ACTA UNIVERSITATIS APULENSIS No 11/2006

Proceedings of the *International Conference on Theory and Application of Mathematics and Informatics ICTAMI 2005* - Alba Iulia, Romania

A STATISTICAL SUPERVISED CLASSIFICATION USING A PERCEPTRON CRITERION FUNCTION

Iuliana Florentina Iatan

Abstract. In this paper we project a statistical classifier, based on perceptron criterion function. To this aim we use a statistical model to set up a programming code in order to test the ECG signal classification for ischemic heart disease diagnosis. This programm uses the database [2] of 40 prototypes (2 classes). The programm recognition score has been of 100% for the training lot and of 95% for the test lot, for almost all the considered variants of feature selection with KLT(Karhunen- Loeve Transformation).

2000 *Mathematics Subject Classification*: 68-xx, 68Txx, 68T10.

1. Introduction

The most of all classification or pattern recognition methods is to decide to which class an unknown pattern belongs. Pattern recognition specially uses mathematical methods of statistical decision theory and this approach one named statistical pattern recognition. In the field of statistical recognition,the classification decision one based on a similarity measure or a discriminant function.

The statistical pattern recognition system is shown in Fig.1 and consists of the following two stages:

a) feature selection,

b) statistical classifier.

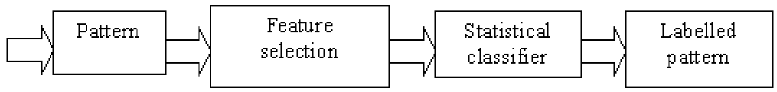

Figure 1: The statistical pattern recognition system

The feature selection strongly affects the statistical classification performances and for this reason it is thought to be one of the main stages of the statistical pattern recognition process.

At this point, it should be mentioned that some of the classification algorithms are more efficient on spaces of low dimension and fail to be practical when applied on spaces of higher dimension; transformations of the original n-dimensional feature space into a space of lesser dimension are therefore strongly recommended.

In this respect, we choose here the Karhunen- Loeve transformation = KLT [1] or Principal Component Analysis= PCA in order to achieve the passing from the initial observation space to a lesser dimension space.

2. Architecture of the statistical perceptron classification model.

We suppose that we have a set of N vectors $\{X_1, X_2, \ldots, X_N\} \subseteq \Re^n$ which belong in two classes $\omega(X_k) \in \{\omega_1, \omega_2\}$,where $\omega(X_k), k = \overline{1, N}$ represents the class whose belongs the vector X_k. Namely, each vector $X_k, k = \overline{1, N}$ is labelled either in the class ω_1 or in the class ω_2.

The projection of a classifier based upon the perceptron criterion function consists in the following stages:

1. the training stage when we adjust the weight vector $W^t = (w_1, w_2, \ldots, w_n)$,
2. the test stage when we realize the choice of the class which will be allotted the vector $X_k, k = \overline{1, N}$, proceeding from the decision rule
$$g(X_k) = W^t X_k \begin{cases} > 0 & \text{if } X_k \in \omega_1 \\ \leq 0 & \text{if } X_k \in \omega_2 \end{cases} \quad (1)$$
after we calculated the discriminant function values $g(X_k) = W^t X_k, k = \overline{1, N}$.

A simple way of determining W is [1] to make the variable change $Y = X$ for $X \in \omega_1$ and $Y = -X$ for $X \in \omega_2$ in order to have a restriction in (1) with a single sign.Then (1) become

$$W^t X_k > 0, (\forall)\ Y_k \in \omega_1 \text{ or } Y_k \in \omega_2 \quad (2)$$

In this case, we will choice the following perceptron criterion function for solving the inequation system $W^t X_k > 0$

$$J(W) = \sum_{Y_k \in G(W)} (-W^t Y_k) \tag{3}$$

which will be minimized, where $G(W)$ is the wrong classified vector set of W ($J(W) = 0$ if none sample was wrong classified).

The recursive estimation equation of the weight vector, upon the negative gradient method [1]is:

$$W_{k+1} = W_k - c \cdot \nabla_W J|_{W=W_K}, c > 0 \tag{4}$$

where $\nabla_W J|_{W=W_K} = (\partial J/\partial W)_{W=W_K}$ is the gradient vector.

From (3) results

$$\nabla_W J = -\sum_{Y_k \in G(W)} Y_k \tag{5}$$

and the recursive equation of the negative gradient become

$$W_{k+1} = W_k + c \cdot \sum_{Y_k \in G(W_k)} Y_k \tag{6}$$

where $G(W_k)$ is the wrong classified vector set of W_k.

If we assume that the samples are sequential observed then the algorithm (6) become:

$$\begin{cases} W_{k+1} = W_k & \text{if } W_k^t Y_k > 0, k = \overline{1, N}, \\ W_{k+1} = W_k + c \cdot Y_k & \text{if } W_k^t Y_k \leq 0, k = \overline{1, N} \end{cases} \tag{7}$$

3.Experimental results

The signals ECG classification is an interesting practical application because the ischemic heart diagnosis occupies a very important place among the heart diseases with a great probability of appearance. We will use a statistical model of perceptron type for ischemic heart diagnosis which admits of entrance pattern, the reduced prototypes of ECG signals [2]. For testing the proposed classifier using the criterion perceptron function we have used the ECG database of Grunwald [2] with 40 prototypes (consist of 128 samples): 20 corresponds of some normal subjects and other 20 belongs to some patients with ischemic heart disease.The ECG database consists of the training lot and the test lot, each of them having 20 subjects from the both categories: 10 normal (fig.2) and 10 with ischemic heart disease (fig.3).

We will present a part of the Matlab program which realizes the classification of the ECG signals in one of the two classes, upon the statistical model from [1].

At first, one assume $w \in \Re^{m+1}$ a weight random vector in [-1,1], thus:
$w = rand(m + 1, 1);$
$w = 2 * w - 1;$
where m is the number of features extracted with PCA.

We will consider the following two stages:

1. the training stage when we will adjust the weight vector, on the basis of (7);
2. the test stage of a training lot and of the test lot when we will not modify the weight vector, obtained from the first stage.

The both stages will finish in a finite number of epochs $Nmax$(an epoch means the crossing of all the vectors).

We consider $\{Y_1, Y_2, \ldots, Y_{20}\} \subseteq \Re^{m+1}$ the vector set of the first class, respectively $\{X_1, X_2, \ldots, X_{20}\} \subseteq \Re^{m+1}$ the vector set from the second class.
$i = 1;$
% beginning of the first stage
while($i <= Nmax$)
p=1;
while($p <= 20$)
if $p == k$
$H = Yk;$
elseif $p == k + 1$
$H = Xk;$
end
V=H(:);
z=fi'*V; % fi is the matrix of the eigenvectors
if $mod(k, 2)\ = 0$
$z = [1\ \ z'];$
else
$z = [-1 - z'];$
end
$z = z';$

```
if w' * z <= 0
w = w + 0.1 * z;
end
p = p + 1;
end
% beginning of the second stage (the test stage of the test lot)
p = 1;
while(k <= 20)
if p == k
H=Yk;
elseif p == k + 1
H = Xk;
end
V = H(:); z = fi' * V;
z = [1 z']; z = z';
if w' * z >= 0
et(k) = 1;
elseif w' * z < 0
et(k) = 2;
end
p = p + 1;
end
er = 0;
for k = 1 : 20
if et(k) == ett(k)
er = er + 1;
end
end
rt(i) = (er * 100)/20;
i = i + 1;
end
```

We denoted by $et(k)$ the real label, by $ett(k)$ the ideal label of the k- the vector and by $rt(i)$ the recognition score over the test lot at the epoch i.

One obtain the results from the Table 1.

- m is the number of features selected with PCA,
- R_1 -recognition score over the training lot,
- R_2 -recognition score over the test lot,

m	R_1	R_2	N_1	N_2	N_3
15	100%	85%	6	-	10
25	100%	95%	3	3	10
50	100%	95%	2	1	10

Table 1: The performances of the statistical model based on perceptron criterion function, experimented to ECG signals classification, with PCA.

- N_1-the number of epochs after which the recognition score over the training lot becomes 100%,
- N_2-the number of epochs after which the recognition score over the test lot becomes $\geq 90\%$,
- N_3-the total crossing epochs number with a view to obtaining an as good as possible recognition score over the test lot.

The next figures present the recognition score over the training lot (fig.2) and, respectively over the test lot, for $m = 15$,$m = 25$,$m = 50$ features extracted with PCA (fig.3).

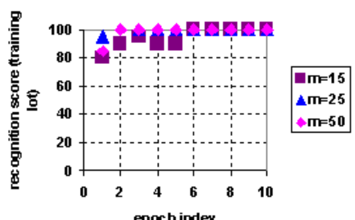

Figure 2: Recognition score over the training lot for $m = 15, m = 25, m = 50$

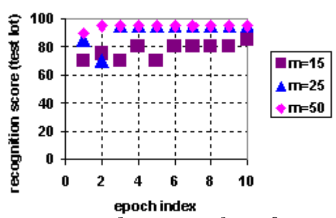

Figure 3: Recognition score over the test lot for $m = 15$, $m = 25$, $m = 50$

4.Conclusion

The aim of this paper is to realize a programming code in order to test the ECG signal classification using a statistical model of type perceptron [1].

We also made a programm for feature selection with PCA, in order to reduce the space dimensionality.

From Table 1, one can observe that our programming recognition score is the best (of 95%) for the test lot the both for $m = 25$ and for $m = 30$ selected features with PCA.

In [2] we obtained the best recognition score of 100% over the test lot, using a neuro- fuzzy classifier (called Fuzzy Gaussian Neural Network), in the case of $m = 50$ features extracted with PCA.

References

[1] V. E. Neagoe, O. Stănasilă,*Teoria recunoaşterii formelor [The theory of pattern recognition, Ro]*, Bucharest, Academia Română Publishers, 1992.

[2] V. E. Neagoe, I.F. Iatan, S. Grunwald, *A Neuro- fuzzy Approach to Classification of ECG Signals fos Ischemic Heart Disease Diagnosis*, the American Medical Informatics Association Symposium (AMIA 2003), Nov. 8- 12 2003, Washington DC, pp. 494-498.

[3] V. E. Neagoe, O. Stănasilă, *Recunoaşterea formelor şi reţele neurale* [Pattern recognition and neural networks, Ro]. Bucharest, Matrix Rom Publishers, 1999.

[4] R. Fuller, *Introduction to Neuro- Fuzzy Systems*, Heidelberg, Physica-Verlag, 2000.

Iuliana Florentina Iatan
Department of Mathematics and Informatics
Technical University of Civil Engineering of Bucharest
124, Lacul Tei Blvd. 020396 Bucharest 38, Romania
email: iuliai@k.ro, iuliafi@yahoo.com