Compact Evolutive Design of Error-Correcting Output Codes

Miguel Angel Bautista, Xavier Baró, Oriol Pujol, Petia Radeva, Jordi Vitrià, and Sergio Escalera

UNIVERSITAT DE BARCELONA

CVC Centre de Visió per Computador

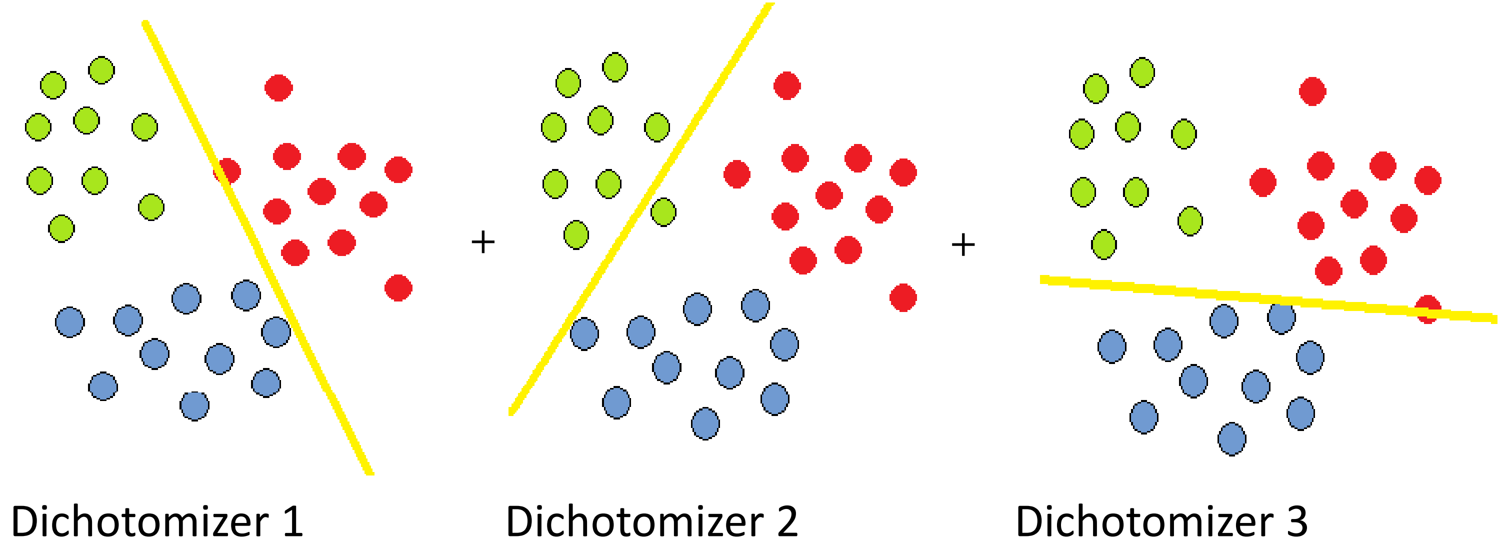
BARCELONA, SPAIN
ECMLPKDD2010
European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases
SUEMA2010 workshop
Supervised and Unsupervised Ensemble Methods and their Applications

Outline

- Error Correcting Output Codes
- SVMs with Gaussian-RBF kernel
- Genetic optimization
- Experiments & results
- Conclusions

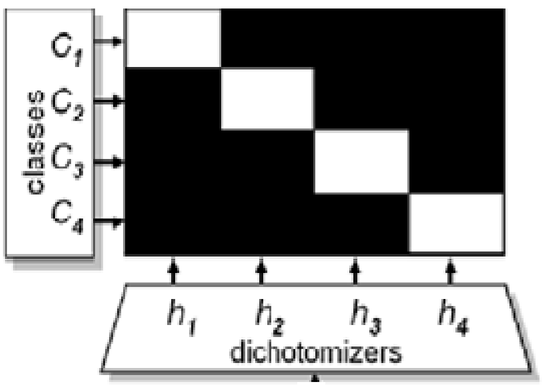
Error Correcting Output Codes (ECOC)

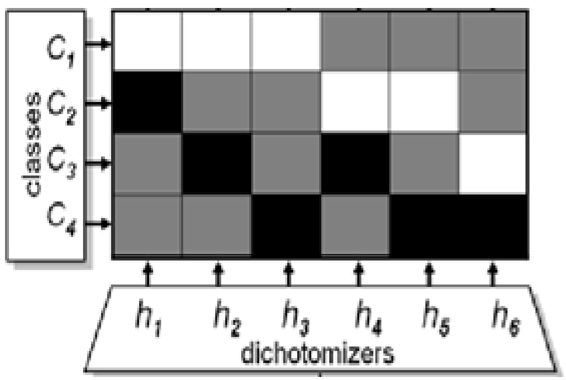
- ECOCs are an ensemble learning methodology which allow to combine dichotomizers to treat multiclass problems.



ECOC coding

- ECOCs can be represented as matrices, which columns represent the different sub-problems to treat.
- Each column has values that distinguish categories in two groups.
- One-versus-All and One-versus-One are the standard codings.

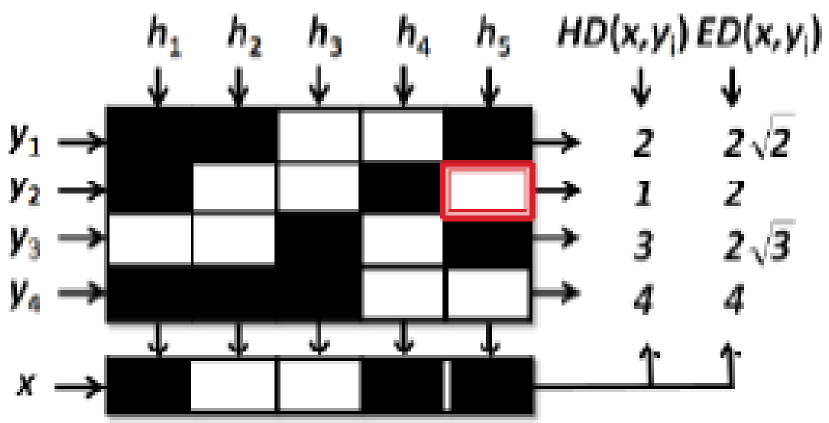




ECOC decoding

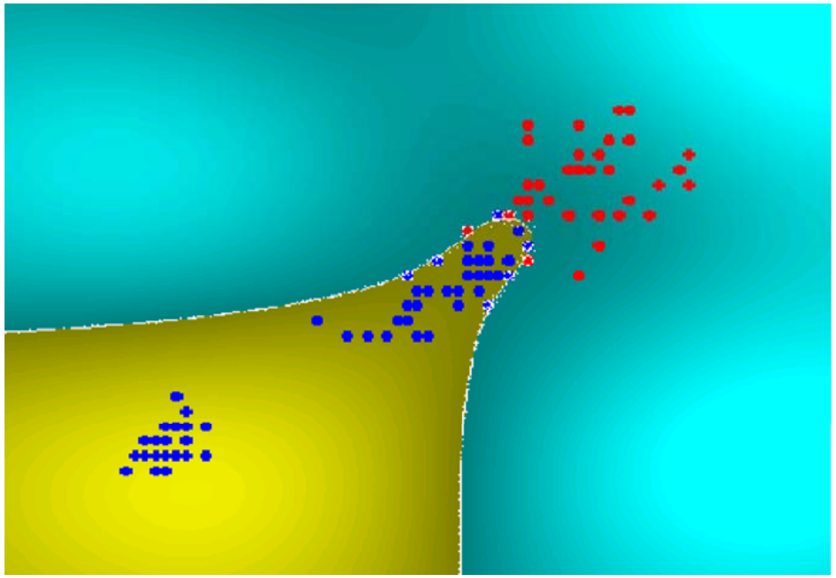
- Each sub-problem is trained and the set of predictions are compared to the codewords.

- Various types of decoding based on Euclidean and Hamming distances (only binary codings).

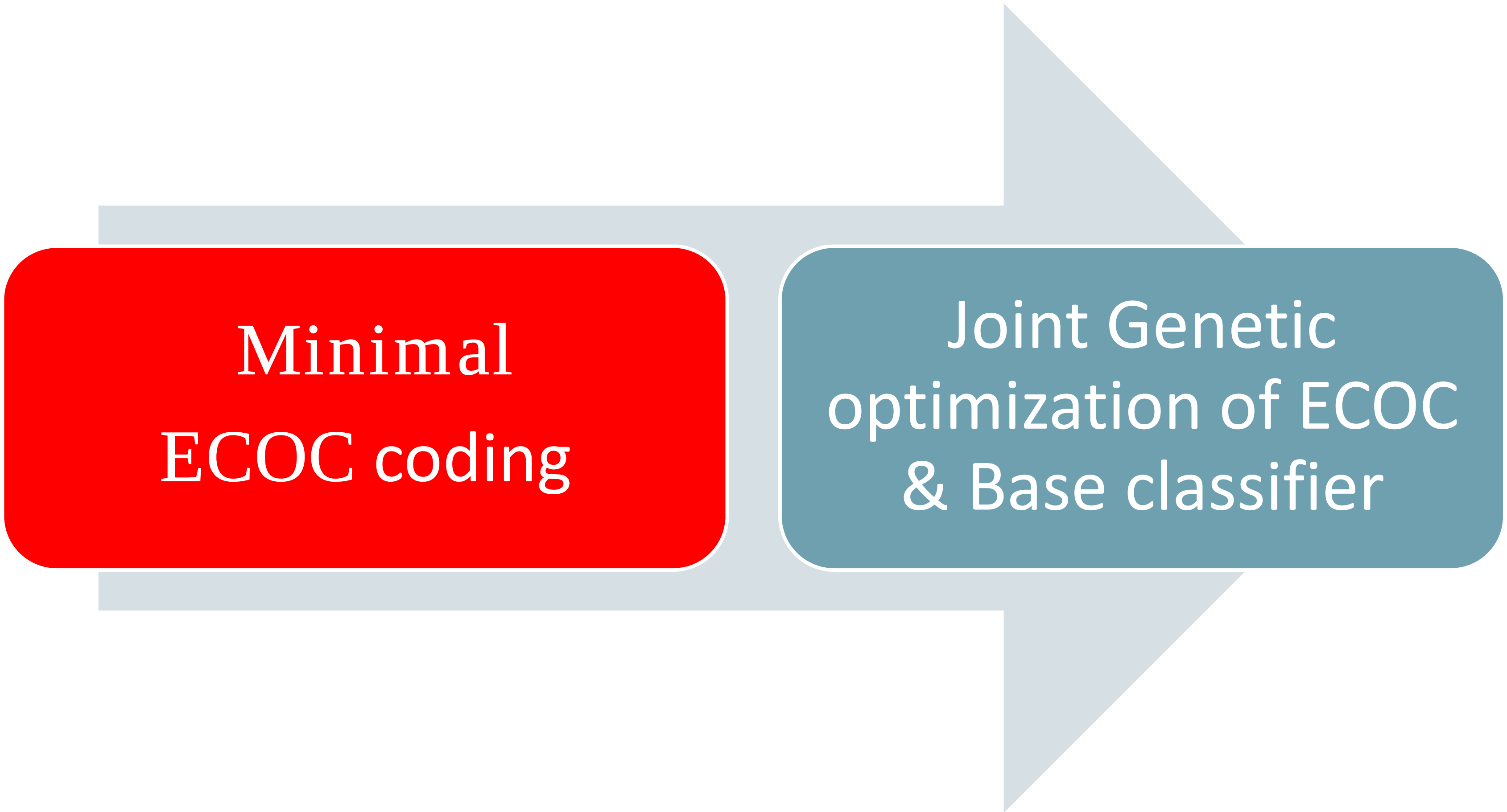


Base classifier: SVM with an RBF kernel

- Each binary problem is learned by a base classifier.
- SVM with RBF kernels have shown a good performance on those kind of problems.
- This type of SVM needs the parameters (C & Gamma) to be optimized.

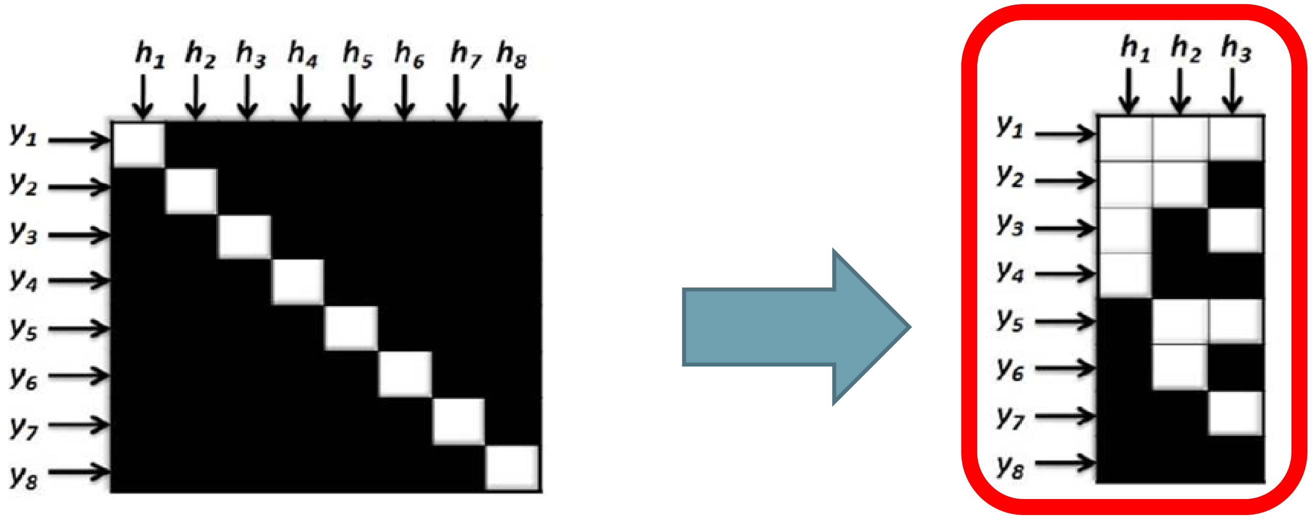


Global overview

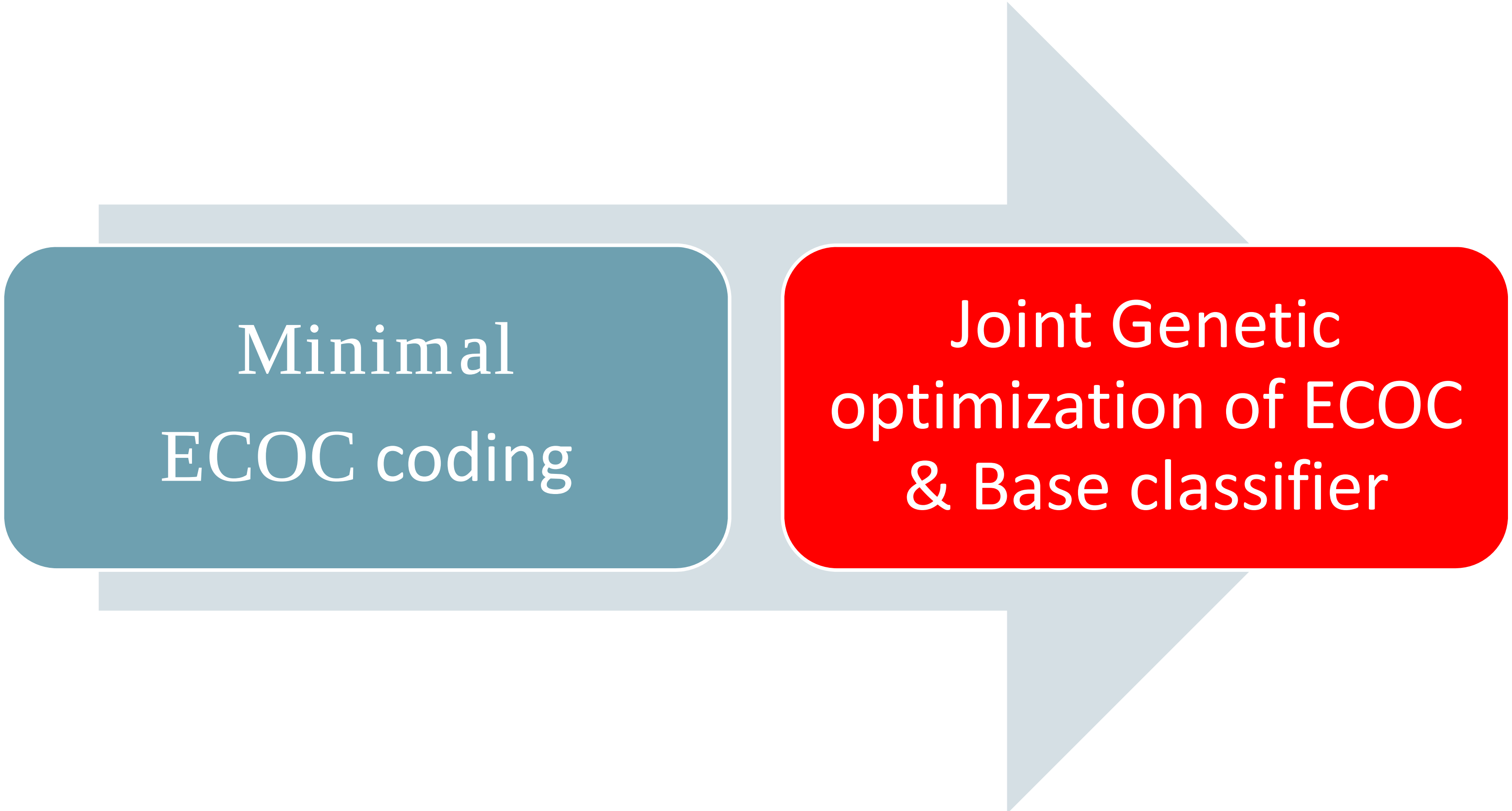


Minimal coding

- Define the minimal number of base classifiers needed to discriminate N categories.
- Taking profit of Information theory only $\log_2 N$ bits are needed to discriminate N categories.

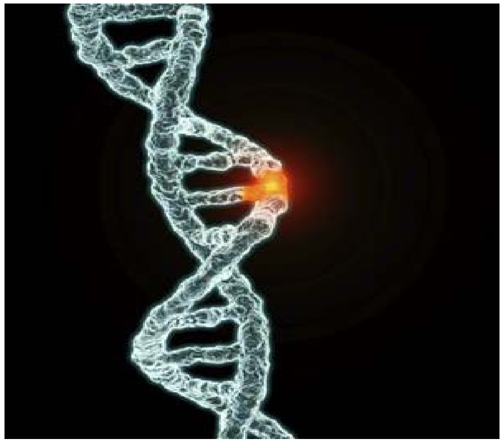


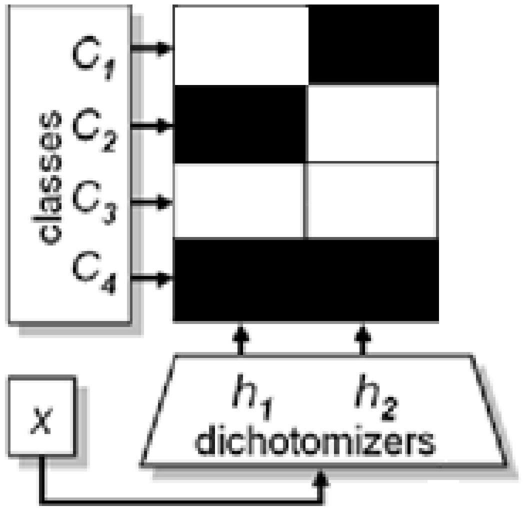
Global overview



Genetic algorithms

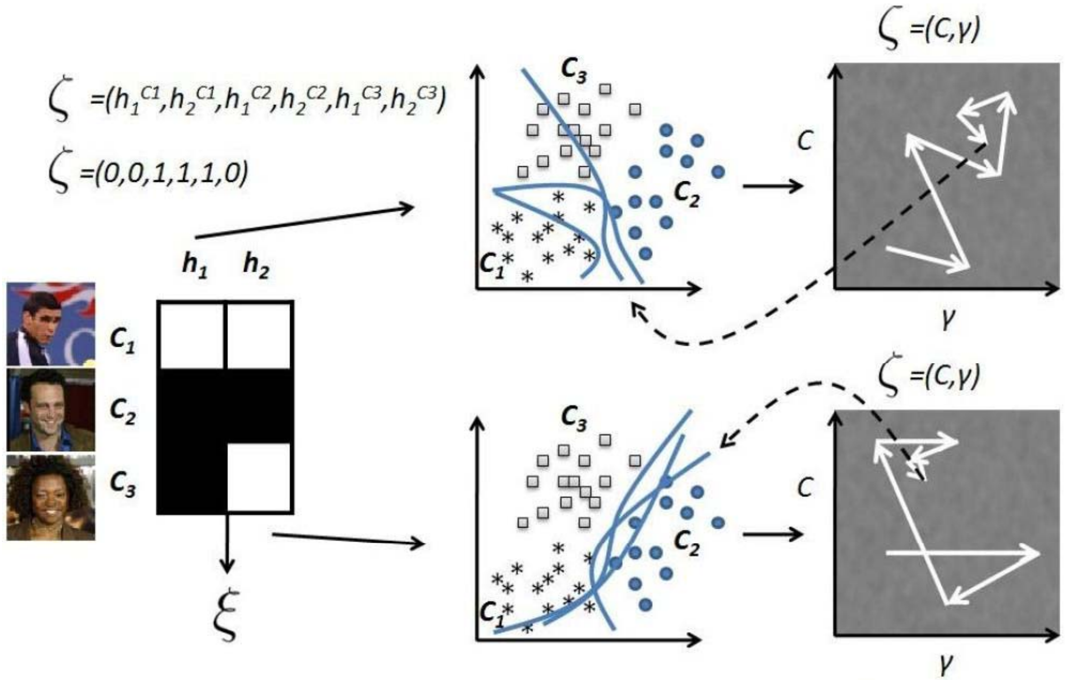
- Optimization algorithms based on the evolution theory of Darwin.
- Recommendable method when the space is not continuous neither differentiable.





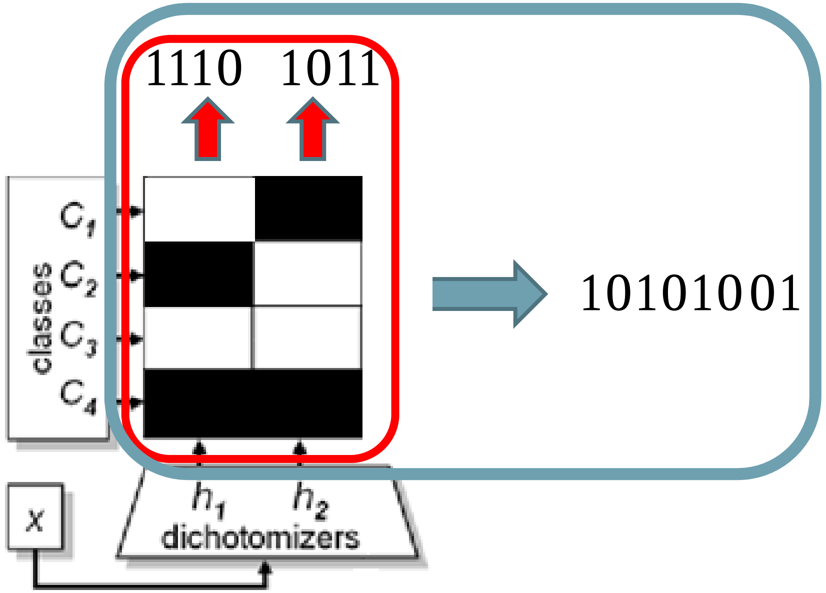
Evolutionary optimization for SVMs

- An optimization process is carried out to tune the parameters of the SVMs.
- SVM-RBF classifiers have mainly 2 parameters (C & Gamma).



Evolutionary optimization for ECOCs

- Each ECOC individual is seen as a binary vector and evaluated by means of its classification error.



A) Optimize the SVMs looking for suitable parameters.

B) Optimize the coding matrix and return to step A.

- Standard genetic operators are used, scattered crossover and gaussian add unit mutation.

Experiments characteristics

- UCI dataset characteristics.

Problem	#Training samples	#Features	#Classes
Dermathology	366	34	6
Iris	150	4	3
Ecoli	336	8	8
Vehicle	846	18	4
Wine	178	13	3
Segmentation	2310	19	7
Glass	214	9	7
Thyroid	215	5	3
Vowel	990	10	11
Balance	625	4	3
Shuttle	14500	9	7
Yeast	1484	8	10

- Labelled Faces in The Wild : 6144x50, 184 classes (http://vis-www.cs.umass.edu/lfw/).

Results on UCI problems

- As we can see the evolutive minimal performs better than the standard codings.

	Binary Minimal ECOC		Evol. Minimal ECOC		one-vs-all ECOC		one-vs-one ECOC	
Data set	Perf.	Classif.	Perf.	Classif.	Perf.	Classif.	Perf.	Classif.
Derma	96.0±2.9	3	**96.3±2.1**	3	95.1±3.3	6	94.7±4.3	15
Iris	96.4±6.3	2	**98.2±1.9**	2	96.9±6.0	3	96.3±3.1	3
Ecoli	80.5±10.9	3	**81.4±10.8**	3	79.5±12.2	8	79.2±13.8	28
Vehicle	72.5±14.3	2	76.99±12.4	2	74.2±13.4	4	**83.6±10.5**	6
Wine	95.5±4.3	2	**97.2±2.3**	2	95.5±4.3	3	**97.2±2.4**	3
Segment	96.6±2.3	3	96.6±1.5	3	96.1±1.8	7	**97.18±1.3**	21
Glass	56.7±23.5	3	50.0±29.7	3	53.85±25.8	6	**60.5±26.9**	15
Thyroid	**96.4±5.3**	2	93.8±5.1	2	95.6±7.4	3	96.1±5.4	3
Vowel	57.7±29.4	3	**81.78±11.1**	3	80.7±11.9	8	78.9±14.2	28
Balance	80.9±11.2	2	87.1±9.2	2	89.9±8.4	3	**92.8±6.4**	3
Shuttle	80.9±29.1	3	83.4±15.9	3	**90.6±11.3**	7	86.3±18.1	21
Yeast	50.2±18.2	4	**54.7±11.8**	4	51.1±18.0	10	52.4±20.8	45
Rank & #	2.9	**2.7**	**2.0**	**2.7**	2.7	5.7	2.2	15.9

Results on LFW dataset

- In this experiment we can see how evolutionary approaches outperform standard ECOC codings while decreasing the number of classifiers dramatically.

	Binary M. ECOC		GA M. ECOC		one-vs-all		one-vs-one	
Data set	Perf.	#	Perf.	#	Perf.	#	Perf.	#
FacesWild	26.4±2.1	10	**30.7±2.3**	10	25.0±3.1	184	-	16836

Conclusions

- The minimal ECOC represents the lower-bound in terms of number of classifiers.
- The evolutive ECOC optimization obtains comparable results to the standard coding designs (sometimes better) while using far less number of dichotomizers.
- This design is suitable when classifying problems with large number of classes.

Thank you

QUESTIONS?