# Classification of Heart Rate Signals using Support Vector Machines

Kampouraki A, Nikou C, Manis G
University of Ioannina
{cnikou, manis}@cs.uoi.gr

*Abstract. We use support vector machines to classify heart rate signals. The signals of twenty youngand twenty elderly subjects were analyzed with eleven common heart rate variability computation methods. All eleven methods failed to classify the signals correctly. We stacked the results of these eleven methods in a vector for each heart rate signal and applied kernelbased SVMs to perform signal classification. This time we achieved to differentiate correctly for the considered dataset all young and elderly subjects using leave-one-out crossvalidation.*

## 1 Introduction

Heart Rate Variability (HRV) analysis is based on measuring the variability of heart rate signals and more specifically, the variability in intervals between R peaks of the electrocardiogram (ECG), referred as RR intervals. A survey of methods for the analysis of signals of RR intervals can be found in [1]. In this study, we discuss the use of Support Vector Machines (SVMs) [2] to classify heart rate signals. Experimental results show perfect categorization of subjects using SVMs where standard methods fail to successfully classify the input data.

## 2 Methods

Support vectors classifiers are based on recent advances on statistical learning theory. They use a hypothesis space of linear functions in a high dimensional feature space, trained with a learning algorithm from optimisation theory that implements a learning bias derived from statistical learning theory. This learning strategy is principled and very powerful method that has outperformed most other systems in a wide variety of applications [3].

The learning machine is given a training set of examples (or inputs), belonging to two classes, with associated labels (or output values). The examples are in form of attribute vectors and the SVM finds the hyperplane separating the input data and being furthest from both convex hulls. If the data are not linearly separable a set of slack variables is introduced representing the amount by which the linear constrained is violated by each data point. Moreover, for many datasets, it is unlikely that a hyperplane will yield a good classifier. Instead, we want a decision boundary with more complex geometry. One way to achieve this is to map the attribute vector into some new space of higher dimensionality and look for a hyperplane in that new space, leading to kernel-based SVMs [2]. The interesting point about kernel functions is that although classification is accomplished in a space of higher dimension, any dot product between vectors involved in the optimization process can be implicitly computed in the low dimensional space [2].

Standard classification techniques applied to the study of HRV rely on signal statistics. In this study we have stacked 11 standard measures in a vector for each ECG and applied kernel-based SVMs to perform signal classification. The measures we have considered are: the standard deviation (SDNN), the standard deviation of the average RR interval calculated over 5 minutes (SDANN), the root mean square of successive differences (RMSSD), the mean of the 5-min standard deviation of the RR interval calculated over 24 hours (SDNNi), the standard deviation of successive differences (SDSD), the number of interval differences of successive RR intervals greater than 50 ms over the number of RR intervals (pNN50), the mean prediction error of the signal using local linear prediction (LLP), the total number of all RR intervals divided by the height of the histogram of all RR intervals measured on a discrete scale with bins of 7·8125 ms (TI), the baseline width of the minimum square difference triangular interpolation of the highest peak of the histogram of all NN intervals (TINN), the signal entropy and the signal autocorrelation.

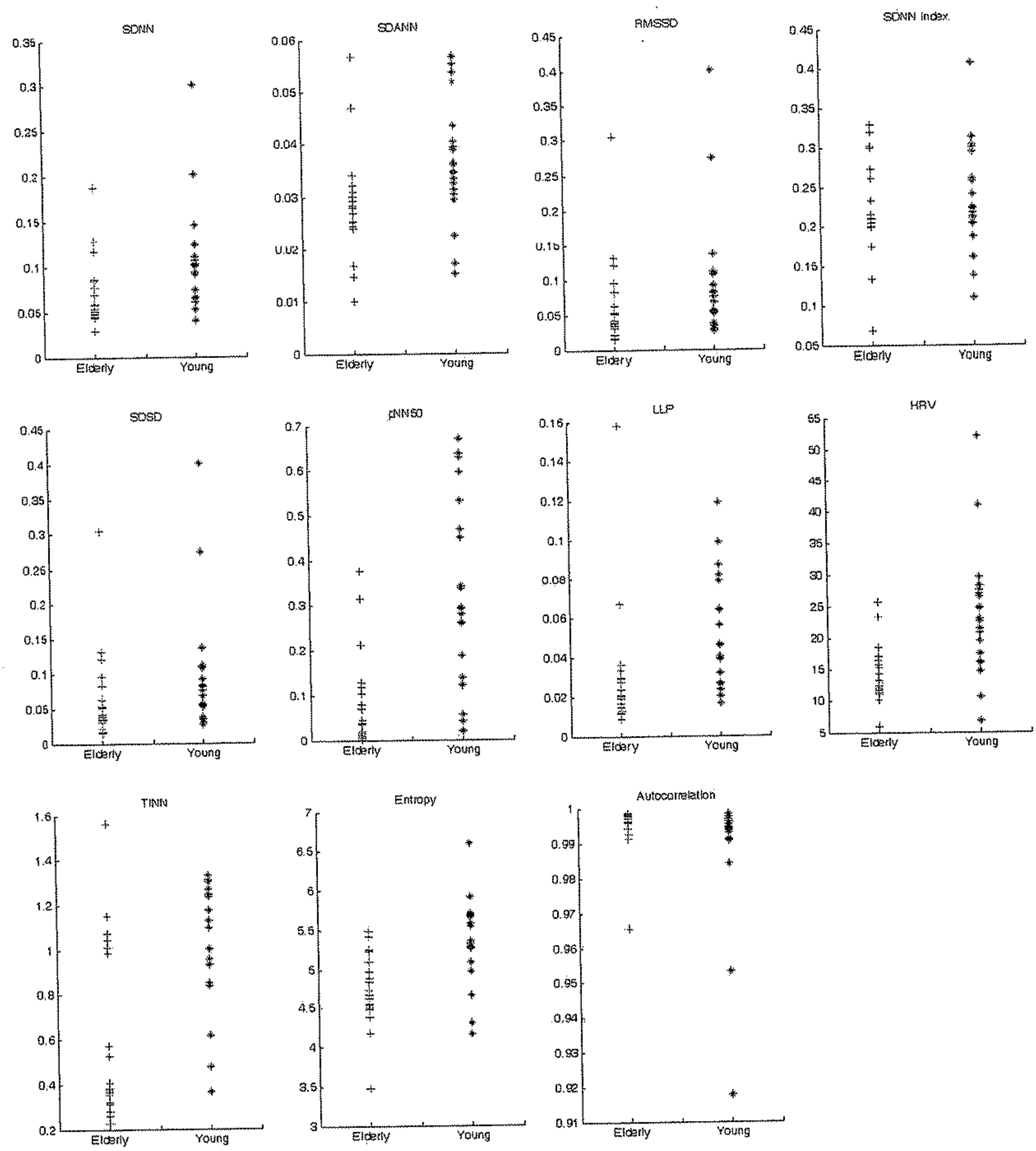


Fig 1. Some standard HRV methods used for classification. Crosses represent elderly subjects expected to show decreased values of HRV, while stars represent young subjects expected to show larger values of HRV. The vertical axis shows the value of the respective feature. Notice that the data are not clearly separable using the features independently. See text for feature acronyms.

## 3 Results

We applied SVM classification to several data sets. We present here results on long-term (about 2 hours in duration) ECG recordings. Twenty young (21 - 34 years old) and twenty elderly (68 - 85 years old) rigorously-screened healthy subjects underwent 120 minutes of continuous supine resting while continuous electrocardiographic (ECG) signals were collected. Each subgroup of subjects includes equal numbers of men and women. All subjects remained in a resting state in sinus rhythm while watching the movie Fantasia (Disney, 1940) to help maintain wakefulness. The continuous ECG signals were digitized at 250 Hz. Each heartbeat was annotated using an automated arrhythmia detection algorithm, and each beat annotation was verified by visual inspection. The data are available and further described in [4-6].

In figure 1, the classification results using some standard measures are presented. Notice that using the standard measures the subjects are not correctly classified. On the other hand, SVM achieves to correctly differentiate young and elderly subjects at 100% for the considered dataset. Leave-one-out cross-validation was used to evaluate the classifier. In all cases the categorization was 100% accurate..

## 4 Conclusions

The categorization of the ECGs into two dinstinct groups according to their heart rate variability can be very accurate using an SVM in cases where standard methods fail to present a satisfactory categorization. Early experiments comparing SVM classification of heart rate signals with the classifications obtained by other non-linear classifiers have also confirmed the effectiveness of the former methodology. Future research work consists of examining the robustness of SVMs to classification of heart rate signals with noise.

## References

[1] Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. Heart rate variability: Standards of measurement, physiological interpretation, and clinical use. European Heart Journal 17:354-381,1996.

[2] K. R. Muller, S. Mika, G. Ratsch, K. Tsuda, B. Scholkopf. An introduction to kernelbased learning algorithms. IEEE Trans. on Neural Networks, 12(2):181-202, 2001.

[3] H. Byun and S. W. Lee. A survey of pattern recognition applications of support vector machines. Int. J. of Pattern Recognition and Artificial Intelligence.17(3):459-486, 2003.

[4] PhysioBank: Physiol. signal archives for biomed. research, http://www.physionet.org.

[5] A. L. Goldberger, LAN Amaral, L. Glass, J. M Hausdorff, PCh Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C. K. Peng, H. E. Stanley. PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. Circulation 101(23):e215-e220, 2000.

[6] N. Iyengar, C. K. Peng, R. Morin, A. L. Goldberger, L. A. Lipsitz. Age-related alterations in the fractal scaling of cardiac interbeat interval dynamics. Am J Physiol 271:1078-1084, 1996.