Support Vector Machines in Statistical Shape Analysis

Amaral, Getulio, J. A. e Carvalho, Jhonnata, B.
1 Universidade Federal de Pernambuco
2 Universidade Federal do Amazonas

Abbreviated abstract: Statistical Shape Analysis (SSA) is a relevant field of statistics. SSA begins placing landmarks on the object image. After that, some mathematical operation are performed to remove the effects of translation, scale and rotation. Supervised learning is one of the tasks of SSA. Some algorithms to carry out supervised learning in SSA are presented. These algorithm are support vector machines (SVM), hill-climbing, kernel k-means and ensemble of methods. The SVM algorithm outperforms the other methods.

Related publications:
– Carvalho, J. B. And Amaral, G. J. A. , Expert Systems with Applications, v. 151, p. 113320 (2020)

gjaa@de.ufpe.br

4th Conference on
Statistics and
Data Science
Salvador, Brazil (online)
December 1-3, 2022

Problem, Data, Previous Works

Suppose that there are G populations. The researcher desires to classify an object into one of the G populations. It is done with a classifier.
Initially, this problem was solved with Bayes discriminant. This classifier uses the Bayes rule to classify the object into one of G groups.
The proposal of this work is to introduce other classifiers for the SSA. The proposed learners need to be adapted to this new context.
One option is the SVM classifier. The SVM is used to obtain two classes separated by a hyperplane. The algorithm needs to be adapted to complex vectors, which are used in SSA. A optimization problem is defined. This problem is solved with the Lagrange multipliers method.
A second option is Bayes density estimation, which uses the Bayes rule with the density estimation for each density defined.

4th Conference on
Statistics and
Data Science
Salvador, Brazil (online)
December 1-3, 2022

Methods

A third method is a kernel K-means algorithm. This algorithm starts with a initial partition. After that, a k-means algorithm is performed. Finally, the cluster algorithm is adapted to the context of supervised learning. It is done by computing the distance between an object and the cluster centers. The object is assigned to the closest group.
The fourth method is a Hill climbing procedure. It is initially a clustering algorithm but it is adapted to supervised learning context. Four test statistics of SSA are used. These statistics are named Hotelling, Goodall, James and Lambda. The classification is performed as follows. For one fixed object, it is allocated to the groups and the value for each statistics is computed.
It is posible to combine classifiers. This kind of method is called ensemble. Suppose that there are four methods to be combined. The final classification of an object is a four dimensional vector. Moreover, a voting scheme is used to choose which group the objected will be allocated. There are three voting squemes: Majority vote, Plurarity vote and Weighted vote. The object is allocated to the group with largest voting.

Results and Conclusions

The five previously described algorithms were compared in simulation studies and real data analysis.

Table 1: Simulation results for the classifiers error rates with different variability values (σ^2).

Classifier	Type	σ^2						
		0.001	0.005	0.009	0.01	0.014	0.018	0.02
SVM_C	L	0.0000	0.0100	0.0500	0.0500	0.1000	0.1000	0.1400
	P	0.0000	0.0100	0.0400	0.0400	0.0900	0.0900	0.1500
	G	0.0000	0.0100	0.0400	0.0400	0.0900	0.0900	0.1400
$SVM_C(d_F)$	L	0.0000	0.0100	0.0500	0.0500	0.1000	0.1100	0.1500
	P	0.0000	0.0100	0.0400	0.0400	0.0900	0.0900	0.1500
	G	0.0000	0.0100	0.0400	0.0500	0.1000	0.1000	0.1400
$SVM_C(d_P)$	L	0.0000	0.0100	0.0500	0.0500	0.0900	0.1000	0.1500
	P	0.0000	0.0200	0.0400	0.0400	0.0900	0.0900	0.1600
	G	0.0000	0.0100	0.0400	0.0400	0.0900	0.0900	0.1400
$SVM_C(\rho)$	L	0.0000	0.0100	0.0500	0.0600	0.0900	0.0900	0.1600
	P	0.0000	0.0200	0.0400	0.0500	0.0900	0.0900	0.1500
	G	0.0000	0.0100	0.0400	0.0400	0.0900	0.0900	0.1500
Kernel k-means	KKM	0.0000	0.0600	0.1100	0.1000	0.1700	0.2100	0.2000
	KKM_p	0.0000	0.0400	0.0900	0.1300	0.1400	0.1800	0.2000
Bayes-DE	Normal	0.0000	0.0300	0.0700	0.0600	0.1100	0.0800	0.1400
	t-Student	0.0000	0.0300	0.0600	0.0500	0.1100	0.0800	0.1400
Bayes	Bingham	0.0000	0.0600	0.1500	0.1500	0.2400	0.1700	0.2700
	Watson	0.0000	0.0600	0.1100	0.1100	0.1400	0.1900	0.2000
Hill-Climbing	James	0.0000	0.0500	0.0900	0.1200	0.2100	0.2100	0.2600
	Hotelling	0.0000	0.0700	0.0800	0.1200	0.1700	0.2200	0.2700
	Goodal	0.0000	0.0600	0.1200	0.1100	0.1400	0.1900	0.2000
	Lambda	0.0000	0.0700	0.1500	0.1300	0.1800	0.1800	0.2100
Ensemble	MV	0.0000	0.0600	0.1200	0.1300	0.1500	0.2200	0.2400
	PV	0.0000	0.0300	0.0700	0.1100	0.1200	0.1300	0.1700
	WV	0.0000	0.0200	0.0500	0.0600	0.1100	0.1000	0.1700

Table 2: Classifiers error rates for the four data sets based on the method *leave-one-out*.

Classifier	Type	Gorila	Orangutan	Mouse	Schizophrenia
SVM_C	L	0.0169	0.0741	0.0435	0.2143
	P	0.0169	0.1111	0.0000	0.1429
	G	0.0169	0.0556	0.0217	0.2500
$SVM_C(d_F)$	L	0.0169	0.0926	0.0652	0.2143
	P	0.0169	0.0926	0.0000	0.1429
	G	0.0169	0.0556	0.0435	0.2500
$SVM_C(d_P)$	L	0.0169	0.0926	0.0652	0.2143
	P	0.0169	0.0926	0.0000	0.1429
	G	0.0169	0.0556	0.0435	0.2500
$SVM_C(\rho)$	L	0.0169	0.0926	0.0652	0.2143
	P	0.0169	0.0926	0.0000	0.1429
	G	0.0169	0.0556	0.0435	0.2500
Kernel k-means	KKM	0.0508	0.1481	0.0870	0.3929
	KKMp	0.0339	0.0926	0.0435	0.2500
Bayes-DE	Normal	0.0339	0.1852	0.0217	0.3929
	t-Student	0.0339	0.1296	0.0217	0.3929
Bayes	Bingham	0.0339	0.0926	0.0870	0.7500
	Watson	0.1017	0.1296	0.0870	0.3214
Hill-Climbing	James	0.0000	0.1111	0.0870	0.5000
	Hotelling	0.0000	0.1296	0.0870	0.5357
	Goodal	0.0847	0.1296	0.0870	0.3571
	Lambda	0.0169	0.0926	0.0652	0.4643
Ensemble	MV	0.0169	0.1667	0.0652	0.4286
	PV	0.0169	0.1667	0.0652	0.4286
	WV	0.0169	0.0556	0.0435	0.1786