



HAL
open science

An Experimental Comparison of Supervised Classifiers on Two Real Datasets for Pattern Recognition

Mohammed Ngadi, A. Amine, B. Nassih, H. Hachimi

► **To cite this version:**

Mohammed Ngadi, A. Amine, B. Nassih, H. Hachimi. An Experimental Comparison of Supervised Classifiers on Two Real Datasets for Pattern Recognition. International Meeting on Advanced Technologies in Energy and Electrical Engineering, Nov 2018, Fès, Morocco. hal-01941989

HAL Id: hal-01941989

<https://hal.science/hal-01941989>

Submitted on 2 Dec 2018

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.

An Experimental Comparison of Supervised Classifiers on Two Real Datasets for Pattern Recognition

¹Ngadi, M., ²Amine, A., ³Nassih, B. & ⁴Hachimi, H.
Systems Engineering Laboratory National School
of Applied Sciences. Ibn Tofail University
Kenitra – Morocco
Ngadi.mohammed@univ-ibntofail.ac.ma

Abstract

In this paper, we present a comparative study between Support Vector Machine (SVM) and Adaboost, as being two decisions based classification tools in the field of shape recognition. The aim of this work is to study their theoretical foundations, their learning algorithms and to see their performance in classification capacity. To compare their performance, we apply them to two famous training datasets, widely used by the community, namely the CBCL MIT face and Wisconsin diagnosis breast cancer (WDBC). The quality of decision of each classifier depends on the choice of its parameters and its implementation.

keywords— supervised learning, unsupervised learning, classification, shape recognition, SVM, Adaboost.

Introduction

The field of pattern recognition [1, 2, 3] knows a revolution since the mid-90s with the statistical learning theory and the advent of the Support Vector Machines [4, 5, 6] (SVM) for the resolution of detection problems, classification and regression. In recent years, appeared a set of highly interdependent disciplines, concerning the information treatment, decision theory and methods of pattern recognition, Boosting methods [7, 8, 9]. Their field of applications is much expanded and extended to several areas, in particular in: shapes recognition, the approximation of functions, image processing, speech recognition, classification. The objective of this paper is to compare their performance in the field of the supervised classification.

The term may refer to two classes of distinct methods: the supervised classification and unsupervised classification. Non supervised methods are intended to constitute examples groups (or groups of instances) based on observed data, without a priori knowledge. On the

other hand, supervised methods use a priori knowledge on the belonging of a sample to a class to build a recognition system based on these classes. In this paper we focus supervised classification.

The goal of supervised classification is to build, using a set of training data (training set), a classification procedure which allows predicting membership of a new example to a class.

Experimental results

We will now test and compare the mentioned classifiers on real datasets. We have used two famous datasets very used by the community, namely the MIT-CBCL face database and the Wisconsin Diagnostic Breast Cancer (WDBC) database.

The method of Boosting is particularly interesting because we can choose the number of classifiers in order to achieve the desired error rates on samples examples. Moreover, we observe that the error rate decreases exponentially with the number of used weak classifiers, figure 1 (L: MIT, R: WDBC).

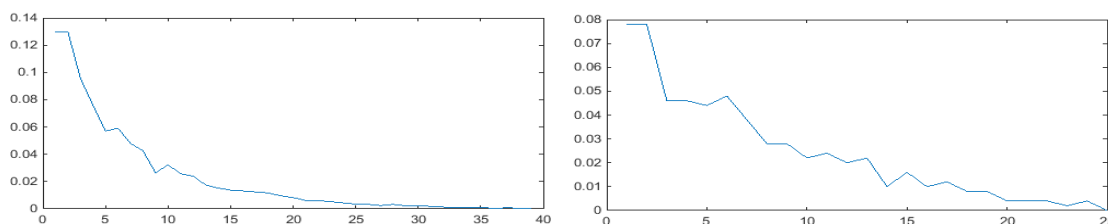


Figure 1. Classification error versus number of weak classifiers

We have used LIBSVM to measure the classification accuracy using SVM. A typical use of LIBSVM involves two steps: first, training a dataset to obtain a model and second using the model to predict information of a testing dataset. As kernel function, we have used a Gaussian radial basis function (RBF).

Table 1 shows that the best classification accuracy of the LIBSVM:

Database	Kernel	Parameter	Accuracy (%)
MIT-CBCL	RBF	Gamma=-20	99.95
WDBC	RBF	Gamma=-20	96.95

Classifiers used for our experiments are Naive Bayes, decision tree, the K Nearest Neighbours (KNN), Support Vector Machines (SVM) and Adaboost (figure 2).

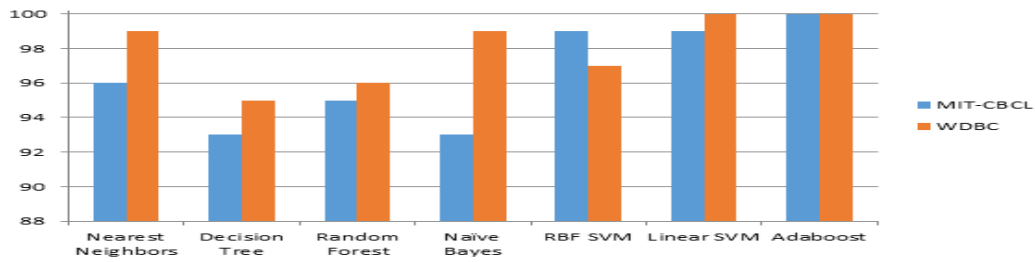


Figure 2. Accuracies of different classifiers methods

Conclusion

In this paper, we have presented a detailed comparative study of two classification algorithms, Adaboost and SVM. We applied them to two famous training datasets, widely used by the community, namely the MIT- CBCL face and the WDBC datasets. The main criteria that we used for comparison is the accuracy of the classification.

The experimental results show that AdaBoost and SVM perform better than other learning algorithms on all the data that we have used.

Our goal in the near future is to continue the study of SVM and Adaboost in order to test the relationships that exist between them.

Références

1. G.Yang,Z.Wang,J.Ren. Facial expression recognition based on adaboost algorithm. Chinese Journal of Computer Application. 25(4), pp. 946-948, (2005).
2. W.Zhu,S.Luo. Face detection based on cascaded boosting algorithm. Chinese Journal of Computer Application. 25(9), pp.2128-2130, (2005).
3. J.Ren,G.Yang. Facial Expression Recognition based on Gabor Transform and AdaBoost Algorithm. Chinese Journal of MicroComputer Information. 23 (3-1), pp.290-292, (2007).
4. Kim, S.Pang, M.Je. Constructing support vector machine ensemble. Pattern Recognition, 36, pp.2757-2767, (2005).
5. I.Buciu. Demonstrating the stability of support vector machines for classification. Signal Processing. 86, pp. 2364-2380, (2006).
6. X.Li, L.Wang, E.Sung. A Study of AdaBoost with SVM Based Weak Learners. Proceedings of International Conference on Neural Networks, pp.196-200, (2005).
7. L.Shen,L.Bai. A review on Gabor wavelets for face recognition. Pattern Anal Applic, (2006).
8. M.Turk, A.Pentland. Eigenfaces for recognition. J Cogn Neurosci, 3(1), pp.71-86, (1991).
9. P.Belhumeur,J.Hespanha,D.Kriegman. Recognition using class specific linear projection. IEEE Trans PAMI, 19(7), pp.711-720, (1997).