ELSEVIER
Journal of Food Engineering xxx (2004) xxx–xxx
JOURNAL OF FOOD ENGINEERING
www.elsevier.com/locate/jfoodeng

Feature selection algorithms using Chilean wine chromatograms as examples

N.H. Beltrán [a], M.A. Duarte-Mermoud [a,*], S.A. Salah [a], M.A. Bustos [a]
A.I. Peña-Neira [b], E.A. Loyola [b], J.W. Jalocha [b]

[a] *Department of Electrical Engineering, University of Chile, Av. Tupper 2007, Casilla 412-3, Santiago 6513027, Chile*
[b] *Enology and Agroindustry Department, University of Chile, Av. Santa Rosa 11315, Santiago, Chile*

Received 3 November 2003; accepted 17 May 2004

Abstract

This work presents the results of applying genetic algorithms, in selecting the more relevant features present in chromatograms of polyphenolic compounds, obtained from a high performance liquid chromatograph with aligned photodiodes detector (HPLC-DAD), of samples of Chilean red wines Cabernet Sauvignon, Carmenere and Merlot. From the 6376 points of the original chromatogram, the genetic algorithm is able to select 37 of them, providing better results, from classification point of view, than the case where the complete information is used. The percent of correct classification reached with these 37 features turned out to be 94.19%.

Keywords: Feature selection; Genetic algorithms; Wine classification; Signal processing

1. Introduction

The wine industry has experienced a remarkable growth in the last few years. Chile has not been out of this growing market and has incorporated new technologies in the harvest and also in the wine making process. The control efforts on this process have assured the quality of the resultant product. In this sense the classification methods of the grape variety used to elaborate wine play an important role.

During the last two decades it has been an increasing interest in the use of wine classification techniques that allow classifying the variety of the wine as well as the production place (origin denomination). This classification has been carried out by processing information corresponding to physical features (color, density, conductivity, etc.), chemical features (phenols, anthocians, amino acids, etc. (Peña-Neira, Hernández, García-Vallejo, Estrella, & Suárez, 2000; Marx, Holbach, & Otteneder, 2000)) and organoleptic features (odors, tasting, etc. (Flazy, 2000)). This information has been processed by several techniques, such as statistical methods (discriminant analysis, principal components, Fisher transformation, etc. (Fukunaga, 1990)), artificial neuronal networks (perceptrons, multilayers ANN, ANN with radial basis functions, etc. (Ripley, 1996)]) and genetic algorithms (Goldberg, 1989; Holland, 1992; Michalewicz, 1996; Mitchell, 1996).

In every classification problem the process of feature selection become important because it allows to eliminate the features that can lead to errors (noisy features), to discard those that do not contribute with information (irrelevant features) and to eliminate those that provide the same information that others (redundant features)

---

[*] Corresponding author. Tel.: +56-2-678-4213; fax: +56-2-672-0162.
E-mail address: mduartem@cec.uchile.cl (M.A. Duarte-Mermoud).

0260-8774/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jfoodeng.2004.05.015

(Blum & Langley, 1997). The main advantages of this process are the reduction of the data processing time, decrement in the requirements of data storage space, decreasing in the cost of data acquirement (by the use of specific sensors) and the most important, it allows to select a subset of the original features which contribute with the largest amount of information for a particular problem (reduction in the dimensionality of the input data).

This work presents a methodology for selecting the most important variables, for classification purposes, contained in the information comprised in a polyphenolic chromatograph of wine samples, obtained by a high performance liquid chromatograph with detector of aligned photodiodes, HPLC-DAD. In Section 2 a brief explanation of feature selection methods currently in use is presented. In Section 3 is described the data used for this study, indicating the general way it was generated. In Section 4 the methodology used to perform the feature selection is described and Section 5 shows the results obtained. Finally, in Section 6 the main conclusions about this work are drawn and some remarks about future developments are presented.

2. Feature selection methods

Generally speaking, in the feature selection procedures four basic stages are distinguished (Dash & Liu, 1997):

1. *Generation procedure*: In this stage a possible subset of features to represent the problem is determined. This procedure is done according to one of the standard methods used for this purpose.
2. *Evaluation function*: In this stage the subset of features selected in the previous stage is evaluated according to some function previously defined (fitness).
3. *Stopping criterion*: It is verified if the evaluation of the selected subset satisfies the stopping criterion defined for the searching procedure.
4. *Validation procedure*: In this step it is checked the quality of the selected subset of features, using a prespecified criterion.

The general feature selection process is illustrated in Fig. 1.

The feature selection methods are classified from the point of view of the way in which the new subset to evaluate is generated, leading to three types of methods (Dash & Liu, 1997).

1. *Complete methods*: These methods examine all the possible feature combinations. They are computationally very expensive (search space of order $O(2^N)$ for N features) but they assure to find the optimal feature subset. As examples of these methods it is possible to mention *Branch and Bound* (Narendra & Fukunaga, 1977) and *Focus* (Almuallin & Dietterich, 1992).
2. *Heuristics methods*: They use a search methodology such that it is not necessary to evaluate all the possible feature subsets. Thus a higher speed of the method is reached, since the search space is smaller than in the previous case. These methods do not assure to find the optimal subset. As examples in this category we can mention the methods *Relief* (Kira & Rendell, 1992) and*DTM* (Cardie, 1993).
3. *Random methods*: These methods do not have a specific way of defining the feature subset to be analyzed, but use random methodologies. Thus, a probabilistic search takes place in the feature space. The results using these types of methods will depend on the number of attempts, without assuring that the optimal subset is attained. The methods presented in LVW (Liu & Setiono, 1996) and some using genetic algorithms (Vafaie & Imam, 1994) belong to this kind.

From the evaluation function viewpoint, the feature selection procedures can be classified into two categories (John, Kohavi, & Pfleger, 1994).

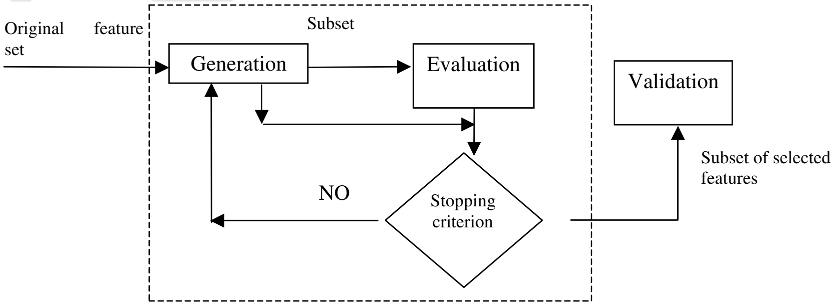

Fig. 1. General procedure of feature selection.

1. *Filtering methods*: These are methods where the selection procedure is made in an independent way of the evaluation function (classification). To this extent is possible to distinguish four different measures: distance, information, dependency and consistency. As an example of these methods we have *Relief* (Kira & Rendell, 1992), *DTM* (Cardie, 1993), *POE* & *ACC* (Mucciardi & Gose, 1971) and *Focus* (Almuallin & Dietterich, 1992) respectively.
2. *Wrapped methods*: In these methods the selection algorithm uses as measure the error rate of the classifier. This method usually provides better results than the previous case, but brings out a large computational cost. In this category we have methods like *Oblivon* (Langley & Sage, 1994).

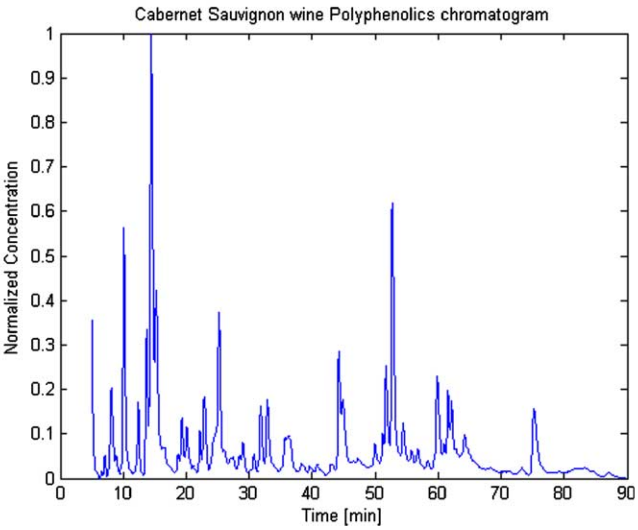

Fig. 2. Typical normalized chromatogram of phenolic compounds of a Chilean Cabernet Sauvignon wine.

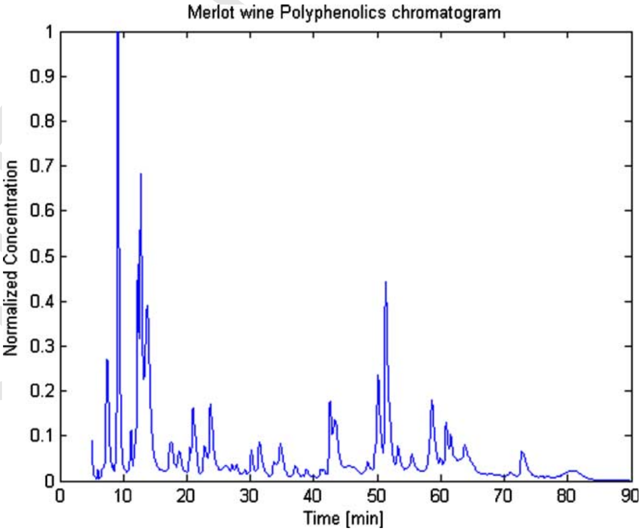

Fig. 3. Normalized chromatogram of phenolic compounds of a typical Chilean Merlot wine.

3. Experimental data

The information used in this study corresponds to that contained in the chromatograms of phenolic compounds of small molecular weight of Chilean red wine samples. They were obtained by means of liquid chromatography analysis using a high performance liquid chromatograph (HPLC) connected to a detector of aligned photodiodes (DAD) (Peña-Neira et al., 2000). The equipment used is an HPLC Merck-Hitachi, model L-4200 UV–VIS Detector with pump model L-600 and columnhold Thermostat. The column used corresponds to a Novapack C_{18}, 300 mm length and 3.9 mm of internal diameter. For the separation of the different phenolics compounds it was used as solvents: A: 98% H_2O, 2% acetic acid; B: 78% H_2O, 20% acetonitrile, 2% acetic acid; C: 100% acetonitrile. The gradient used was: 0–55 min. 100% of A (flow of 1 ml/min); 55–57 min. 20% of A and 80% of B (flow of 1 ml/min); 57–90 min. 10% of A and 90% of B (flow of 1.2 ml/min). Each chromatogram contains 6751 points and each one presents peaks corresponding to a specific phenolic compound. These compounds have been completely studied and identified by agronomists and enologists working in this area (Alamo, 2002; Muñoz, 2002; Peña-Neira et al., 2000).

Typical phenol chromatograms for Chilean Cabernet Sauvignon, Merlot and Carmenere, are shown in Figs. 2–4, respectively.

In order to avoid distortions, before processing the data contained in the chromatograms the information was normalized, since the amplitude of the peaks depends on the volume of the wine injected into the chromatograph. In some cases were injected 20 ml whereas in other cases were injected up to 100 ml of prepared sample and consequently the peak amplitude (corresponding to each component concentration) have different magnitudes. For the normalization process, in each chromatography was subtracted the minimum value and the result was divided by the maximum value, obtaining final values in the interval 0 and 1. This allows a fair comparison of different chromatographs.

The first five minutes of each chromatogram were deleted (375 points) analyzing only the 6376 remaining points. This is done since during the first instants of time appears information in the chromatogram regarding the effluents used to obtain the liquid chromatography and they are chemical compounds which are not present in the wine sample.

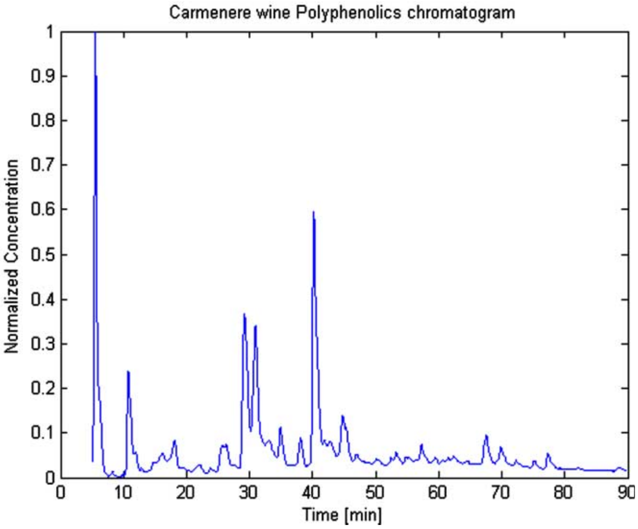

Fig. 4. Typical normalized chromatogram of phenolic compounds of a Chilean Carmenere wine.

Table 1
Distribution of samples used in the study

Wine	Sample no.	Number of samples
Cabernet Sauvignon	1–80	80
Carmenere	81–137	57
Merlot	138–172	35
	Total	172

For this study it was used data of Chilean wines from different valleys, years and different vintages, which are summarized in Table 1.

4. Proposed methodology

Due to the large number of features contained in each sample (6376), the search of those features providing the largest amount of information for classification problems may become a large computational process with too many calculations. On the other hand the number of wine samples used in this work is relatively low to apply some classifiers, since the number of features by sample (6376) is far larger than the total number of samples (172). This disadvantage is avoided by the genetic algorithm used in this work, since the methodology uses a parallel processing of the possible solutions.

Genetic algorithms (GA) (Holland, 1992) correspond to one of the techniques known under the name of evolutionary computing, which is inspired in the concept of biological evolution. The main idea is that each individual of a population represents a possible solution of the optimization problem to be solved and according to the adaptation of each individual to the problem under consideration (measure given by an objective function), the evolution of the population will take place. Thus, new possible solutions are generated corresponding to the result of recombination processes (crossover) of the different initial individuals. New solutions can also be considered by introducing mutations on the existing individuals.

There are different ways of applying genetics algorithms. According to the selection method used to select the individuals leading to the next generation, the most common are the methods of proportional selection (Goldberg, 1989), selection by tournaments (Michalewicz, 1996) and selection by ranking (Mitchell, 1996). It can also be used different crossover methodologies (where two new individuals are created from two parents) and different mutation processes (where existing individuals are randomly modified).

In order to determinate the best feature subset a genetic algorithm of niches was used, particularly one denominated "Deterministic Crowding" (Mahfoud, 1995), which allows to find multiples solutions. In this particular case the use of this method is in particular attractive, since obtaining a feature subset, that corresponds to a group of specific phenolics compounds of the wines, might be experimentally easier to determinate and could have a similar performance, from classification point of view, that another subset of phenols with harder experimental difficulties.

In order to use the genetic algorithm there is a sequence of steps that is described in what follows.

Individual codification: In making each individual codification, that corresponds to a possible feature subset, binary strips of length N are used, where N corresponds to the number of original features describing the problem. In this case each feature corresponds to a point of the chromatogram and $N = 6376$. The existence of a 1 in the ith position indicates that feature i of the sample should be considered in the classification process.

In order to start the algorithm, the number of individuals N_i that will be considered in each generation must be defined. This definition will have a direct relationship with the amount of calculations to be performed and with the convergence of the algorithm to an optimal solution. Given a larger number of individuals, a wider exploration will be made in each generation, needing a larger number of calculations. In the case of a simple genetic algorithm not of a niche type, this number N_i will also have relation with the selective pressure that can be measured according to the number of generations that the best individual of a population takes to fill out all the population with copies of itself.

In the particular case of the methodology used here, the number of individuals will be $N_i = 150$ and it was defined arbitrarily. According to the evaluation function considered in the method, the amount of calculations to determine the fitness of each individual will vary. Thus the algorithm could take larger or smaller time in making the calculations for each generation depending on different evaluation functions.

Another factor to be considered is the "genetic diversity" that is introduced in the initial population. For example, if in a feature selection problem with $N=50$ we use $N_i=4$ individuals representing subsets of at most 10 features each one, in the best case (if all the subsets were disjoint) the algorithm will consider 40 features in the search (10 for each individual) being 10 features not considered in the search space. That is why the larger is the initial population the larger will be the genetic diversity. If in the same previous example $N_i=10$ individuals were considered and their features were chosen randomly, the probability that all the features were included in the search space would be very high. Another tool that can be used to introduce this genetic diversity is the mutation.

Once defined the number of individuals N_i of the population, the initial features of individuals must be defined. This was made in a random form, in such a way that each individual did not include a feature number greater than 80 ($N_c \leqslant 80$). This is because for this particular case there were only 172 samples and having a larger number of features than the number of samples, the Linear Discriminant Analysis (LDA), the classifier selected for this study, cannot be used.

Adaptation function (*fitness*): In order to define the adaptation function or fitness, for each individual and since the aim is to select those features providing more information, the performance of an LDA classifier was used, which corresponds to the linear Fisher classifier (Fukunaga, 1990). This classifier uses the linear Fisher Transformation that maximizes the distance between classes and minimizes the inter classes distance. In addition, a leave-one-out methodology was considered, consisting in designing the classifier using all the samples except one and using the sample that was left out in classification. This procedure is performed excluding all the samples, one by one, and determining the error computed as the number of samples classified wrong divided by the total number of samples. In this study this form was chosen since the number of samples was not sufficiently large and is the limit of the cross-validation methodology when the sets are set to $N-1$ for training and 1 for validation. The methodology corresponds to a wrapped methodology, since the right classification percentage of the classifier is used as a performance measurement.

Next generation selection methodology: With the aim of selecting the population of the next generation, the denominated Deterministic Crowding (Mahfoud, 1995) was used. It consists in making a random selection of two parents, allowing every individual in the population being selected as father only once, so that every individual of a generation can be considered as parent for the next generation. In the following generation the parents recombine themselves in discrete form (Uniform Crossover), that is to say for every variable of each individual of the intermediate population the variable belonging to one of the parents is chosen randomly and equiprobably. The intermediate population corresponds to a population of individuals which are possible candidates to be considered as part of the following generation and is generated in the middle of generation n and $n+1$. Each pair of parents will arise two individuals of the intermediate population, which are evaluated in the sense of similarity with the parents, using the Hamming distance of the individuals (Mahfoud, 1995). For the next generation are chosen the two individuals of best performance in the comparisons between parents and individuals of the intermediate population. This assures the continuity of the different possible feature subsets solving the problem, since if exist individuals considering similar feature subsets (i.e. their Hamming distance is small) these will be compared amongst them and not with individuals considering extremely different feature subsets.

5. Classification results

Applying the methodology explained in Section 4 to the information described in Section 3, a series of results were obtained, which are described in what follows.

Initially the algorithm was run to a point in which an 88.4% of correct classification was obtained (20 samples incorrectly classified), considering only 64 out of the 6376 features. At that moment was not possible to continue the execution of the algorithm, because the feature subsets found have a non invertible correlation matrix, indicating that those features were linearly dependent. In order to solve this problem, whenever a situation like this was found the subset was eliminated, because is of no interest to find feature subsets considering correlated features, see Fig. 5.

With these considerations in mind a result considering only 36 features was obtained, providing a correct classification percentage of 91.86% (14 samples wrong

Charact N°1 Charact N°6376

1	0	0	1	1	0	...	1	1	0	0	1	0

Fig. 5. Example of an individual in the population.

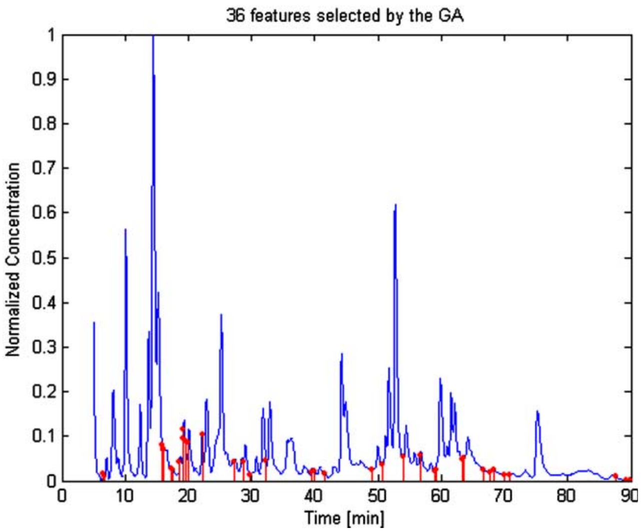

Fig. 6. The 36 best features selected by the GA plotted on the chromatograph for sample no. 1, providing a correct classification percentage of 91.86%.

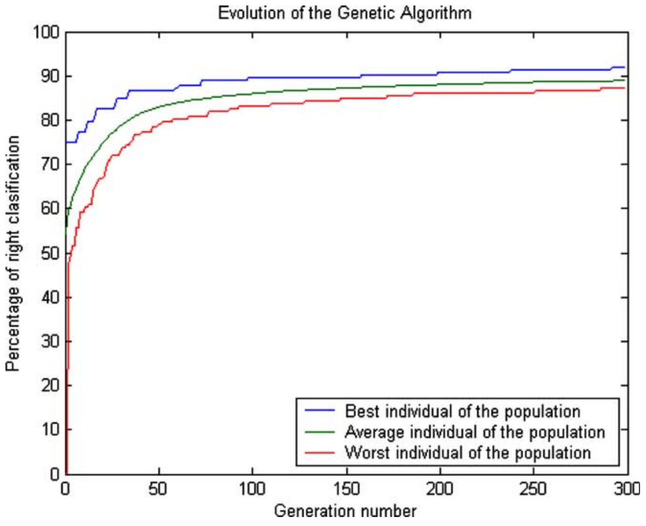

Fig. 8. Evolution of the percentage of correct classification as function a of the generations.

classified). In Fig. 6 the 36 features selected by the genetic algorithm are shown by vertical lines plotted on the chromatograph corresponding to sample no. 1 (a Cabernet Sauvignon sample wine). Another possible solution found applying this methodology was two subsets of 35 features giving a correct classification percentage of 91.27%. One of the 35 selected features by the GA are indicated in Fig. 7 and plotted over the chromatograph corresponding to sample no. 1.

It is important to point out that once the GA selected the 35 features (or 36), these 35 (or 36) points were considered in each chromatogram (172) and then the Leave One Out (LOO) procedure was used to evaluate the classifier (using an LDA classifier). 172 test were performed, each time leaving one sample out of the total set and training the classifier (LDA) with 171 remaining samples. Then the sample left out was presented to the classifier for classification in one of the three classes. In this procedure (repeated 172 times) all the samples but a few ones presented to the classifier were correctly classified as Merlot, Carménére or Cabernet Sauvignon, giving an average correct classification rate.

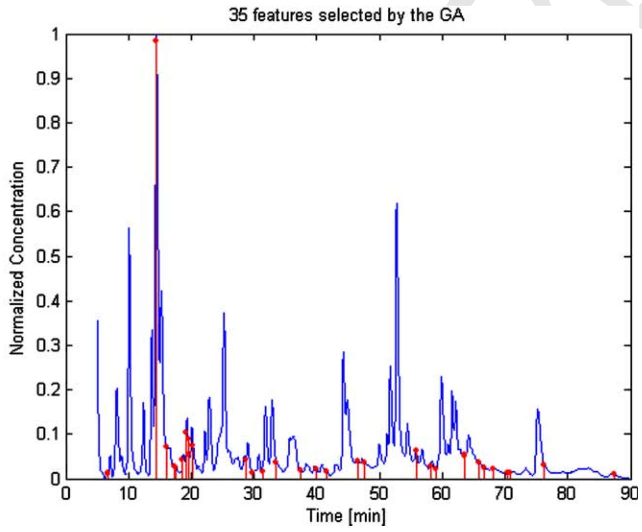

Fig. 7. The 35 best features selected by the GA plotted on the chromatograph for sample no. 1, providing a correct classification percentage of 91.27%.

In Fig. 8 it is possible to appreciate the evolution of the genetic algorithm and how the performance improves as the generations increase. The upper curve corresponds to the percentage of correct classification of the best individual of each generation, the second curve corresponds to the average of correct classification of the whole population, and the lower curve corresponds to the percentage of correct classification of the worse individual of each generation.

Fig. 9 displays an histogram appearing the frequency that each feature is present in the solution that gives the best classification percentage in each generation, considering the first 300 generations. It is possible to observe that some features never appear in the solution and it is may be conjectured that these features does not contain important information to the wine classification problem.

As it was mentioned before, the benefit of using a niche type genetic algorithm is that it is possible to find more than one feasible solution to the optimization problem. It is important to point out, that due to the choice of the size population (150) and to the large number of features (6376), perhaps not all the subset were considered for the search of the optimal or maybe they were eliminated early. To avoid this situation a new population was defined in a random way, where the

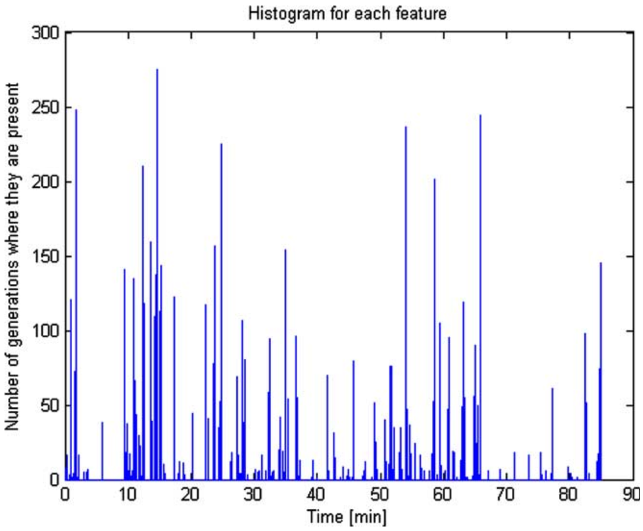

Fig. 9. Number of opportunities that each feature appears in the best subset found by the GA, for the first 300 generations.

three best individuals of the previous simulations were included (best individuals obtained after 300 generations, since as can be appreciated from Fig. 8, an increase of the generations beyond 300 does not improve the percentage of correct classification). The same effect could have been obtained using mutation of the population. After this process, a new possible feature subset was found with 37 features giving a 94.19% of correct classification rate (10 samples incorrectly classified).

Fig. 10 shows the subset including 37 features obtained after including the three best individuals of the first simulation in a new population, obtaining a different feature subset that gives a correct classification percentage of 94.19%. In Fig. 10 the vertical lines correspond to the features selected by the GA and they are plotted on the chromatograph corresponding to sample no. 1 (a Cabernet Sauvignon sample wine).

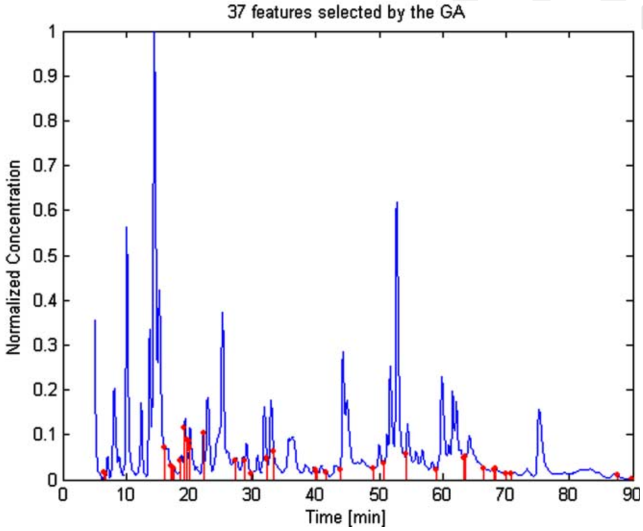

Fig. 10. The 37 best features selected by the GA plotted on the chromatograph for sample no. 1, providing a correct classification percentage of 94.19%.

Table 2
Samples wrong classified for the case of 37 features selected by the GA

Sample no.	Real variety	Classification given by classifier
71	1	2
128	2	1
134	2	3
135	2	1
137	2	1
138	3	1
139	3	1
154	3	1
162	3	1
172	3	1

1=Cabernet Sauvignon, 2=Carménére, 3=Merlot.

Table 3
Confusion matrix for the classifier with 37 features

Real sample/classified sample	1	2	3
1	0.9875	0.0125	0
2	0.0526	0.9298	0.0175
3	0.1429	0	0.8571

1=Cabernet Sauvignon, 2=Carménére, 3=Merlot.

The 10 cases that were wrong classified, using the leave-one-out methodology, by the classifier considering only 37 features, occurred when the classifier is trained with all the samples but the samples given in Table 2 and later, when the samples are presented to the classifier those samples were wrongly classified as indicated in Table 2. This situation is summarized in the confusion matrix shown in Table 3.

6. Conclusions

A methodology of feature selection based on genetic algorithms has been proposed in this paper for wine classification purposes.

It was demonstrated that the application of this methodology to Chilean wine variety classification problems gives percentages of correct classification of 94.19%.

From the results presented in this work it is possible to choose a small feature subset of the original features (6376) to suitably discriminate the classes of the Chilean wine samples. In this case 37 features were selected, each one corresponding to a chemical compound, that contain the best information to differentiate between the wine samples with a percentage of correct classification of 94.19%. That is to say, with only the 0.58% of the original features it is possible to reach percentage of right classification of 94.19%.

Another interesting point arising from this study is the fact that several different feature subsets can be found, providing the same percentage of right classification. The features selected by the GA corresponding to one possible solution, will give information to enologists as to what compounds are more relevant for wine classification purposes.

From this study arise the necessity of having a larger number of samples to generalize and validate these results. Currently we are processing about 200 new samples of Chilean wines, including some of the 2003 vintage, to increase the number of samples of our data base to a total of about 350.

An interesting alternative to the proposed methodology is to incorporate an objective function penalizing the number of features considered in the solution subset. Thus, besides considering in the solution the correct classification percentage will also be considered the smallest number of features satisfying the selection objective.

Acknowledgment

The results reported in this paper have been supported by CONICYT-Chile, through the grant FONDEF D01-1016 “Identificación varietal de vinos chilenos mediante instrumentación inteligente”.

References

Alamo, V. S. (2002). Characterization of phenolic compounds in 2002 commercial merlot and sauvignon blanc from five Chilean valleys. Agronomy Engineer Thesis, Faculty of Agronomical Sciences, University of Chile.

Almuallin, H. & Dietterich, T. G. (1992). Learning with many irrelevant features. In *Proceedings of Ninth National Conference on Artificial Intelligence* (pp. 547–552). MIT Press, Cambridge, MA, USA.

Blum, A. L., & Langley, P. (1997). Selection of relevant features and examples in machine learning. *Artificial Intelligence, 97*(1–2), 245–271.

Cardie, C. (1993). Using decision trees to improve case-based learning. In *Proceedings of Tenth International Conference on Machine Learning* (pp. 25–32). Morgan Kaufmann Publishers, University of Massachusetts, Amherst, USA, June 1993.

Dash, M., & Liu, H. (1997). Feature selection for classification. *Intelligent Data Analysis, 1*(3), 131–156.

Flazy, C. (2000). *Enología: Fundamentos científicos y tecnológicos*. Madrid, Spain: Mundi Prensa.

Fukunaga, K. (1990). *Introduction to statistical pattern recognition* (second ed.). San Diego, USA: Academic Press.

Goldberg, D. (1989). *Genetic algorithms in search, optimization and machine learning*. New York, USA: Addison-Wesley.

Holland, J. H. (1992). *Adaptation in natural and artificial systems* (second ed.). Cambridge, MA, USA: MIT Press.

John, G. H., Kohavi, R. & Pfleger, P. (1994). Irrelevant features and the subset selection problem. In *Proceedings of the Eleventh International Conference on Machine Learning* (pp. 121–129). New Brunswick, Morgan Kaufmann, USA, 1994.

Kira, K. & Rendell, L. A. (1992). The feature selection problem: Traditional methods and a new algorithm. In *Proceedings of Ninth National Conference on Artificial Intelligence* (pp. 129–134). MIT Press, Cambridge, MA, USA.

Langley, P. & Sage, S. (1994). Oblivious decision trees and abstract cases. In *Working Notes of the AAAI94 Workshop on Case-Based Reasoning* (pp. 113–117). Seattle, WA, USA. AAAI Press, 1994.

Liu, H. & Setiono, R. (1996). Feature selection and classification—A probabilistic wrapper approach. In *Proceedings of Ninth International Conference on Industrial and Engineering Applications of AI and ES* (pp. 419–424). Fukuoka, Japan, June 1996.

Mahfoud, S. W. (1995). Niching methods for genetic algorithms. Ph.D. Thesis, University of Illinois at Urbana-Champaign, Illinois Genetic Algorithms Laboratory (IlliGAL) Report No. 95001, May 1995.

Marx, R., Holbach, B. & Otteneder, H. (2000). Determination of nine characteristics anthocyanins in wine by HPLC. Off. Int. Vigne Vin. Bulletin. Paris. August 2000. F.V. No. 1104 2713/100200.

Michalewicz, Z. (1996). *Genetic algorithms + Data structures = Evolution programs* (third ed.). New York, USA: Springer-Verlag.

Mitchell, M. (1996). *An introduction to genetics algorithms*. Cambridge, MA, USA: MIT Press.

Mucciardi, A. N., & Gose, E. E. (1971). A comparison of seven techniques for choosing subsets of pattern recognition. *IEEE Transactions on Computers, 20*(9), 1023–1031.

Muñoz, L. P. (2002). Characterization of phenolic compounds in 2002 commercial cabernet sauvignon and chardonnay from five Chilean valleys. Agronomy Engineer Thesis, Faculty of Agronomical Sciences, University of Chile.

Narendra, P. M., & Fukunaga, K. (1977). A branch and bound algorithm for feature selection. *IEEE Transactions on Computers, 26*(9), 917–922.

Peña-Neira, A. I., Hernández, T., García-Vallejo, C., Estrella, I., & Suárez, J. (2000). A survey of phenolic compounds in Spanish wines of different geographical origins. *European Food Research and Technology, 210*, 445–448.

Ripley, B. D. (1996). *Pattern recognition and neural networks*. Cambridge, UK: Cambridge University Press.

Vafaie, H. & Imam, I. F. (1994). Feature selection methods: Genetic algorithm vs. greedy-like search. In *Proceedings of the 3rd International Fuzzy Systems and Intelligent Control Conference*. Louisville, KY, March 1994.