorithm [20].

3.3.2. Outliers

To detect noisy or outlier data, two methods of box diagram and displaying the probability distribution of data are used, and finally, outlier data that reduce the quality of the algorithms are removed .

3.3.3. Data normalization

Data normalization is one of the most important pre-processing steps in data mining. Since the used features may have different intervals and lower the quality of clustering or supervised algorithms, in order to achieve more accurate results, it is necessary that the intervals of different features be somehow the same or close to each other. For this research, minmax and z-score methods are used and the results are compared with each other.

3.3.4. Data aggregation

Since the data used in this research are in the form of time series, in order to be able to use machine learning algorithms, summarization is obvious. For this reason, aggregation is done using two aggregation methods based on maximum and average, and the results are compared with each other.

3.3.5. Data Correlation

Correlation is a statistical term that examines whether two variables are related to each other or not. Features with high correlation are linearly related to each other and have approximately the same effect on the dependent variable; Therefore, when two features have a high correlation, one of these two features can be left out. In this research, the correlation between the characteristics is checked and the characteristics that have a high correlation are removed.

3.4. Model selection

At this stage, a model must be selected that can be used to separate and categorize the data. In this research, since the labeled data are available, however, the goal is to build an unsupervised method, various unsupervised algorithms are applied to the data and the results obtained are compared with the results of supervised algorithms. Different clustering algorithms are used for data classification, which can be based on distance (k-means), density (DBSCAN and mean shift) and rank order (agglomerative). In the supervised method, neural network, decision tree, support vector machine and ensemble learning algorithms are used and the obtained results are compared.

3.5. Selection of hyperparameters

In k-means and agglomerative algorithms, since our goal is to distinguish miners from non-miners, the number of clusters is equal to 2; But in other algorithms, other hyperparameters must be determined. For example, in the average transfer algorithm, the bandwidth is estimated based on the k-nearest neighbor method. To put it more clearly, the average distance between samples located in a cluster is calculated. In this paper the bandwidth in mean-shift algorithm is chosen as a value between 0 and 2 according to KNN. And in the DBSCAN algorithm, hyperparameter selection is done according to the accuracy of the silhouette index; The higher the index, the better hyperparameters are selected. In this paper the hyper-parameters of DBSCAN are set based on the grid search algorithm [21]. In fact, eps has been selected as a number between 0.1 and 2 and minimum samples are numbers between 1 and 10.

3.6. Prediction

In this step, the model built by clustering and supervised algorithms is applied to new data to predict their labels. In the following, the work done in this research will be examined in the form of several models.

In the next step, according to the mentioned points, unsupervised models NF, FNF, MNF, ZNF, CNF were built based only on the recorder features from the AMI. Then, the features created with the use of DMD method were used and unsupervised models MO, EV, MOEV were created which are based on mode features, eigenvalues and a combination of the two. Finally, the unsupervised models MONF, NFEV, MONFMEM, MONFME, which are based on a combination of recorded features and DMD ones were built. Mean shift, agglomerative k-means, and DBSCAN algorithms are used in all these models and their only difference was in the used features and pre-processing methods. Finally, in order to compare the results obtained with the supervised models, MONFMEM, MONF, NFEV, MNF, NF were built. In these models, neural network algorithms, ensemble learning, support vector machine, and decision tree are used. Figure 5 shows the structure of this research. According to this figure, first the appropriate features for the Distribution Company of Markazi Province data set are selected, then they are prepared using pre-processing methods such as data normalization, data aggregation, filling missing values and determining the correlation between features for applying unsupervised algorithms. Then, in order to evaluate the obtained results, supervised algorithms are also applied and the obtained results are compared.

4. Evaluation

In this section, the results obtained from the models built in the previous chapter are evaluated. The results obtained with the help of DMD features have the highest quality. Also, the agglomerative algorithm, which is a type of hierarchical algorithm, had the highest accuracy in data classification, and this accuracy was evaluated at 74%. By evaluating the results of supervised models, it was found that DMD features performed best and the highest accuracy achieved was 85%. By examining the two results obtained from the unsupervised and supervised models, it can be seen that there is almost a ten

percent difference between the two aproaches. Figures 6 and 7 show the comparison of the results of the supervised and unsupervised models. In the models built with the help of unsupervised methods, when only features recorded from the AMI were used, the highest accuracy of 62% was achieved. With the addition of built-in features using DMD, this value has increased to 74%. For monitoring models, when only AMI features were used, the highest accuracy achieved was 79%. With the addition of constructed features with the help of DMD, this value increased to 85%. Therefore, it can be concluded that the accuracy in both cases was higher when the DMD features were used. In Figures 8 and 9, a comparison between true positive and true negative labels is made. As it can be seen, when the features made from DMD were used along with the AMI features, there were more true positive labels, while when only AMI features were used, there were more true negative labels. In Figures 10 and 11, a comparison is made between false positive and false negative labels. As it can be seen, when the features made from DMD were used along with the AMI features, there were more false positive labels, while when only AMI features were used, there were more false negative labels .

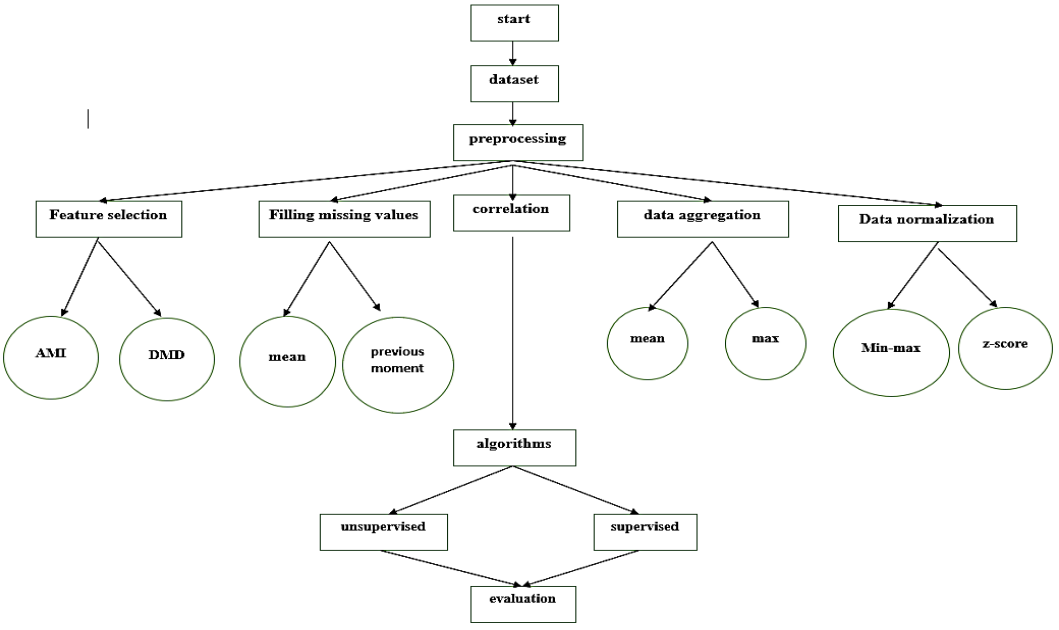


Figure 5. The procedure of the miner detection based on DMD and unsupervised learning.

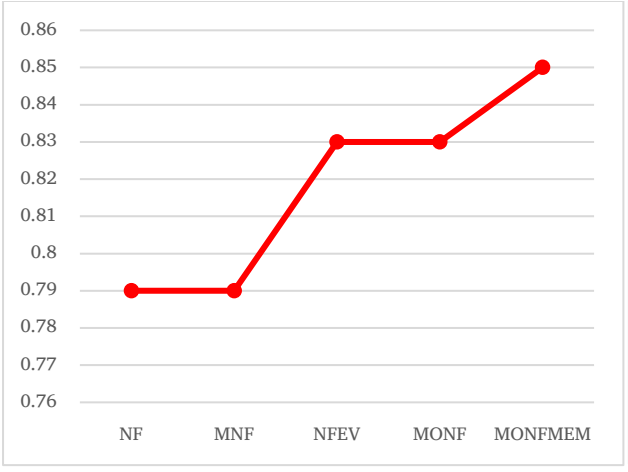


Figure 6. Accuracy of supervised models.

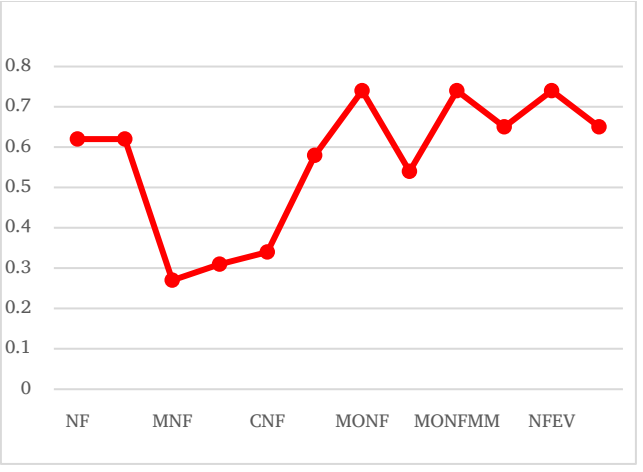


Figure 7. Accuracy of unsupervised models.

In Figures 12, 13 and 14, a comparison has been made between precision, recall and support of unsupervised models. In this research, precision measures how good the model is at assigning positive samples to the positive class. That is, how accurate the miner prediction is. Recall measures how good the model is in detecting positive samples. The measures provide valuable information, but the objective is to improve recall without affecting precision. Sensitivity measures how apt the model is to detect samples in the positive class. Therefore, given that miners are a positive class, sensitivity quantifies how many of the actual miners are correctly predicted as miners. As it is known, the values of sensitivity, recall and precision in the three models MONF, NFEV and MONFMEM had the highest value, which means that these models have performed well in detecting subscribers with miners.

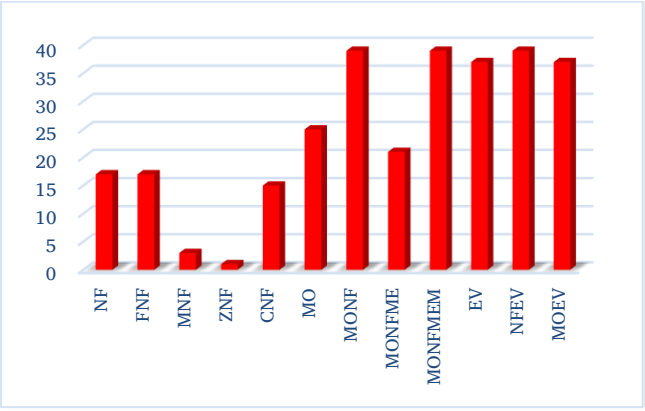


Figure 8. Number of true positive labels of unsupervised models.

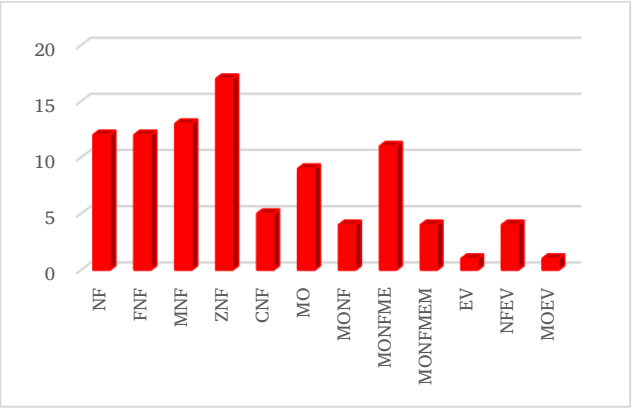


Figure 9. Number of true negative labels of unsupervised models.

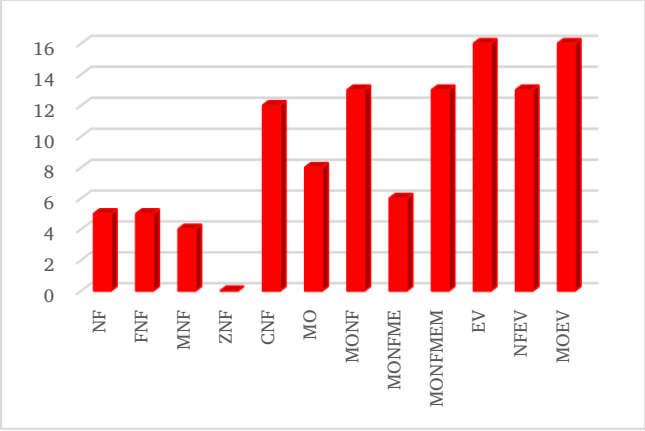


Figure 10. Number of false positive labels of unsupervised models.

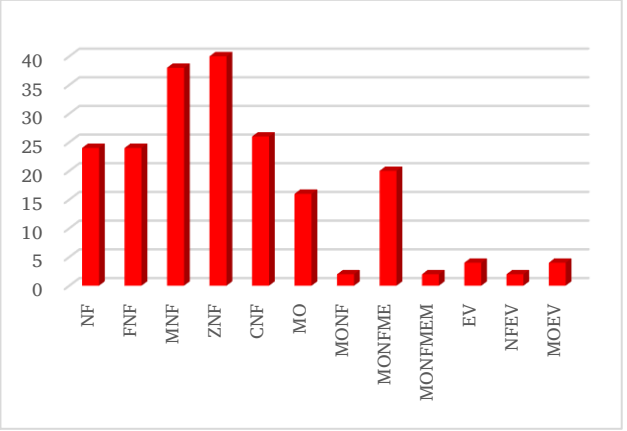


Figure 11. Number of false negative labels of unsupervised models.

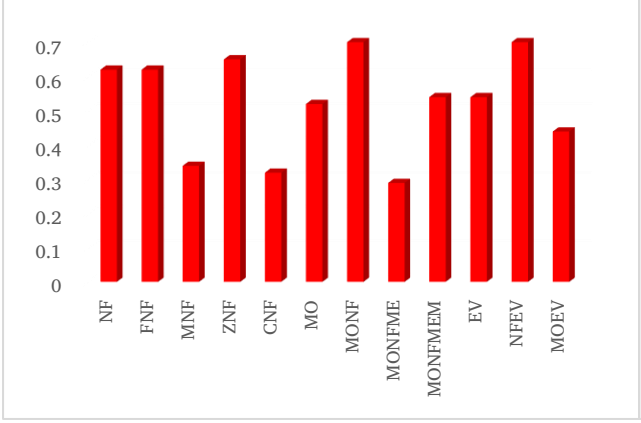


Figure 12. Precision for unsupervised models.

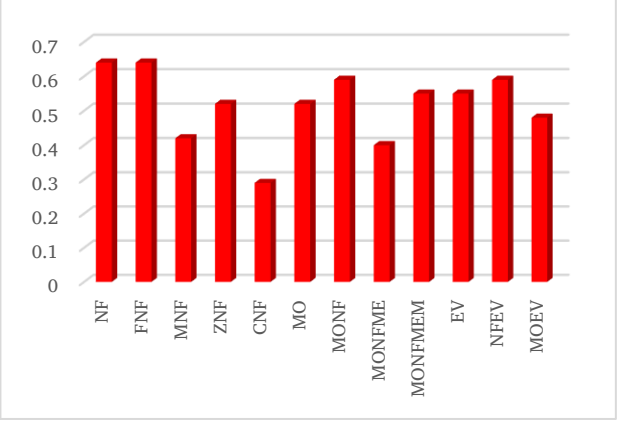


Figure 13. Recall for unsupervised models.

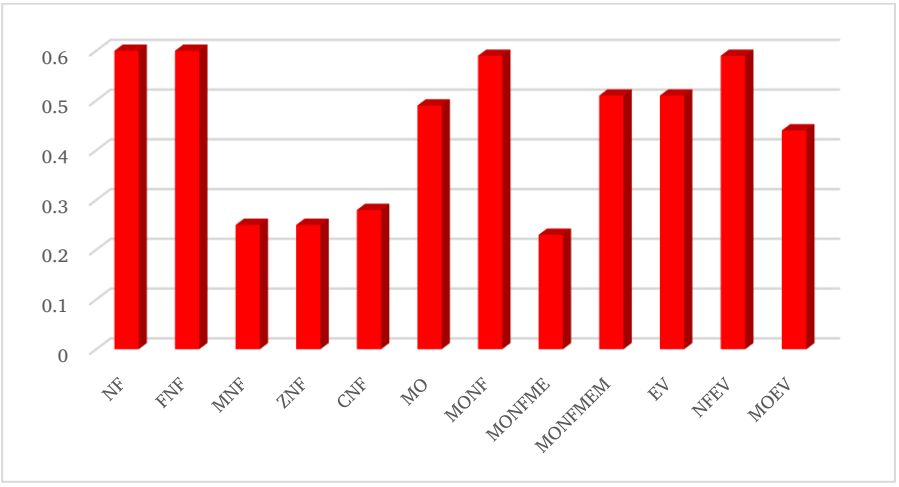


Figure 14. Sensitivity for unsupervised models.

5. Conclusion

This research seeks to provide a method to provide an efficient method to identify unauthorized miners. The proposed models which used AMI and DMD features together produced the most superior results in detecting miners by using unsupervised algorithms. Unsupervised algorithms MONF, NFEV and MONFMEM showed 74% accuracy which is significant compared to the supervised algorithms. The DMD method has produced useful features that have increased the accuracy of the algorithms. Since the data of the consumers collected at distribution companies are highly confidential and accessibility to these data is difficult, data augmentation for improving model generalizability and scalability is suggested as future work.

References

[1] B. Dindar, and Ö. Gül, "The Detection of Illicit Cryptocurrency Mining Farms with Innovative Approaches for the Prevention of Electricity Theft," *Energy & Environment*, vol. 33, no. 8, pp. 1663-1678, 2022.
[2] C. Chahla, H. Snoussi, L. Merghem, and M. Esseghir, "A Deep Learning Approach for Anomaly Detection and Prediction in Power Consumption Data," *Energy Efficiency*, vol. 13, no. 8, pp. 1633-1651, 2020.
[3] M. Kardi, T. AlSkaif, B. Tekinerdogan, and J. P. S. Catalao, "Anomaly Detection in Electricity Consumption Data using Deep Learning,"*2021 IEEE International Conference on Environment and Electrical Engineering and 2021 IEEE Industrial and Commercial Power Systems Europe,* pp. 1-6, 2021.
[4] C. Zhou, Z. Fang, et al., "Using Long Short-Term Memory Networks To Predict Energy Consumption Of Air-Conditioning Systems," *Sustainable Cities and Society*, vol. 55, 102000, 2020.
[5] M. Li, K. Zhang, J. Liu, H. Gong, and Z. Zhang, "Blockchain-Based Anomaly Detection of Electricity Consumption in Smart Grids," *Pattern Recognition Letters*, vol. 138, pp. 476-482, 2020.
[6] Z. Ouyang, X. Sun, J. Chen, D. Yue, and T. Zhang, "Multi-View Stacking Ensemble for Power Consumption Anomaly Detection in the Context of Industrial Internet of Things," *IEEE Access*, vol. 6, pp. 9623-9631, 2018.
[7] P. Nerurkar, "Illegal Activity Detection on Bitcoin Transaction using Deep Learning," *Soft Computing*, vol. 27, no. 9, pp. 5503-5520, 2023.
[8] S. Wu, Q. Liu, and J. Li, "Research on the Detection of Illegal Transactions in Currency Transactions Based on Blockchain Technology," *International Journal of Network Security*, vol. 26, no. 1, pp. 19-24, 2024.
[9] K. Kolesnikova, O. Mezentseva, and T. Mukatayev, "Analysis of Bitcoin Transactions to Detect Illegal Transactions using Convolutional Neural Networks," *IEEE International Conference on Smart Information Systems and Technologies,* pp. 1-6, 2021.
[10] A. Wahrstätter, J. Gomes, S. Khan, and D. Svetinovic, "Improving Cryptocurrency Crime Detection: Coinjoin Community Detection Approach," *IEEE Transactions on Dependable and Secure Computing*, vol. 20, no. 6, pp. 4946-4956, 2023.
[11] M. Amiri, and H. Askari, "Illegal Miner Detection Based on Pattern Mining: A Practical Approach," *Journal of Computing and Security*, vol. 9, no. 2, pp. 1-10, 2022.
[12] M. Ahmed, R. Seraj, and S. M. S. Islam, "The K-Means Algorithm: A Comprehensive Survey and Performance Evaluation," *Electronics*, vol. 9, no. 8, p. 1295, 2020.
[13] J. Qian, Y. Zhou, et al. "MDBSCAN: A Multi-Density DBSCAN Based on Relative Density," *Neurocomputing*, vol. 576, 127329, 2024.
[14] C. Cariou, S. Le Moan, and K. Chehdi, "A Novel Mean-Shift Algorithm for Data Clustering," *IEEE Access*, vol. 10, pp. 14575-14585, 2022.
[15] Y. A. Ghassabeh, and F. Rudzicz, "Modified Subspace Constrained Mean Shift Algorithm," *Journal of Classification*, vol. 38, pp. 27-43, 2021.
[16] X. Ran, Y. Xi, Y. Lu, X. Wang, and Z. Lu, "Comprehensive Survey on Hierarchical Clustering Algorithms and the Recent Developments," *Artificial Intelligence Review*, vol. 56, no. 8, pp. 8219-8264, 2023.
[17] G. J. Szekely, and M. L. Rizzo, "Hierarchical Clustering via Joint Between-Within Distances: Extending Ward's Minimum Variance Method," *Journal of classification*, vol. 22, no. 2, pp.151-184, 2005.
[18] M. R. Ackermann, J. Blömer, D. Kuntze, and C. Sohler, "Analysis of Agglomerative Clustering," *Algorithmica*, vol. 69, pp. 184-215, 2014.
[19] P. J. Schmid, "Dynamic Mode Decomposition and its Variants," *Annual Review of Fluid Mechanics*, vol. 54, pp. 225-254, 2022.
[20] J. A. Rosenfeld, R. Kamalapurkar, L. F. Gruss, T. T. ND Johnson, "Dynamic Mode Decomposition for Continuous Time Systems with the Liouville Operator," *Journal of Nonlinear Science*, vol. 32, pp. 1-30, 2022.
[21] K. Khan, S. U. Rehman, K. Aziz, S. Fong, and S. Sarasvady, "DBSCAN: Past, Present and Future," *The Fifth International Conference on the Applications of Digital Information and Web Technologies*, pp. 232-238, 2014.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. The ethical issues, including plagiarism, informed consent, misconduct, data fabrication and/or falsification, double publication and/or submission, redundancy, have been completely observed by the authors.

Bibliography

Alireza Simorgh was born in 1997 in Boroujerd, Iran. He obtained his Bachelor's and Master's degrees in Electrical Engineering and Electrical Control Engineering, respectively, from Ayatollah Boroujerdi University in Boroujerd, Iran, and Arak University in Arak, Iran, in 2021 and 2024. Currently, he is working as an Industry Liaison Officer at Shoroue Innovation Center. His specialized interests include robotics, multi-agent systems, data science and machine learning, and fuzzy systems.
Email: alirezasimorgh76@gmail.com
ORCID: 0009-0000-8996-0355
Contribution Statement: Data curation, Formal analysis, Investigation, Methodology, Resources, Software, Validation, Visualization, Roles/Writing - original draft.

Khosro Khandani received his BS degree in electrical engineering from Sahand University of Technology, Tabriz, Iran, in 2007; MS degree in electrical engineering from Iran University of science and Technology, Tehran, Iran, in 2011; Ph.D. degree in electrical engineering from Tarbiat Modares University, Tehran, Iran, in 2016. He is currently an associate professor in electrical engineering department, faculty of engineering, Arak University, Arak, Iran. His research interests include cooperative control of multi-agent systems, fuzzy systems, fractional-order systems and machine learning.
Email: k-khandani@araku.ac.ir
ORCID: 0000-0002-1762-3541
Contribution Statement: Conceptualization, Investigation, Project administration, Supervision, Roles/Writing-original draft, Writing-review & editing.

Maryam Amiri received the BS degree in computer engineering from Arak University, Arak, Iran, in 2009; the MS degree in computer engineering from the Bu-Ali Sina University, Hamedan, Iran, in 2012; the Ph.D. degree in computer engineering from the University of Tabriz, Tabriz, Iran, in 2018. She is currently an assistant professor in the Department of Computer Engineering, the faculty of engineering, Arak University, Arak, Iran. Her research interests include cloud computing, machine learning, and data mining.
Email: m-amiri@araku.ac.ir
ORCID: 0000-0002-7411-9552
Contribution Statement: Conceptualization, Data curation, Project administration, Resources, Writing-review & editing.