

# Support Vector Machines Versus Multi-Layer Perceptrons for Reducing False Alarms in Intensive Care Units

BEN REJAB Fahmi  
BESTMOD, High Institute of  
Management,  
41, City Bouchoucha, 2000  
Bardo, Tunis, Tunisia

NOUIRA Kaouther  
BESTMOD, High Institute of  
Management,  
41, City Bouchoucha, 2000  
Bardo, Tunis, Tunisia

ABDELWAHED Trabelsi  
BESTMOD, High Institute of  
Management,  
41, City Bouchoucha, 2000  
Bardo, Tunis, Tunisia

## ABSTRACT

This paper presents a comparative study between two well-known classification techniques in the machine learning area namely the Multi-Layers Perceptrons (MLP) and the Support Vector Machines (SVM) applied in the medicine field. More precisely, our aim in this paper is to reduce the rate of false alarms in the Intensive Care Units (ICU) using the SVM and the MLP techniques. To this end, we have performed an appropriate setting of parameters for both SVM and MLP techniques to guarantee the good monitoring of patients' states. Then, we have made a comparison between the adapted classification techniques i.e. the SVM and the MLP and the current system using different evaluation criteria. Results of comparative experiments show that the true alarms can be identified with high accuracy by the SVM technique. Compared with the MLP and the current system, the SVM technique shows its potential to reduce the rate of false alarms.

## General Terms

Reducing false alarms in ICU using SVM technique.

## Keywords

Machine Learning, Support Vector Machines, Multi-Layers Perceptrons, Classification, Intensive Care Units, Monitoring.

## 1. INTRODUCTION

Monitoring system are important tools in ICU. Their role is to monitor dynamics of a large number of variables relevant to patient such as physiological variables, laboratory data, device parameters, etc. [1].

The function of monitoring system is also to report physiological signal that can be measured. Indeed, ICU monitors are based on fixed thresholds in order to determine the current state of the patient. The exceeding of thresholds, even for a particular variable, means that the patient is in critical state. But, most of existing alarm systems used in ICU produce a large percentage of false alarms; it is about 86% [2] [3]. False alarms are due to measurement artifacts like patient movements [4] or transient fluctuations [5]. Moreover, according to Donchin and Seagul [6], such alarms engender difficult working conditions in the ICU and this defect causes medical errors, resulting in thousands of deaths each year. Unfortunately, the percentage of false alarms by devices can rise to 90% [7] and it means the presence of frequent but insignificant alarms.

In order to overcome this problem in ICU, many works have been done in this field. We can mention online trend extraction methodology [8], digital signal processing [9], artificial intelligence [10], etc. Moreover, an overview of alarm algorithms in critical care monitoring has been detailed in [11]. Despite these different works, the rate of false alarms is still high in ICU.

In order to avoid this issue, we will consider it as a classification problem to easily differentiate between normal and critical states of patient. As a result, we will apply two well-known classification techniques that have been successfully applied in several fields namely the Multi-Layers Perceptrons (MLP) and the Support Vector Machines (SVM). The structure of this paper is as follows; Section 2 and Section 3 review the applied techniques namely the SVM and the MLP. Section 4 illustrates the monitoring system in ICU with the proposed classification techniques i.e. SVM and MLP. The used databases and the experimental results are described along with detail in Section 5.

## 2. SUPPORT VECTOR MACHINES

Support vector machines [12] is considered as a fundamental supervised machine learning technique [13] that has attracted a lot of researchers [14][15]. Moreover, the SVM is considered among the best classifiers, it is successfully used in many fields such as medical image processing, feature selection for cancer classification [16], protein classification [17], etc.

Generally, learning algorithms are unable to solve problems when we use a huge database. However, SVM can find an adequate solution to this issue by using the notion of support vectors which represents the nearest features from hyperplane. Therefore, the complexity of an SVM model is unaffected by the number of features encountered in the training data since the number of support vectors selected by the learning algorithm of SVM is usually small. For this reason, SVM is well suited to deal with learning tasks where the number of features is large with respect to the number of training instances.

Besides, the SVM technique is characterized by its ability to handle linearly and non-linearly separable data.

### 2.1 Case of linearly separable data

If the training data is linearly separable, then a pair  $(w, b)$  exists such that:

$$w^T \cdot x_i + b \geq 1, \text{ for all } x_i \in P.$$

$$w^T \cdot x_i + b \leq -1, \text{ for all } x_i \in N.$$

with the decision rule given by:

$$f_{w,b}(x) = \text{sign}(w^T \cdot x_i + b).$$

Where  $w$  is termed the weight vector,  $b$  the bias (or  $-b$  is termed the threshold),  $x_i$  is an observation and  $P$  and  $N$  present respectively positive and negative data. When it is possible to linearly separate two classes, an optimum separating hyperplane can be found by minimizing the squared norm of the separating hyperplane. The minimization can be set up as a convex quadratic programming (QP) problem:

$$\begin{cases} \min \frac{1}{2} \|w\|^2, w \in R^d, b \in R, \\ \text{subject to} \\ y_i(w \cdot x_i + b) \geq 1, \\ \text{for } i = 1, \dots, m. \end{cases}$$

with  $y_i$  the class of the observation  $x_i$ ,  $m$  the number of observations and  $R^d$  the dimension number.

Figure 1 shows the optimal hyperplane relative to the decision function.

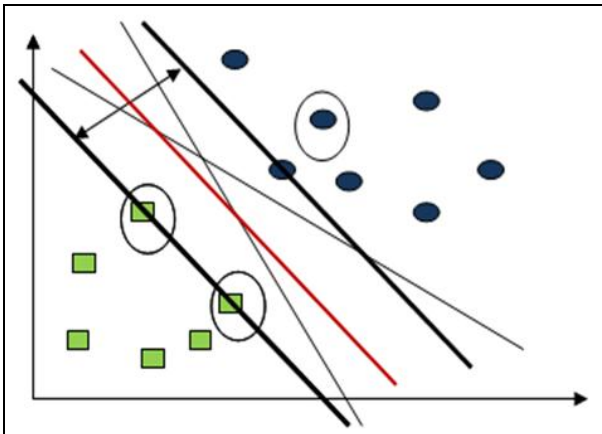


Fig 1: Optimal hyperplane [18]

## 2.2 Case of non-linearly separable data

Generally, in real problems, data is not linearly separable and there is no separation between positive and negative instances in the training set. The solution given to this problem is to map the data into a higher dimensional space and define a separating hyperplane there.

This high-dimensional space is called the transformed feature space, as opposed to the input space occupied by the training instances.

As a result, a linear separation is obtained in the transformed features space which can contain an infinite dimension described in Figure 2.

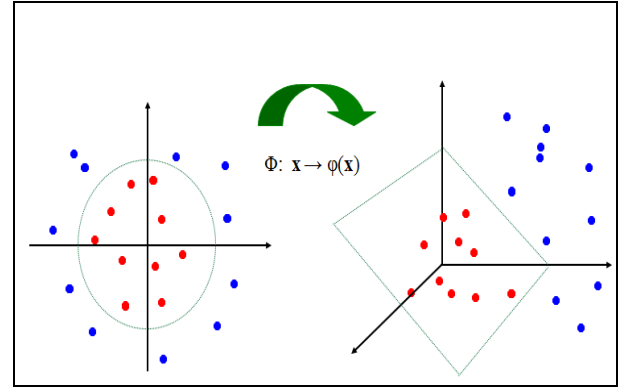


Fig 2: Mapping data

This mapping is assured using the function  $\Phi$  defined such as  $R^d \rightarrow R^D, (D \gg d)$ .

with  $R^d$  presents HILBERT space.

However, it is not necessarily to determinate  $\Phi$  since the kernel function  $K$  is only used in the training algorithm. The kernel function is defined by  $K(x_i, x_j) = \Phi(x_i) \cdot \Phi(x_j)$ .

Within the most common kernel functions, we can cite [12]:

- Linearly:

$$K(x, y) = x \cdot y.$$

- Polynomial of degree:

$$K(x, y) = [(x \cdot y) + 1]^d.$$

Radial Basis Function (RBF):

$$K(x, y) = \exp\left(-\frac{\|x-y\|^2}{2\sigma^2}\right).$$

Using the Lagrangian formulation, these optimization problems are solved by introducing a new unknown scalar variable  $\alpha_i$  called the Lagrange multiplier.  $\alpha_i$  is introduced for each constraint and forms a linear combination involving the multipliers as coefficients. The problem will be expressed as follows:

$$f(x) = \sum_{i=1}^{n_s} \alpha_i y_i k(x_i, x) + b.$$

The optimization problem using a penalty term which penalizes slack variables with high values is defined as follows:

$$\begin{cases} \min \frac{1}{2} \|w\|^2 + C \sum_{i=1}^m \xi_i, \\ \text{subject to} \\ y_i(w \cdot x_i + b) \geq 1 - \xi_i, \\ \xi_i \geq 0, \text{ for } i = 1, \dots, m, \end{cases}$$

### 3. MULTI-LAYER PERCEPTRON

Multi-layer perceptrons (MLP) is a linear classifier belonging to the Neuron Networks family. This classifier is composed of several layers and where the input layer communicates with the output layer by sending to it information; it consists of a feed-forward network type [19].

Figure 3 shows the feed-forward process in the MLP.

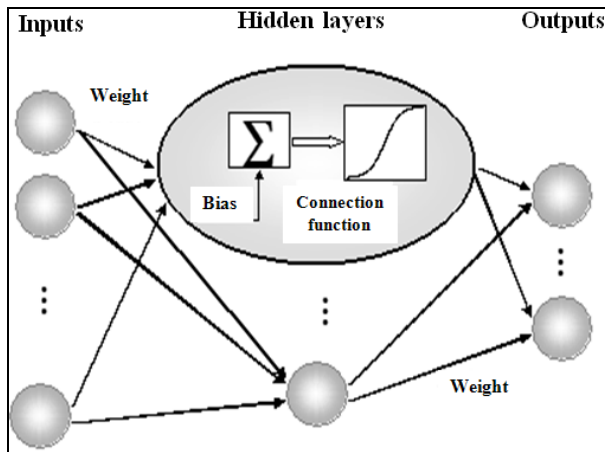


Fig 3: The MLP architecture

Each layer in the MLP classifier is formed of a different number of neurons and the system output corresponds always to the number of the neurons in the output layer.

This limitation has been avoided only through the gradient back propagation of the error, in 1984 by Paul J Werbos, by David Rumelhart in 1986 secondly.

In the MLP method based on the back propagation process neurons relative to a layer are connected to the most adjacent layers' neurons.

The weight of each connection is a fundamental key for the good function of the network. As a result, the establishment of a multi-layer perceptrons consists of the setting of the best weights adapted to the different inter-neuronal connections. The weights are fixed using the back propagation algorithm which is illustrated as follows:

1. Present a pattern to the network.
2. Compare the network output to the target output.
3. Compute the output error relative to each neuron in the network.
4. Calculate the correct output value for each neuron that is supposed be obtained.
5. Set the difference between the network output and the correct output value i.e. precise the local error.
6. Adjust the weight of each connection to the lowest local error.
7. Assign "error signal" to all previous neurons.
8. Repeat these steps from (4) to (8) by applying them to the previous neurons and by using "error signal" as error.

### 4. MONIYORING SYSTEM USING MLP AND SVM

System in ICU is based on the setting of thresholds for detecting critical state of patient. In fact, the medical staff modifies these thresholds for the different measured variables relative to the different patients. On the current system, when the values of parameters are beyond the limits, already fixed, an alarm is trigged.

However these numerous alarms do not necessarily translate a critical state of patient since, there is no relation between the measured variables when thresholds are fixed. In fact, it is not sufficient that a particular variable reaches its limits to trigger alarms because it can be caused by the disconnection of the sensor relative to the patient for example.

Besides, the current system used in ICU is no more a real monitoring system but, an ordinary measuring tool.

Our aim is to resolve this problem by applying two classification techniques in the medical field. More precisely, we propose to use the SVM and the MLP methods in the ICU in order to reduce the rate of false alarms and to detect in a significant way the true alarms. As a result, the application of SVM and MLP techniques will guarantee the improvement of the work conditions of the medical staff and the good monitoring of patients' states.

As we propose to reduce the rate of false alarms in ICU using classification approaches, it is primary to describe the structure of the system when we apply the machine learning techniques namely the MLP and the SVM. Figure 4 shows the system structure.

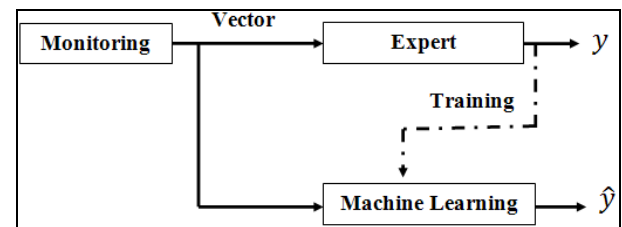


Fig 4: Structure of the proposed system

A system that learns from examples is composed of three main modules [20] mainly:

1. A generator which provides random data called input vectors. In our proposed system these vectors present the observations obtained from the monitoring.
2. A supervisor which associates to each input vector  $x$  an output  $y$  presenting the class. In our case, the expert gives his opinion about the trigged alarms relative to each database and precise if these alarms are correct or not.
3. A machine learning which finds the most adequate function that associates to each input vector  $x$  an output  $y$ . In this work, we propose to apply the SVM and MLP techniques.

In order to successfully apply the SVM and the MLP techniques, we should follow these steps described in detail as follows:

1. Setting of parameters: An appropriate setting of parameters for both of the SVM and MLP approaches can improve the classification accuracy. In order to choose the best parameters to our system, we have applied the grid search approach proposed by Chen and Lin in [21] and based on cross-validation.
2. Division of data set on training and test sets: We have used the stratified sampling and we have divided the data set into two sets. The first one contains 70% of the total number of the observations and corresponding to the training set. However, 30% of the total number of the observations is relative to the test set.

3. Building the new system from training data: After setting the best parameters and fixing the learning sample, we build the model of SVM or MLP.
4. Testing the proposed system using test set: To evaluate our system and prove its effectiveness, we make several tests using test set.

## 5. EXPERIMENTS

### 5.1 Data description

It is essential to test and evaluate the monitoring system after applying the SVM and the MLP techniques. To this end, we have used MIMICII (Multiparameter Intelligent Monitoring for Intensive Care) database taken from Physiobank [22].

This database contains data from hemodynamically unstable patients hospitalized in 1996 in ICU of the cardiology division in the Teaching Hospital of Harvard Medical School. It includes 100 patients' records of continuous data recorded each second. Each recording is accompanied with detailed annotations or labels made by an expert in order to precise if the current state of the patient is critical or not.

To assure the training and the testing of the SVM and the MLP methods, we used 14 databases. Each one evaluates the state of a particular patient. For each patient, several variables are measured such as HR (Heart Rate), SpO2 (Oxygen Saturation), Non-Invasive Blood Pressure (NBP), Respiratory rate (Resp), ABP (Artery Blood Pressure), Pulmonary Artery Pressure (PAP) [10]. Table 1 gives more detail for the used databases with #Attributes presents the total number of measured parameters and #Instances presents the values of each measured variable at a moment  $t$ .

**Table 1. Description of the used data sets**

Databases	Attributes	Instances
Patient 01	6	4101
Patient 02	8	42188
Patient 03	8	42188
Patient 04	7	42188
Patient 05	9	42188
Patient 06	9	5350
Patient 07	7	11300
Patient 08	7	10600
Patient 09	12	5700
Patient 10	5	42188
Patient 11	7	42188
Patient 12	7	42188
Patient 13	9	42188
Patient 14	7	42188

### 5.2 Evaluation criteria

Our aim is to evaluate the new monitoring system and compare its performance to the existing one.

In order to evaluate and compare the MLP and the SVM techniques to the current system used in ICU, we have used three evaluation criteria described as follows:

- The rate of false alarms expressed by the error rate(ER) as follows:

$$ER = \frac{(FP + FN)}{(TP + TN + FP + FN)}$$

With FP, FN, TP and TN present respectively False Positive, False Negative, True Positive and True Negative alarms.

The rate of false alarms is considered as a very important criterion allowing us to know how efficient are the SVM and the MLP techniques compared to the current system.

- The sensitivity (S) of the system presents the ratio of TP to all significant alarms which are declared by experts as critical and they are correctly identified by the system:

$$S = \frac{TP}{TP + FN}$$

- The number of true alarms illustrated by the ability of the SVM to detect the true alarms

### 5.3 Experimental results

In this section, we apply the SVM and the MLP, and then we show the performance of SVM and MLP models. After that, we compare them to the current monitoring system.

In order to evaluate the improved system that uses the SVM and the MLP, we used the free software Tanagra [23].

#### Step 1: Setting model parameters.

The most important step before building the SVM or the MLP model is the setting parameters.

For the SVM Model: It consists of the setting of the model cost and the kernel parameters namely  $\gamma$  (when using RBF) and  $\sigma$  (when using sigmoid kernel). To find the best parameters we used the grid search approach [21].

To this end, we used the function tune.svm from the e1071 package developed with R in these 14 patient databases. The obtained parameters are described in Table 2.

**Table 2. Best parameters of SVM**

Kernel function	RBF
Cost C	1000
Gamma $\gamma$	600

For the MLP Model: we apply the same approach used in the SVM model [21] to find the right parameters namely:

- The number of hidden layers.
- The neuron number in the hidden layers.
- The iteration number.

So we used the function tune.nnet from the same package and we obtained the best combination of parameters minimizing the rate of training and testing detailed in Table 3.

**Table 3. Best parameters of MLP**

Number of hidden layers	2
Iteration number	100
Neuron number in hidden layers	10

**Step 2: Loading and stratified sampling (training and test sets).**

We have divided each database into two sets: training set and test set. This division is assured by the stratified sampling which allows the generation of the correct classification model. We use for the training 70\% of the whole database.

**Table 4. Results of stratified sampling**

Parameters	
Sampling proportion	70%
Stratification attribute	class
Results	
Sample	Size
Input	4101
Output	2870

Table 4 illustrates the output result of stratified sampling relative to the first database using the software Tanagra. As indicated on this table (Table 4) 70% of observations (i.e. 2870 observations) are used for the training however, 30% of observations (i.e. 1231 observations) are for the test.

**Step 3: Learning with training data (building model).**

Table 5 presents the confusion matrix of the training phase relative to the SVM and the MLP techniques.

**Table 5. Results of the training phase of SVM and MLP**

	SVM			MLP		
Error rate	0.0003			0.11		
	Confusion matrix			Confusion matrix		
	No	Yes	Sum	No	Yes	Sum
No	2654	0	2654	2377	277	2654
Yes	1	215	216	19	197	216
Sum	2655	215	2870	2396	474	2870

From Table 5 we can notice that the SVM has an important ability to learn compared to the MLP. The SVM has detected 215 true alarms from 216. However, the MLP technique has detected 474 alarms where 277 are false. It has also missed 19 true alarms.

**Step 4: Testing SVM and MLP using test set.**

Table 6 details the results of the SVM and the MLP techniques in the test phase.

**Table 6. Results of the test phase of SVM and MLP**

	SVM			MLP		
Error rate	0.011			0.1031		
	Confusion matrix			Confusion matrix		
	No	Yes	Sum	No	Yes	Sum
No	1130	8	1138	1020	118	1138
Yes	6	87	93	10	83	93
Sum	1136	95	1231	1030	201	1231

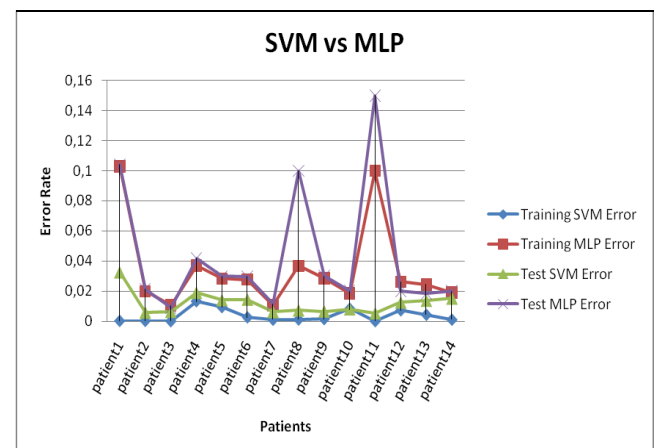
Looking at Table 6, we can conclude that SVM technique has proved its ability to improve the current system since it has successfully detected most of the true alarms and considerably reduced the number of false alarms (8 false alarms). However, the MLP has triggered 118 false alarms.

**5.4 Comparative results**

In this section, we will compare the proposed classification techniques i.e. the MLP and the SVM to the existing system. Note that, the same learning and test sets are used in both techniques.

Generally, we notice that the SVM is more efficient than the MLP in learning and test sets. This result is due to the generation capacity of the SVM in the learning phase through the maximization of the margin.

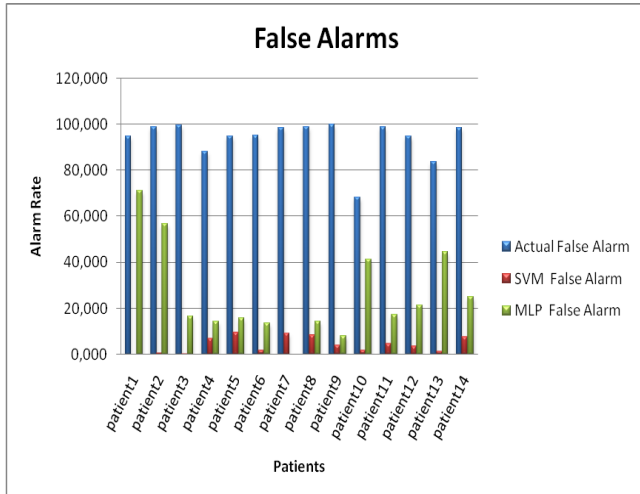
Figure 5 illustrates the results of both the SVM and the MLP techniques in the learning and test phases. It shows that the error rate relative to the SVM technique is lower than the error rate of the MLP technique for the different patients' databases.



**Fig 5: Comparative error rate of SVM and MLP (training and test phases)**

Figure 6, detailed as follows, presents the rate of false alarms for the different databases relative to the patients.



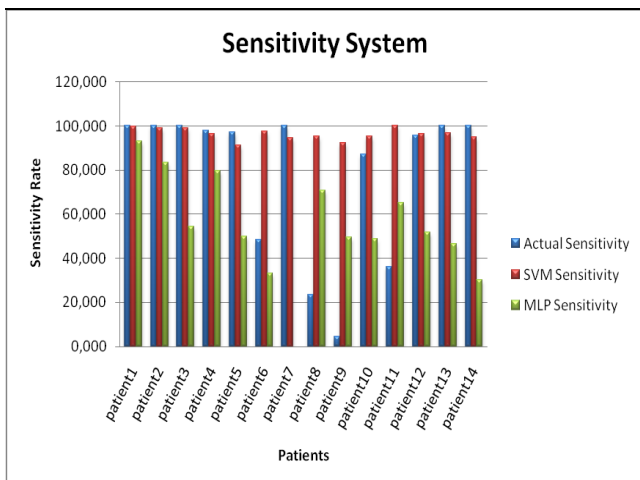


**Fig 6: Comparative performance of SVM to reduce the rate of false alarms**

From Figure 6, we can conclude the efficiency of the SVM since it minimizes and even interestingly annuls the rate of false alarms.

Moreover, it produces more important results than the MLP which has also reduced the rate of false alarm compared to the current system. As both of the SVM and MLP techniques have reduced the false alarm rate, we can conclude that the chosen techniques are efficient and produce good results compared to the current system.

Moving to the sensitivity criterion, Figure 7 explains the sensitivity of the existing system compared to the developed systems i.e. the SVM and the MLP.



**Fig 7: The sensitivity of SVM, MLP and existing system**

Based on Figure 7, we can notice that the existing system can lose its sensitivity to detect alarm especially in the cases of the sixth and ninth patients because the medical staff has increased the threshold; thus, decreasing the number of triggered alarms. In fact, the staff changed the thresholds of the variable ABPSYS and respiration respectively from [80; 140] to [80; 220] and from [5; 25] to [8; 35].

Besides, we observe that the sensitivity is stable for the SVM but it is variable for the MLP and that prove again the efficiency of Vapnik technique.

**Table 7. True positive alarms of the three systems compared to the expert**

Data Sets	CS-TA	SVM-TA	MLP-TA	Expert
Patient1	309	308	288	309
Patient2	1076	1066	898	1076
Patient3	740	734	402	740
Patient4	4533	4476	3690	4636
Patient5	1942	1823	996	195
Patient6	883	1783	607	1826
Patient7	54	51	0	54
Patient8	24	98	73	103
Patient9	27	543	290	587
Patient10	2002	2189	1120	2301
Patient11	378	1042	680	1042
Patient12	1232	1245	670	1290
Patient13	6909	6685	3210	6909
Patient14	697	663	210	697

Table 7 compares the number of true positive alarms of the three systems and the expert where CSTA, SVM-TA and MLP-TA denotes the number of true alarms relative respectively to the current system, the SVM and the MLP.

This table (Table 7) proves the ability of the SVM to detect true alarms for all patients' databases. However, the MLP technique does not produce interesting results when it deals with small bases. Furthermore, the result, provided in Table 7 and relative to the sixth patients' databases, proves that the SVM gives a better quality of result than the MLP.

Generally, we can conclude that the SVM results are close to the expert opinion which proves the ability of the SVM to simulate the expert approach.

## 6. CONCLUSION

In this work, we have adapted two techniques, namely, the SVM and the MLP to a medical area. More precisely, we have tested these approaches in the ICU in order to minimize the rate of false alarms, improve the work conditions of medical staff and guarantee the safety of the patients.

The SVM method has been successfully applied to monitoring system in the ICU by regulating its parameters which are fewer than the parameters of the other classification method (MLP).

Furthermore, all obtained results have emphasized the effectiveness of the SVM approach compared to both of the MLP method and the existing system in the ICU. Besides, the reliability of the models has been examined and compared to expert opinion. It has been shown that the results of the SVM model are in agreement with expert classification and they indicate the ability of the SVM model to detect critical patient states.

However, though its interesting results, the proposed system has been only applied off-line. Besides, it needs huge databases for the learning task and to make a good setting of parameters. Thus, as future work we propose to improve this system and our aim is to make it on-line.

## 7. ACKNOWLEDGMENTS

Authors would like to thank the staff of “military hospital” for their help and for providing interesting information.

## 8. REFERENCES

- [1] Fried, R., Gather, U., Imhoff, M. and Bauer, M. 2000. Some statistical methods in intensive care online monitoring- a review, in: Proceedings of the First International Symposium on Medical Data Analysis (2000), 67–77.
- [2] Tsien, C. 1997. Reducing false alarms in the intensive care unit: A systematic comparison of four algorithms, in: Proceedings of the AMIA Annual Fall Symposium, American Medical Informatics Association, 1997.
- [3] Chambrin, M. C. 2001. Alarms in the intensive care unit: how can the number of false alarms be reduced, *Critical Care* 5 (2001) 184–188.
- [4] Fried, R., Gather, U. and Imhoff, M. 2001. Pattern recognition in intensive care online monitoring, Univ., SFB 475, 2001.
- [5] Bauer, M., Gather, U. and Imhoff, M. 1999. The Identification of Multiple Outliers in Online Monitoring Data, 1999.
- [6] Yoel, D. and Jacob, S. F. 2001. The hostile environment of the intensive care unit, *Current Opinion in Critical Care* 8 (2002), 316–320.
- [7] Salatian, A. and Oborkhale, L. 2011. Filtering of icu monitor data to reduce false alarms and enhance clinical decision support, *International Journal of Bio-Science and Bio-Technology* 3 (2011).
- [8] Charbonnie, S. and Gentil, S. 2007. A trend-based alarm system to improve patient monitoring in intensive care units, *Control Engineering Practice* 15 (2007), 1039–1050.
- [9] Borowski, M., Siebig, S., Wrede, C. and Imhoff, M. 2011. Reducing false alarms of intensive care online-monitoring systems: An evaluation of two signal extraction algorithms, *Comp. Math. Methods in Medicine* 2011 (2011).
- [10] Noura, K. and Trabelsi, A. 2011. Intelligent monitoring system for intensive care units, *Journal of Medical Systems* (2011).
- [11] Imhoff, M., and Kuhls, S. 2006. Alarm algorithms in critical monitoring, *Anesthesia and Analgesia* 102 (2006), 1525–1537.
- [12] Cristianini, N. and Shawe-Taylor, J. 2000. *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*, 1 edition, 2000.
- [13] Cortes, C. and Vapnik, V. 1995. Support vector networks, *Machine Learning* 20 (1995).
- [14] Ming, H. T. and Vojislav, K. 2005. Gene extraction for cancer diagnosis by support vector machines, *Artificial Intelligence in Medicine* 35 (2005).
- [15] Ivanciuc, O. 2007. *Applications of Support Vector Machines in Chemistry*, 2007.
- [16] El Sayed, A. Wahed, I. and Al Emam A. B. 2012. Feature Selection for Cancer Classification: An SVM based Approach, *International Journal of Computer Applications IJCA* (2012) 46(8):20-26.
- [17] Zhang, S.-W., Pan, Q., Zhang, H. C., Zhang, Y. L., and Wang, H. Y. 2003. Classification of protein quaternary structure with support vector machine, *Bioinformatics* 19 (2003) 2390–2396.
- [18] Kotsiantis, S. B. 2007. Supervised machine learning: A review of classification techniques, in: *Emerging Artificial Intelligence Applications in Computer Engineering: Real World AI Systems with Applications in eHealth, HCI, Information Retrieval and Pervasive Technologies*, volume 31, (2007), 3–24.
- [19] Bishop, C. 1995. *Neural networks for pattern recognition*, 1995.
- [20] Vapnik, V. 1998. *Statistical learning theory*, *IEEE Transactions on Neural Networks* 10 (1998) 988–999.
- [21] Yi-wei, C. and Chihjen, L. 2006. Combining svms with various feature selection strategies, *Strategies* 324 (2006) 1–10.
- [22] Goldberger, A., Amaral, L., Glass, L., Hausdorff, J., Ivanov, P., Mark, R., Mietus, J., Moody, G., Peng, C. and Stanley, H. 2000. *PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals*, volume 101, 2000.
- [23] Rakotomalala, R. 2005. Tanagra: un logiciel gratuit pour l’enseignement et la recherche, in *Actes de EGC’2005*, 2 (2005).