# Implementation of SVM Kernels for Identifying Irregularities Usage of Smart Electric Voucher

*by* Armin Lawi

# Implementation of SVM Kernels for Identifying Irregularities Usage of Smart Electric Voucher


1st Edy Budiman
*department of informatics*
*mulawarman university*
Samarinda, Indonesia
edybudiman.unmul@gmail.com
http://orcid.org/0000-0002-3164-5157

2nd Armin Lawi
*department of computer science*
*hasanuddin university*
Makassar, Indonesia
armin@unhas.ac.id
http://orcid.org/0000-0003-1023-6925

3rd Supriyadi La Wungo
*Departement of informatics*
*stmik kreatindo monokwari*
Monokwari, Indonesia
supriyadi.la.wungo@gmail.com



***Abstract*— Statistical methods and machine learning have been widely used to identify deviations in the use of electrical energy for prepaid services. The paper applies the Support Vector Machine method to identify prepaid electricity usage irregularities that can overcome classification and regression problems with linear or nonlinear kernels with high accuracy and relatively small error rates. The results showed that the predictions of morbidity of electricity voucher purchase transactions, the amount of test data used did not affect the accuracy, precision, and memory values of the Linear and Polynomial kernels, the values obtained were all 100%. This shows that the addition of test data, the value of False Positive and False Negative remains 0. Thus, in each additional test data value of precision, accuracy and memory do not change. However, in the RBF kernel, the value of accuracy and precision decreases as the amount of test data increases.**



***Keywords*— *support vector machine, prepaid electricity, RBF, polynomial kernels.***


## I. Introduction

Prepaid electricity is a product issued by. The Indonesian government-owned electricity service company has several advantages and ease of use due to the electricity usage payment system from the service provider, where the user pays first then can use electricity from the service provider. The use of Prepaid Electricity in Indonesia to replace Postpaid Electricity previously applied in residential housing where users will pay electricity bills every month in accordance with the use of electricity.

The use of smart electric voucher in Indonesia is intensively encouraged in order to replace the post-paid electrical consumption in which customers will pay utility bills each month in accordance with their electricity consumption [1]. Despite the ease of smart electric vouchers that have been provided by energy providers, there are still many customers found and indicated to have electricity theft in the field. For instance, the customer changed the wiring pattern on kWh meter indicator and by illegally connecting cables with direct cables, tapping or jumping electricity [2], [3].

In general, the methods used by service providers to detect electricity theft are physical condition detection methods and customer consumption investigation methods. This physical condition detection method requires checking the physical condition of all electrical components in the house and finding out if there are any peculiarities. Meter seal checks, service cables, meter accuracy and additional cables are required. The purpose of this method is to find out what might have happened, whether it was indeed a victim of electric theft or an oddity in the electric power and the bill was caused by damage to its components.

The customer consumption investigation method is an investigation method if there is a possibility that the perpetrators of electricity theft are the customers themselves. Customer's electricity consumption habits will be known. If the electricity bill is lower than usual, the service provider will certainly question it. Although this method is less effective, it can lead to further investigations to avoid the possibility of electricity theft. And usually, this method will be carried out by the officer of the service provider as the authorized party in managing electricity problems in Indonesia.

Several methods are used by service providers to monitor electricity theft, such as through smart networks (Smart Grid) using the Automatic Meter Reading (AMR) system that can record electricity consumption in detail, the service provider can find out quickly where there is electricity theft. This method is in the form of installation of electricity consumption logging devices which are installed on the customer's meter and integrated into the recording system at the service provider's centre.

AMR has now been installed on customer meters with capacities above 40 kV. This system approach is where the service provider no longer needs to send officers to record electricity consumption to its customers. So no need to read the meter. With these strict records, you can quickly detect the theft of electricity. Because if there is electricity consumption that is not as usual, the service provider can immediately prove.

Statistical methods and machine learning have been widely used to identify abuse in the use of electrical energy such as Logistic Regression, Artificial Neural Networks, k-Nearest Neighbors, Decision Trees, Fuzzy C-Means, Naïve Bayes and Genetic Algorithms. In previous studies, the method of machine learning has a good level of accuracy in identifying the abuse of the use of electrical energy when compared with statistical methods. This is due to the machine learning method being able to map input data into higher dimensions (nonlinear) [4], [5].

One of the machine learning methods that recently received much attention is Support Vector Machine (SVM). SVM is a linear classifier and subsequently developed to work on non-linear problems by incorporating the concept of kernel tricks in high-dimensional workspaces [6]. Support Vector Machine has been successfully applied in real-world problems and generally provides better solutions than conventional methods such as Artificial Neural Networks [7]. SVM can overcome the problem of classification and regression with linear or nonlinear kernels which can become a learning algorithm capability for classification and regression [7]. SVM also has high accuracy and a relatively small error rate, the ability to overcome overfitting does not require data that is too large and can be used to identify and predict. From the description above, this study will use the Support Vector Machine (SVM) classification to identify deviations from the Use of Smart Electric Vouchers.

## II. Related Work

### A. Support Vector Machine (SVM)

Support Vector Machine (SVM) is known as the most advanced machine learning technique after the previous machine learning known as Neural Network (NN) [8]. The basic principle of SVM is a linear classifier and subsequently developed to work on non-linear problems by incorporating the concept of kernel tricks in high-dimensional workspaces. Both SVM and NN have been successfully used in pattern recognition. Support Vector Machine is a learning machine method that works on the principle of Structural Risk Minimization (SRM) with the aim of finding the best hyperplane that separates two classes in the input space [9].

The best hyperplane is a hyperplane that is located halfway between two sets of objects from two classes. The best hyperplane separator between the two classes can be found by measuring the hyperplane's margin and finding its maximum point. Margin is the distance between the hyperplane and the closest pattern of each class. The closest pattern is referred to as a support vector [10]. According to Li Guangcai et al. [11], SVM is a classification method for finding the best hyperplane values that are capable of finding optimal global solutions. So the accuracy value is not easy to change.

The concept of SVM can be explained simply as an attempt to find the best hyperplane that functions as a separator of two classes in the input space. The best hyperplane separator between the two classes can be found by measuring the margin of the hyperplane and finding its maximum point. Margin is the distance between the hyperplane and the closest data from each class. The closest subset of training data sets is called a support vector. The solid line shows the best hyperplane, which is right in the middle of the two classes, while the points of a square and circle in a black circle are support vectors. The effort to find the optimal hyperplane location is the core of the learning process in SVM. It is assumed that both class-1 and +1 can be completely separated by a dimensionless hyperplane, which is defined in "equation (1)".

$$\vec{w}.\vec{x} + b = 0 \tag{1}$$

A pattern $\vec{x}$ belonging to class -1 (negative sample) can be formulated as a pattern that satisfies inequality ("Equation 2").

$$\vec{w}.\vec{x} + b \leq -1 \tag{2}$$

A pattern $\vec{x}$ belonging to class +1 (positive sample) can be formulated as a pattern that satisfies inequality ("Equation 3").

$$\vec{w}.\vec{x} + b \geq -1 \tag{3}$$

The largest margin can be found by maximizing the value of the distance between the hyperplane and its closest point, which is $1 / \|\vec{w}\|$. This can be formulated as a Quadratic Programming (QP) problem, which is to find the minimum point "Equation (4)", taking into account the constraint "Equation (5)".

$$\min \tau(w) = \frac{1}{2}\|\vec{w}^2\| \tag{4}$$

$$y_i(\vec{w}.\vec{x} + b) - 1 \geq 0, \forall i \tag{5}$$

Data training which is used when constructing the SVC model in reality mostly cannot be separated linearly. To resolve the classification problem linearly, some slack variables will be mapped into high dimensions using the RBF and Polynomial kernels in SVM so that errors and disturbances during dataset training can be reduced [2]. Linear, RBF (radial basis function) and Polynomials kernels which are used in this paper as follows [12]

### B. Kernel Function

In general, problems in the real world domain (real world problem) are rarely linear in nature and most are non-linear [13], [14]. To solve non-linear problems, SVM is modified by include the Kernel function. In non-linear SVM, the $\vec{x}$ data is first mapped by the function $\Phi$ ($\vec{x}$) to a higher dimensional vector space. In this new vector space, the hyperplane that separates the two classes can be constructed. This is in line with the cover theory which states "If a transformation is non-linear and the dimensions of the feature space are high enough, then the data in the input space can be mapped to a new feature space, where the patterns at high probability can be linearly separated.

Kernel Trick, which is formulated in "Equation (6)".

$$K(x_i, x_j) = \Phi(x_i).\Phi(x_j) \tag{7}$$

Kernel Trick provides various facilities, because in the learning process of SVM, to determine the support vector, and only enough to know the function of the kernel that is used, and do not need to know the form of non-linear functions $\Phi$. The various types of Kernel functions are known, as summarized in Table 1.

TABLE I. TYPE OF KERNEL FUNCTION

| Type of Kernel | Definition |
|---|---|
| *Polynomial* | $K(\vec{x}_i, \vec{x}_j) = K(\vec{x}_i, \vec{x}_j + 1)^p$ |
| *Gaussian* | $\mathbf{K}(\vec{\mathbf{x}}_i, \vec{\mathbf{x}}_j) = \boldsymbol{exp}(\frac{\|\vec{\mathbf{x}}_i, \vec{\mathbf{x}}_j\|^2}{\mathbf{2\sigma^2}})$ |
| *Linear* | $\mathbf{K}(\vec{\mathbf{x}}_i, \vec{\mathbf{x}}_j) = \boldsymbol{x_i^T x_j}$ |
| *Sigmoid* | $\mathbf{K}(\vec{\mathbf{x}}_i, \vec{\mathbf{x}}_j) = \boldsymbol{tanh}(\boldsymbol{\alpha}\vec{\mathbf{x}}_i, \vec{\mathbf{x}}_j \boldsymbol{\beta})$ |

Where $x_i$ is support vector and $x_j$ is the data value of the attribute and $\gamma$ is the kernel parameter. The main purpose of this kernel is to make the best decision limit when doing the classification of training sets into two parts [15].

### C. Performance Evaluation

To measure performance, use confusion matrix. Confusion matrix provides decisions obtained in training and testing [16].

Confusion matrix provides an assessment of the performance classifications based on true or false objects [16], [17]. Confusion matrix is a 2-dimensional matrix that illustrates comparison between predicted results and reality [18]. For unbalanced data, accuracy is more dominated by accuracy in minority class data, then the right metrics are AUC (Area Under the ROC Curve), F-Measure, G-mean, overall accuracy, and accuracy for minority classes [19]. Minority class accuracy can use the TP-rate/recall (sensitivity) metric. G-Mean and AUC are more comprehensive evaluators of predictors in the context of imbalance [20]. To do the calculations used “equation (4)” [21].

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \tag{8}$$

### D. Postpaid Electricity Services

During this time electricity customers get postpaid electricity services, i.e. customers use electricity first and pay later the following month. With postpaid electricity services, every month service providers must record meters, calculate and issue accounts to be paid by customers, bill customers who are late or not paying, and cut off electricity if consumers are late or do not pay electricity bills after a certain time. This mechanism is not implemented in postpaid electricity service systems (Smart Electric).

Smart Electricity is a prepaid electricity service that allows customers to control their own electricity usage according to their needs and abilities. As with refill pulses on cell phones, on a smart electricity system, customers first purchase credit (vouchers / tokens) for refillable electricity through ATM outlets in a number of banks or through online electricity bill payment counters [22].

Prepaid electricity is the latest product to replace postpaid electricity. Prepaid electricity is issued by electric energy supply companies that have several advantages and ease of use because electricity consumption is fully controlled by the customer so that it is more convenient and controlled. With Prepaid Electricity the use of electricity is more controlled, there are no sanctions to cut off electricity, there is no meter recording, no charge is charged, electricity tokens are easy to get, there are no late fees and the active period is unlimited. Whereas postpaid electricity usage without a grace period, sanctions for power outages, recording of electricity meters, are subject to charges and delays [23], [22].

## III. Result and Discussion

### A. Dataset

The research dataset uses a history of buying customer transaction pulses originating from the electricity service provider (State Electricity Company - PLN), in one of the provincial capitals in Sulawesi-Indonesia. The Table II following is a sample of historical data on prepaid credit purchase transactions.

TABLE II. Minimum Usage Per Power

| Installed Power (VA) | Minimum usage (kWh) |
|---|---|
| 450 | 18 |
| 900 | 36 |
| 1300 | 52 |
| 2200 | 88 |
| 4400 | 176 |

The minimum usage per power is the minimum usage determined by the company for each power consumption. Table III is a sample of historical data on prepaid credit purchase transactions.

TABLE III. Samples of Prepaid Pulse Purchase Transaction History Data

| Installed Power (VA) | Minimum usage (kWh) | Total Usage (kWh) | Total credit (kWh) | Status |
|---|---|---|---|---|
| 1300 | 52 | 10 (B) | 62 (B) | *reasonable* |
| 900 | 36 | -10 (B) | 20 (A) | *unreasonable* |
| 1300 | 52 | -21 (A) | 31 (B) | *unreasonable* |
| 3500 | 140 | -78 (A) | 62 (B) | *unreasonable* |
| 1300 | 52 | -11 (A) | 41 (B) | *unreasonable* |
| 3500 | 140 | -56(C) | 82 (C) | *unreasonable* |
| 900 | 36 | -13 (B) | 23 (A) | *unreasonable* |
| 900 | 36 | 78 (C) | 114 (C) | *reasonable* |
| 2200 | 88 | -57 (A) | 23 (A) | *unreasonable* |
| 3500 | 140 | 57 (C) | 197 (C) | *unreasonable* |

In Table III, the Value 1300 is the installed power used by the customer. A value of 52 is the minimum usage standard of 1300 power used by customers. Value 10 is the total power consumption of the customer for a month obtained from the difference between the total credit in the voucher and the minimum usage. Value 62 is the total credit transaction in the voucher carried out by the customer for a month. While the status is fairness of the customer's electricity usage in using vouchers. If the difference between the value of the column on total pulses in the voucher and the column value at minimum usage is negative in the total power consumption column, the status will be unreasonable and vice versa if the value of the difference between the total column pulses in the voucher and the minimum usage column is positive in the column the total power consumption then the status will be reasonable.

### B. Accuracy level, Precision and calling Results

The analysis used in the classification of the identification of the misuse of electricity is the Support Vector Machine method. The training process will use a portion of the total data in the dataset to be used.

The level of accuracy is obtained from the amount of data whose class is accurately predicted by each model. The level of accuracy in this study is different for each model used. The percentage comparison between testing data and training data used for each model is presented in Table IV

TABLE IV. Samples of Prepaid Pulse Purchase Transaction History Data

| Kernel | Performance | Data Testing (%) | | | |
|---|---|---|---|---|---|
| | | 20 | 40 | 60 | 80 |
| Rbf | *accuracy* | 1 | 1 | 1 | 1 |
| | *precision* | 1 | 1 | 1 | 0.99 |
| | *recall* | 1 | 1 | 1 | 0.99 |
| Linear | *accuracy* | 1 | 1 | 1 | 1 |
| | *precision* | 1 | 1 | 1 | 1 |
| | *recall* | 1 | 1 | 1 | 1 |
| Polynomial | *accuracy* | 1 | 1 | 1 | 1 |
| | *precision* | 1 | 1 | 1 | 1 |
| | *recall* | 1 | 1 | 1 | 1 |

Based on the experiments in table 4 and “Fig. 1”, it shows that by adding the testing data of the Support Vector Machine classification method to the kernel: Rbf, linear and polynomial can predict the unreasonable use of electrical energy with an accuracy value of 99% - 100%. The highest accuracy, precision and recall values are generated when comparing the amount between testing and training data 20%, 40%, 60%, and 80%. The more the amount of testing data is used, the higher the resulting accuracy value is 1. Similarly, the recall value when the amount of testing data increases, the recall value also increases.

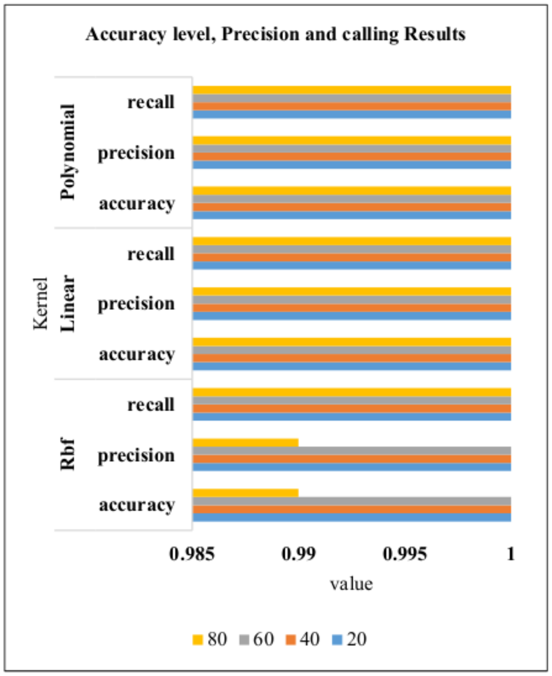


Fig. 1. Accuracy level, precision and calling results

In Table IV and “Fig. 1” shows the test results from the Linear and Polynomial kernels showing values that do not affect the value of accuracy, precision, and recall. The accuracy, precision, and memory values obtained are all 100%. This shows that with the addition of test data, the value of False Positive and False Negative remains 0. Thus, in each additional test data value of precision, accuracy and recall do not change. These results reveal that in each additional test data the value of precision, accuracy or recall did not change.

However, in the RBF kernel, the value of accuracy and precision decreases as the amount of test data increases. This is because the TN value has decreased. While the recall value does not change, it is equal to 100%. It can be said that the SVM method can predict morbidity accurately.

## IV. CONCLUSION

Based on experimental results from the research Implementation of the kernel Support Vector Machine (RBF, precision and accuracy) to identify irregularities in the use of Smart Electric Vouchers shows that the Support Vector Machine method can accurately predict morbidity with an accuracy level of 0.99 - 1 or 99.9% - 100%. This research can be one of the main references as a means of approaching identification and prediction of the problem of irregularities in the use of Smart Electric Vouchers.

## REFERENCES


[1] I. S. Jha, S. Sen, and V. Agarwal, “Advanced metering infrastructure analytics - A Case Study,” in *2014 18th National Power Systems Conference, NPSC 2014*, 2015.

[2] B. Dangar and S. K. Joshi, “Electricity theft detection techniques for metered power consumer in GUVNL, GUJARAT, INDIA,” in *2015 Clemson University Power Systems Conference, PSC 2015*, 2015.

[3] V. Ford, A. Siraj, and W. Eberle, “Smart grid energy fraud detection using artificial neural networks,” in *IEEE Symposium on Computational Intelligence Applications in Smart Grid, CIASG*, 2015.

[4] M. Awad and R. Khanna, *Efficient learning machines: Theories, concepts, and applications for engineers and system designers*. 2015.

[5] E. Budiman, N. Dengen, Haviluddin, and W. Indrawan, “Integrated multi criteria decision making for a destitute problem,” in *2017 3rd International Conference on Science in Information Technology (ICSITech)*, 2017, vol. 2018-Janua, pp. 342–347.

[6] S. S. Keerthi, O. Chapelle, and D. DeCoste, “Building support vector machines with reduced classifier complexity,” *Journal of Machine Learning Research*, 2006.

[7] N. Barakat and A. P. Bradley, “Rule extraction from support vector machines: A review,” *Neurocomputing*, 2010.

[8] S. M. Guzman, J. O. Paz, M. L. M. Tagert, and A. Mercer, “Artificial neural networks and support vector machines: Contrast study for groundwater level prediction,” in *American Society of Agricultural and Biological Engineers Annual International Meeting 2015*, 2015.

[9] T. Bellotti and J. Crook, “Support vector machines for credit scoring and discovery of significant features,” *Expert Systems with Applications*, 2009.

[10] I. Aydin, M. Karakose, and E. Akin, “A multi-objective artificial immune algorithm for parameter optimization in support vector machine,” *Applied Soft Computing Journal*, 2011.

[11] G. Li, J. You, and X. Liu, “Support Vector Machine (SVM) based prestack AVO inversion and its applications,” *Journal of Applied Geophysics*, 2015.

[12] A. R. Almeida, O. M. Almeida, B. F. S. Junior, L. H. S. C. Barreto, and A. K. Barros, “ICA feature extraction for the location and classification of faults in high-voltage transmission lines,” *Electric Power Systems Research*. 2017.

[13] E. Budiman, M. Wati, D. Indra, D. Moeis, and M. Jamil, “QoE and QoS Evaluation for Academic Portal in Private Higher Education Institution,” in *2018 International Conference on Computer Engineering, Network and Intelligent Multimedia, CENIM 2018 - Proceeding*, 2019.

[14] E. Budiman, N. Pusnitasari, M. Wati, Haeruddin, J. A. Widians, and A. Tejawati, “Mobile learning media for computer science course,” *International Electronics Symposium on Knowledge Creation and Intelligent Computing, IES-KCIC 2018 - Proceedings*, pp. 262–267, 2019.

[15] A. Jindal, A. Dua, K. Kaur, M. Singh, N. Kumar, and S. Mishra, “Decision Tree and SVM-Based Data Analytics for Theft Detection in Smart Grid,” *IEEE Transactions on Industrial Informatics*, 2016.

[16] E. Budiman, Haviluddin, N. Dengan, A. H. A. H. Kridalaksana, M. Wati, and Purnawansyah, “Performance of Decision Tree C4.5 Algorithm in Student Academic Evaluation,” in *Lecture Notes in Electrical Engineering*, vol. 488, R. Alfred, H. Iida, A. A. Ag. Ibrahim, and Y. Lim, Eds. Singapore: Springer Singapore, 2018, pp. 380–389.

[17] E. Budiman, U. Haryaka, J. R. J. R. Watulingas, and F. Alameka, “Performance rate for implementation of mobile learning in network,” in *2017 4th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI)*, 2017, vol. 2017-Decem, pp. 1–6.

[18] M. Wati, N. Novirasari, E. Budiman, and Haeruddin, “Multi-Criteria Decision-Making for Evaluation of Student Academic Performance Based on Objective Weights,” in *2018 Third International Conference*

*on Informatics and Computing (ICIC)*, 2018, pp. 1–5.

[19] H. Zhang and Z. Wang, "A normal distribution-based over-sampling approach to imbalanced data classification," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2011.

[20] S. Wang and X. Yao, "Using class imbalance learning for software defect prediction," *IEEE Transactions on Reliability*, 2013.

[21] F. Gorunescu, "Data mining: Concepts, models and techniques," *Intelligent Systems Reference Library*, 2011.

[22] A. Lawi, S. La Wungo, and S. Manjang, "Identifying irregularity electricity usage of customer behaviors using logistic regression and linear discriminant analysis," in *Proceeding - 2017 3rd International Conference on Science in Information Technology: Theory and Application of IT for Education, Industry and Society in Big Data Era, ICSITech 2017*, 2018.

[23] P. PLN, "Listrik Pintar," *Listrik Pintar*, 2019. [Online]. Available: https://www.pln.co.id/pelanggan/listrik-pintar. [Accessed: 27-Aug-2019].

Implementation of SVM Kernels for Identifying Irregularities Usage of Smart Electric Voucher

ORIGINALITY REPORT

| 20% | 10% | 19% | 14% |
| --- | --- | --- | --- |
| SIMILARITY INDEX | INTERNET SOURCES | PUBLICATIONS | STUDENT PAPERS |

PRIMARY SOURCES

| | | |
| --- | --- | --- |
| 1 | Yuslena Sari, Puguh Budi Prakoso, Andreyan Rezky Baskara. "Road Crack Detection using Support Vector Machine (SVM) and OTSU Algorithm", 2019 6th International Conference on Electric Vehicular Technology (ICEVT), 2019<br>Publication | 3% |
| 2 | Andi Hendra, Gazali Gazali. "Support Vector Machine (SVM) For Toddler's Nutritional Classification in Palu City", INSIST, 2016<br>Publication | 3% |
| 3 | M Irfan, A R Nurhidayat, A Wahana, D S Maylawati, M A Ramdhani. "Comparison of K-Nearest Neighbour and support vector machine for choosing senior high school", Journal of Physics: Conference Series, 2019<br>Publication | 3% |
| 4 | Submitted to Padjadjaran University<br>Student Paper | 2% |
| 5 | Submitted to University of Basrah - College of | |

Science
Student Paper
2%

6 Sinarring Azi Laga, Riyanarto Sarno. "Optimal Sample Temperature of Electronic Nose For Detecting Beef and Pork Mixture", 2019 International Conference on Information and Communications Technology (ICOIACT), 2019
Publication
1%

7 Muhammad Haqqi Ghufran Rifaldi, Erwin Budi Setiawan. "Competence Classification of Twitter Users Using Support Vector Machine (SVM) Method", 2019 7th International Conference on Information and Communication Technology (ICoICT), 2019
Publication
1%

8 "Effect of Service Quality on Customer Satisfaction Prepaid Electricity at PT PLN (Persero) Region S2JB Palembang area, Rayon Mariana", International Journal of Management and Humanities, 2019
Publication
1%

9 Submitted to Universitas Gunadarma
Student Paper
1%

10 Cahya Rahmad, Isna Fauzia Rahmah, Rosa Andrie Asmara, Supriatna Adhisuwignjo. "Indonesian traffic sign detection and recognition using color and texture feature
1%

extraction and SVM classifier", 2018 International Conference on Information and Communications Technology (ICOIACT), 2018
Publication

| | | |
|---|---|---|
| 11 | E Sugiharti, A T Putra, Subhan. "Facial recognition using two-dimensional principal component analysis and k-nearest neighbor: a case analysis of facial images", Journal of Physics: Conference Series, 2020<br>Publication | 1% |
| 12 | A Saifudin, S W H L Hendric, B Soewito, F L Gaol, E Abdurachman, Y Heryadi. "Tackling Imbalanced Class on Cross-Project Defect Prediction Using Ensemble SMOTE", IOP Conference Series: Materials Science and Engineering, 2019<br>Publication | 1% |
| 13 | ijies.sie.telkomuniversity.ac.id<br>Internet Source | 1% |

| | | | |
|---|---|---|---|
| Exclude quotes | On | Exclude matches | < 1% |
| Exclude bibliography | On | | |