



**HAL**  
open science

## **A Smart Algorithm for the Diagnosis of Short-Circuit Faults in a Photovoltaic Generator**

Wail Rezgui, Leila-Hayet Mouss, Nadia Kinza Mouss, Mohamed Djamel Mouss, Mohamed Benbouzid

► **To cite this version:**

Wail Rezgui, Leila-Hayet Mouss, Nadia Kinza Mouss, Mohamed Djamel Mouss, Mohamed Benbouzid. A Smart Algorithm for the Diagnosis of Short-Circuit Faults in a Photovoltaic Generator. IEEE ICGE 2014, Mar 2014, Sfax, Tunisia. pp.139-143. hal-01023491

**HAL Id: hal-01023491**

**<https://hal.science/hal-01023491>**

Submitted on 13 Jul 2014

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# A Smart Algorithm for the Diagnosis of Short-Circuit Faults in a Photovoltaic Generator

Wail Rezgui, Leïla-Hayet Mouss, Nadia Kinza Mouss  
and Mohamed Djamel Mouss  
University of Batna, LAP-lab  
Batna, Algeria  
r-wail@hotmail.fr

Mohamed Benbouzid  
University of Brest, EA 4325 LBMS  
Brest, France  
Mohamed.Benbouzid@univ-brest.fr

**Abstract**—This paper deals with a smart algorithm allowing short-circuit faults detection and diagnosis of PV generators. The proposed algorithm is based on the hybridization of a support vector machines (SVM) technique optimized by a k-NN tool for the classification of observations on the classifier itself or located in its margin.

To test the proposed algorithm performance, a PV generator database containing observations distributed over classes is used for simulation purposes.

**Keywords**—Photovoltaic generator, SVM, k-NN, short-circuit fault, smart classification, linear programming.

## NOMENCLATURE

|          |   |                          |
|----------|---|--------------------------|
| PV       | = | Photovoltaic;            |
| SVM      | = | Support Vector Machines; |
| k-NN     | = | k-Nearest Neighbor;      |
| CO       | = | Class of Observation;    |
| $X$      | = | SVM input vector,        |
| $Y$      | = | SVM output vector,       |
| $f$      | = | Linear function;         |
| $\Phi$   | = | Kernel function;         |
| $w$      | = | Weight vector;           |
| $x$      | = | Problem variable;        |
| $x^*$    | = | New problem variable;    |
| $C$      | = | Class addition;          |
| $\alpha$ | = | Lagrange multipliers;    |
| $c_i$    | = | Class number $i$ ;       |
| $D$      | = | $Y$ diagonal.            |

## I. INTRODUCTION

The performance of a photovoltaic generator is the ratio between the amount of energy supplied by the generator - which is in the form of continuous electrical energy converted by a PV converter to an alternative electrical energy used by consumers - and the solar energy - which is in the form of solar radiation received by the solar panels in the principle of the photovoltaic effect [1].

Unfortunately, this performance is unstable, thanks to a series of factors among them: the cell material, the cell average temperature, the tilt of the solar panel to the sun, and the presence of defects which may cause significant power losses, especially short-circuit faults [2-3]. The PV generator faults diagnosis [4-7] and prognosis [8-12] can stabilize its

performance and ensure its availability and reliability. In this context, the paper objective is the development of an algorithm for the fault detection and diagnosis for a photovoltaic generator. Indeed, the paper contributions are twofold: 1) Development of a fault detection and localization algorithm; in particular short-circuits in a cell, bypass, and blocking diodes. 2) Development of a smart classifier based on observation collected from the control system. It is based on the support vector machines (SVM) and k-NN tools [13-22].

## II. CLASSICAL DIAGNOSIS ALGORITHM

The proposed diagnosis algorithm is designed for the detection and the localization of short-circuits in a PV generator cells, bypass, and blocking diodes. This is a critical fault as it creates hot spots in addition to the voltage degradation.

The diagnosis algorithm consists of four main hierarchical steps.

### A. Step I

If a PV generator voltage is less than the one provided by a healthy generator, and its current is equal or greater than the current supplied by a healthy generator; this therefore means that there exist some short-circuited components. A zoom must be carried-out on the generator components characterization to check string states.

### B. Step II

This step consists in the characterization of a string state (short-circuit). In this context, there is numerous condition dealing for a short-circuit detection. The main one is to consider a string short-circuited when its voltage is zero. Another specific fault is the blocking diode short-circuit that will deteriorate the PV generation. This fault is particularly difficult to detect.

### C. Step III

This step consists in the characterization of a module state. A module is considered as short-circuited if its voltage is zero and its current is maximal. This situation means that all its

groups are short-circuited. Otherwise, it is an indication that the module contains at least one short-circuited group.

#### D. Step IV

This step consists in the characterization of a group state. A group is considered as short-circuited if its voltage is zero and its current is equal to the short-circuit value. In this case, all the components are considered as short-circuited. Otherwise, if the group voltage is strictly greater than zero and strictly less than its normal operation value, and its current is equal to the sum of its cells currents, then this group contains at least one good cell and the others are short-circuited. Moreover, when the group voltage is zero and its current is equal to the sum of the cells short-circuit current and the bypass current, this situation means that all the group cells are defectives. Finally, if the group voltage is zero and its current is equal to the sum of the healthy cells currents and the current through the bypass diode, this last situation means that this group is connected by a short-circuited bypass diode.

### III. INTELLIGENT DIAGNOSIS ALGORITHM

#### A. SVM Algorithm

Support vector machines are supervised learning techniques, for solving classification problems in many fields such as pattern recognition, text categorization, or medical diagnosis [23-26].

The SVM technique highlight compared to traditional learning ones is that it does not depend on the data representation space dimension. However, the choice of the kernel function adapted to the problem, or calculation time based on the data number to be treated, can make the use of this tool more complicated. The analysis and study of this method brings out two major drawbacks for its use which are: 1) Binary classification, which means that it cannot handle more than two classes at the same time; 2) Support vectors classification. In other words, observations located at the margin and especially those located on the separator itself are not well-determined.

In this particular restrictive context, our objective is to find a solution to the problem of multi-class SVM classifier. For that purpose, we propose to treat the problem by solving a set of equations, each presenting a classifier between a class and its complement. If the observation does not belong to this class, we repeat the separation. The following SVM algorithm summarizes our proposal.

1) **Step 1:** *Construct the SVM classifier.* Our objective is to construct a function  $f$  which, for each input value  $x$  in a set  $\mathbb{R}^d$ , will match an output value  $y \in \{-1, 1\}$ . The goal is to try to learn  $f$  from a set of pairs  $(x_i, y_i)$ . In the linear case, a discriminate function is obtained by a linear combination of the input vector  $x$ , allowing a quick access to our objective. In a nonlinear case, a change in the data space becomes mandatory with a linear separation in a new space of a larger

dimension. In this case, the use of an implicit function  $\Phi$  becomes necessary.

2) **Step 2:** *Determining the hyper-plane.* There are many separating hyper-planes; the best is the one that maximizes the margin. This one is defined as the distance between a hyper-plane and the closest sample points that are called support vectors. The maximum margin hyper-plane, which is intended to classify any new observation, is given by a shape resulting from linear programming.

SVM classifier second drawback is observations classification: for support vectors and data situated in the margin, side classification error, especially if the margin is not wide which is based on the distribution of data space, and also side classification of observations located on the classifier himself. This paper proposes then a solution for the above-described SVM disadvantage. Indeed, it is proposed the use of the well-known tool in classification; the so-called k-NN.

#### B. k-NN Method

k-NN is one of the unsupervised learning algorithms that solve the problem of classifications used in many areas such as diagnostic. This algorithm consists in three main steps. In this work, it is proposed mathematical modeling by matrices for flexibility and ease to use purposes [27-29].

1) **Step 1.** This first step aims to determine a representative for each observation  $X = \{x_1, x_2, \dots, x_N\}$  in the learning space  $X: n \times m$ . There is a lot of works concerning this step. In particular, calculating the gravity centers for each observation and therefore deriving the resulting vector center  $(x)$  is the most used approach.

2) **Step 2.** This step concerns the observation classification. It also needs a representative by calculating its gravity center, to obtain a vector center  $(x^*)$ .

3) **Step 3.** This last step concerns the determination of the minimum value index class of the Manhattan type Euclidean distance between  $(x^*)$  and  $(x)$  as given by (1).

Even if the above described method has several advantages as robustness and efficiency, it has, as traditional classification methods, some disadvantages. Among them, two main problems: remotely and ambiguity discharges. This work is not concerned by the first problem as it is already classified by the SVM. However, in the second situation, there are several minimum distances. In other words, there are more than one distance between the new observation and the old ones. In this case, we have to check if the old and new observations belong to the same class; otherwise, the new observation belongs to the class of ambiguity discharges.

#### C. The Proposed Smart Algorithm

In this paper, it is particularly proposed a smart algorithm allowing smart classification of observations retrieved from the studied PV generator operation.

This algorithm is considered smart as it results from the hybridization of the SVM technique (artificial intelligence family) optimized by the k-NN tool.

$$f(x^*) = \begin{cases} \text{If the problem is linear separable} \\ \left[ \begin{array}{c} \text{sign} \left( \alpha_{c_1, \delta(c_1)}^T X x^* - \frac{1}{2} \left( \max_{y_i=-1} (\langle w_{c_1, \delta(c_1)}, x_i \rangle) + \min_{y_i=1} (\langle w_{c_1, \delta(c_1)}, x_i \rangle) \right) \right) \\ \vdots \\ \text{sign} \left( \alpha_{c_{m-1}, c_m}^T X x^* - \frac{1}{2} \left( \max_{y_i=-1} (\langle w_{c_{m-1}, c_m}, x_i \rangle) + \min_{y_i=1} (\langle w_{c_{m-1}, c_m}, x_i \rangle) \right) \right) \end{array} \right] \\ (1-\xi) \text{ Else} \\ \left[ \begin{array}{c} \text{sign} \left( K(x^*, X^T) D \alpha_{c_1, \delta(c_1)} - \frac{1}{2} \left( \max_{y_i=-1} (\langle w_{c_1, \delta(c_1)}, \phi(x_i) \rangle) + \min_{y_i=1} (\langle w_{c_1, \delta(c_1)}, \phi(x_i) \rangle) \right) \right) \\ \vdots \\ \text{sign} \left( K(x^*, X^T) D \alpha_{c_{m-1}, c_m} - \frac{1}{2} \left( \max_{y_i=-1} (\langle w_{c_{m-1}, c_m}, \phi(x_i) \rangle) + \min_{y_i=1} (\langle w_{c_{m-1}, c_m}, \phi(x_i) \rangle) \right) \right) \end{array} \right] \\ \text{End} \end{cases} + \xi * CO \left( \text{index} \min \left[ \begin{array}{c} \frac{1}{N} \sum_{j=1}^N |x_{j1} - x_j^*| \\ \frac{1}{N} \sum_{j=1}^N |x_{j2} - x_j^*| \\ \vdots \\ \frac{1}{N} \sum_{j=1}^N |x_{jM} - x_j^*| \end{array} \right] \right) \quad (1)$$

Where  $\xi = 1$  if  $x$  is in the margin of the SVM classifier, else  $\xi = 0$ .

This tool is particularly used to increase the classification rate against observations on the classifier itself or located in the margin which is not wide.

#### IV. SIMULATIONS RESULTS

To illustrate the performance of the proposed fault detection and diagnosis algorithms several simulations have been carried-out on a typical PV generator.

##### A. Faulted PV Generator Characterization

The main simulation results are shown by Fig. 1 to 3. Figure 1 clearly illustrates the PV generator power performance under short-circuited cells. In particular, this figure shows that the power decreases proportionally to the number of short-circuited cells (the current is independent of the number of defective cells).

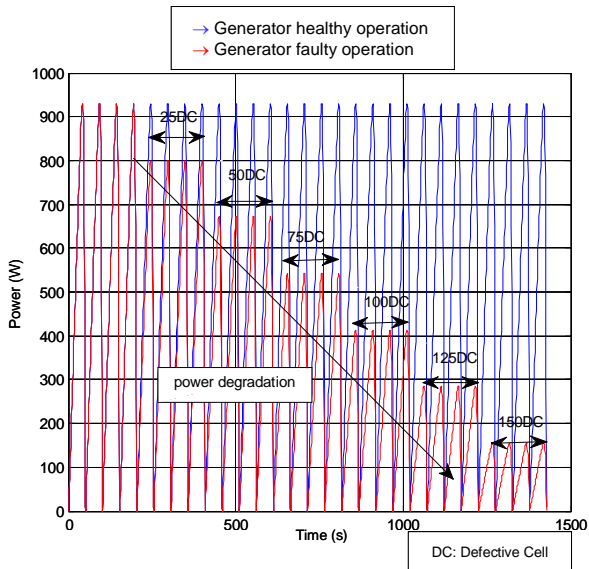


Fig. 1. Short-circuited cells influence on the PV generator operation.

Figure 2 shows the influence of the bypass diodes short-circuits on the generated power. These diode short-circuits have a greater impact on the PV generator performance. Indeed, bypass diodes short-circuits affect a group voltage. Again, the current remains independent of fault type unless all string groups are failed.

Finally, Fig. 3 illustrates the influence of blocking diode short-circuits on the PV generator operation. It is clearly shown that when this type of fault occurs, the generated power clearly exhibits a big drop. Indeed, the faulty strings will consume the power generated by the other healthy strings.

##### B. Smart Algorithm Tests

The proposed smart fault detection and diagnosis algorithm is tested using a PV generator database containing observations distributed over classes.

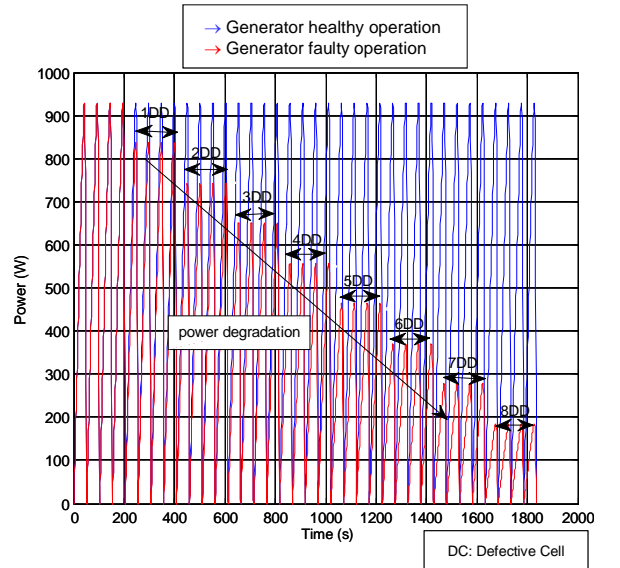


Fig. 2. Short-circuited bypass diode influence on the PV generator operation.

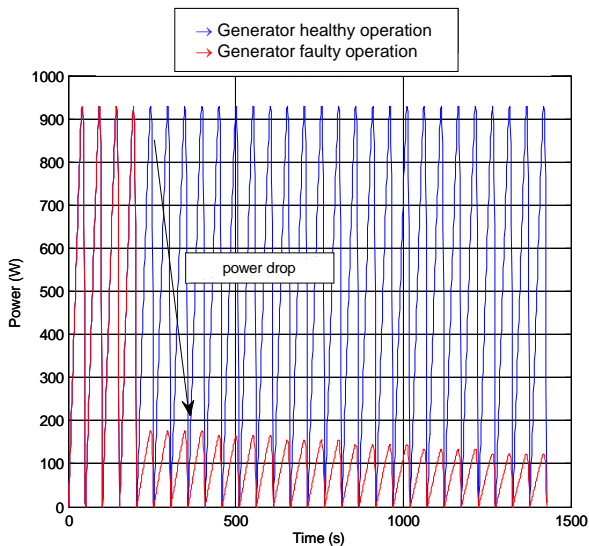


Fig. 3. Short-circuited blocking diode influence on the PV generator operation.

For that purposes, three indicators are used: the rate of good classified observations (Fig. 4), the classification error rate (Fig. 5), and the computation time to classify any new observation (Fig. 6). In this context, 60 samples are selected, each containing 1632 observations. For the classification of each observation of each sample, three classifiers are used for comparison purposes: SVM, k-NN, and the proposed SVM optimized by k-NN.

With the SVM classifier, the achieved results show that the rate of good classified observations is between 60 to 69.9% (Fig. 4), with a classification error rate between 0.8 to 0.9% (Fig. 5), and its computation time is between 5.5 to 7 time units (Fig. 6).

With k-NN classifier, the achieved results show that the rate of good classified observations is between 50 to 53.5% (Fig. 4), with a classification error rate between 1.5 to 2% (Fig. 5), and its computation time is between 2 to 3 time units (Fig. 6).

With the proposed classifier (hybrid SVM\_k-NN), the achieved results show that the rate of good classified observations is between 68 to 75.8% (Fig. 4), with a classification error rate between 0.36 to 0.55% (Fig. 5), and its computation time is between 5 to 10 time units (Fig. 6).

The analysis of the above achieved results shows that the proposed hybrid classifier has the following specific features: a high classification rate with a low error rate but it is a little bit time consuming due the mathematical computations.

## V. CONCLUSION

This paper dealt with a smart algorithm allowing short-circuits detection and diagnosis in PV generators. The proposed algorithm is considered smart as it results from the hybridization of the SVM technique (artificial intelligence family) optimized by the k-NN tool that was used to increase the classification rate against observations on the classifier itself or located in the margin which is not wide.

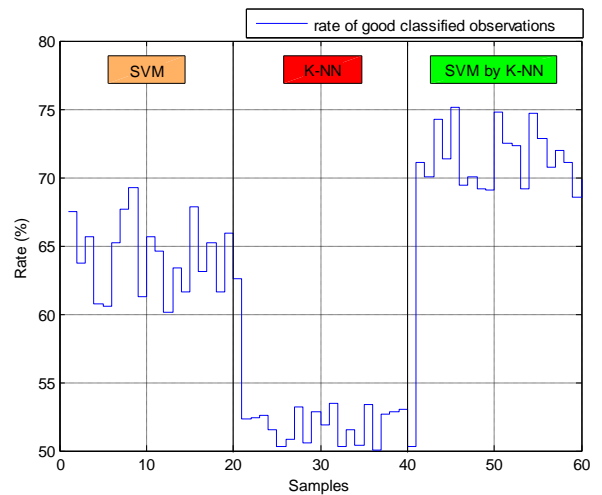


Fig. 4. Observation classification rate vs tools.

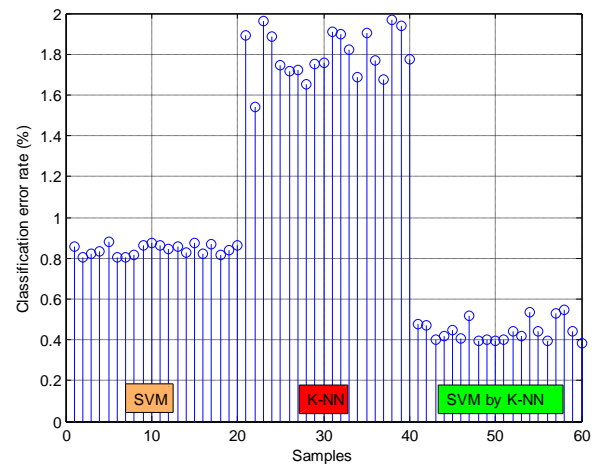


Fig. 5. Observation classification error rate vs tools.

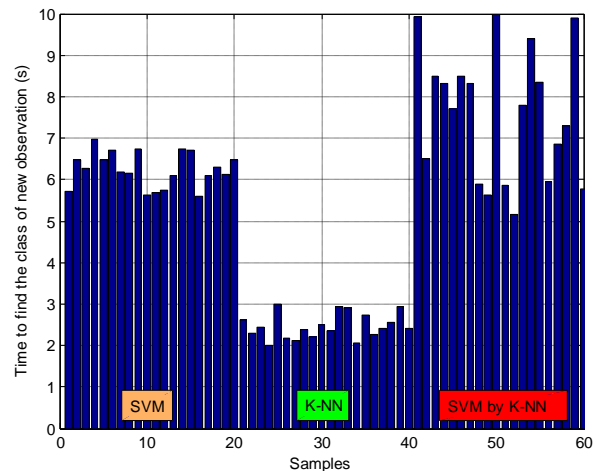


Fig. 6. New observation classification elapsed time vs tools.

The proposed smart fault detection and diagnosis algorithm was tested using a PV generator database containing observations distributed over classes.

The analysis of the achieved results shows that the proposed hybrid classifier has the following specific features: a high classification rate with a low error rate but it is a little bit time consuming due the mathematical computations.

## REFERENCES

- [1] K. Ardani K, G. Barbose, R. Margolis and D. Feldman, "Quantifying non-hardware balance of system costs for photovoltaic installations in the United States using a combined annual expenditure-labor hour productivity approach," in *Proceedings of the 2012 IEEE PVSC*, Austin (USA), pp. 1762-1767, June 2012.
- [2] O. Gana, C. Vasar, M. Babescu and A. Vartosu, "PV control system based on short-circuit current," in *Proceedings of the 2013 IEEE SACI*, Timisoara (Romania), pp. 387-390, May 2013.
- [3] S. Phuttapattimok, A. Sangswang and K. Kirtikara, "Effects on short circuit level of PV grid-connected systems under unintentional islanding," in *Proceedings of the 2008 IEEE ICSET*, Singapore, pp. 928-932, November 2008.
- [4] Y. Hirata, S. Noro, T. Aoki and S. Miyazawa, "Diagnosis photovoltaic failure by simple function method to acquire I-V curve of photovoltaic modules string," in *Proceedings of the 2012 IEEE PVSC*, Austin (USA), pp. 1340-1343, June 2012.
- [5] M. Gonzalez, B. Raison, S. Bacha and L. Bun, "Fault diagnosis in a grid-connected photovoltaic system by applying a signal approach," in *Proceedings of the 2011 IEEE IECON*, Melbourne (Australia), pp. 1354-1359, November 2011.
- [6] T. Jianeng, Z. Yongqiang and W. Wenshan, "Fault diagnosis method and simulation analysis for photovoltaic array," in *Proceedings of the 2011 IEEE ICECE*, Yichang (China), pp. 1569-1573, September 2011.
- [7] X. Xu, H. Wang and Y. Zuo, "Method for diagnosing photovoltaic array fault in solar photovoltaic system," in *Proceeding of the 2011 IEEE APPEEC*, Wuhan (China), pp. 1-5, March 2011.
- [8] B.E. Olivares, M.A. Cerda Munoz, M.E. Orchard and J.F. Silva, "Particle-filtering-based prognosis framework for energy storage devices with a statistical characterization of state-of-health regeneration phenomena," *IEEE Trans. Instrumentation and Measurement*, vol. 62, n°2, pp. 364-376, February 2013.
- [9] Y. Wang, M. Simon, P. Bonde, B.U. Harris, J.J. Teuteberg, R.L. Kormos and J.F. Antaki, "Prognosis of right ventricular failure in patients with left ventricular assist device based on decision tree with SMOTE," *IEEE Trans. Information Technology in Biomedicine*, vol. 16, n°3, pp. 383-390, May 2012.
- [10] J. Son, Q. Zhou, S. Zhou, X. Mao and M. Salman, "Evaluation and comparison of mixed effects model based prognosis for hard failure," *IEEE Trans. Reliability*, vol. 62, n°2, pp. 379-394, June 2013.
- [11] D.M. Sow, J. Sun, A. Biem, J. Hu, M.L. Blount and S. Ebadollahi, "Real-time analysis for short-term prognosis in intensive care," *IBM Journal of Research and Development*, vol. 56, n°5, pp. 3:1-3:10, September-October 2012.
- [12] P. Arpaia, C. Manna, G. Montenero and G. D'Addio, "In-time prognosis based on swarm intelligence for home-care monitoring: A case study on pulmonary disease," *IEEE Sensors Journal*, vol. 12, n°3, pp. 692-698, March 2012.
- [13] Z. Huang and M.L. Shyu, "k-NN based LS-SVM framework for long-term time series prediction," in *Proceedings of the 2010 IEEE IRI*, Las Vegas (USA), pp. 69-74, August 2010.
- [14] X. Shi, Y. Zhao and X. Dong, "Web page categorization based on k-NN and SVM hybrid pattern recognition algorithm," in *Proceedings of the 2008 IEEE FSKD*, Shandong (China), vol. 2, pp. 523-527, October 2008.
- [15] P. Yuan, Y. Chen, H. Jin and L. Huang, "MSVM-kNN: Combining SVM and k-NN for multi-class text classification," in *Proceedings of the 2008 IEEE WSCS*, Huangshan (China), pp. 133-140, July 2008.
- [16] D. Han, C. Han, Y. Yang, Y. Liu and W. Mao, "Pre-extracting method for SVM classification based on the non-parametric K-NN rule," in *Proceedings of the 2008 IEEE ICPR*, Tampa (USA), pp. 1-4, December 2008.
- [17] X. Han and R. Ruonan, "The method of medical named entity recognition based on semantic model and improved SVM-KNN Algorithm," in *Proceedings of the 2011 IEEE SKG*, Beijing (China), pp. 21-27, October 2011.
- [18] K.M. Thilina, K.W. Choi, N. Saquib and E. Hossain, "Pattern classification techniques for cooperative spectrum sensing in cognitive radio networks: SVM and W-KNN approaches," in *Proceedings of the 2012 IEEE GLOBECOM*, Anaheim (USA), pp. 1260-1265, December 2012.
- [19] K. Gayathri and A. Marimuthu, "Text document pre-processing with the KNN for classification using the SVM," in *Proceedings of the 2013 IEEE ISCO*, Tamil Nadu (India), pp. 453-457, January 2013.
- [20] C. Zanchettin, B.L.D Bezerra and W.W. Azevedo, "A KNN-SVM hybrid model for cursive handwriting recognition," in *Proceedings of the 2012 IEEE IJCNN*, Brisbane (Australia), pp. 1-8, June 2012.
- [21] T. Li, "On KNN and SVM text classification technology in knowledge management," in *Proceedings of the 2011 IEEE EMEIT*, Harbin (China), vol. 8, pp. 3923-3926, August 2011.
- [22] T.A. Gulyaeva, D.Y. Korotenko and A.A. Popov, "Application of the hidden Markov models, kNN and SVM for a classification problem modes of power supply system," in *Proceedings of the 2012 IEEE APEIE*, Novosibirsk (Russia), pp. 36-40, October 2012.
- [23] Y. Mao and D. Wei, "Weight balanced linear programming SVM on skewed distribution and its evaluation," in *Proceedings of the 2009 IEEE ICIS*, Shanghai (China), pp. 430-435, June 2009.
- [24] F. Cai and V. Cherkassky, "Generalized SMO algorithm for SVM-based multitask learning," *IEEE Trans. Neural Networks and Learning Systems*, vol. 23, n°6, pp. 997-1003, June 2012.
- [25] Z. Xin, H. Hailong, L. Haien and L. Jianwei, "P-norm regularized SVM classifier by non-convex conjugate gradient algorithm," in *Proceedings of the 2013 IEEE CCDC*, Guiyang (China), pp. 2685-2690, May 2013.
- [26] Q. Hua, D. Liduo, S. Xin and M. Zuqiang, "The optimization of large-scale SVM using nest template ant clustering in kernel space," in *Proceedings of the 2012 IEEE CECNet*, Yichang (China), pp. 682-686, April 2012.
- [27] R. Isola, R. Carvalho and A.K. Tripathy, "Knowledge discovery in medical systems using differential diagnosis, LAMSTAR and k-NN," *IEEE Trans. Information Technology in Biomedicine*, vol. 16, n°6, pp. 1287-1295, November 2012.
- [28] H. Kucuk, C. Tepe and I. Eminoglu, "Classification of EMG signals by k-nearest neighbor algorithm and support vector machine methods," in *Proceedings of the 2013 IEEE SIU*, Haspolat (Turkey), pp. 1-4, April 2013.
- [29] K. Ray, P. Rajamani and A. Mukherjee, "k-nearest neighbor algorithm based classification and localization of seven different types of disc-to-disc impulse insulation failures in power transformer," in *Proceedings of the 2012 IEEE ICPADM*, Bangalore (India), pp. 1-5, July 2012.