Methodology for the recognition and diagnosis of student performance by discriminant analysis and artificial neural networks

J. Sánchez-Monedero, P.A. Gutiérrez, C. Hervás-Martínez, M. Cruz-Ramírez, J.C. Fernández-Caballero, F. Fernández-Navarro

Department of Computer Science and Numerical Analysis, University of Córdoba, Spain

emails: i02samoj@uco.es, pagutierrez@uco.es, chervas@uco.es, i42crram@uco.es, jcfernandez@uco.es, i22fenaf@uco.es

Abstract

This paper describes a methodology for classifying students through their performance and their opinions. The performance was evaluated by the results of three different tests and the opinions were collected in two surveys that were related to the teaching-learning process, all with respect to the same subject. The dataset was composed of 587 students belonging to 10 different educational centres. The samples were obtained during two consecutive academic years and a clustering process was applied to group students. Once the groups were obtained, a supervised classification algorithm was applied. For this classification, we decided on a stepwise discriminant analysis and a feed forward artificial neural network with sigmoidal basis functions. The architecture and weights were tuned through the use of an evolutionary algorithm. An analysis of the most relevant features of the three student classes considered was performed by studying the discriminant functions learned by the neural network. The proposed methodology could be introduced in any educational environment and is valid for any feature set describing educational performance and opinion.

Key words: student performance, statistical models, neural networks, expert systems, evolutionary programming

1. Introduction

Every teaching professional is concerned not only about giving lessons and evaluating the students' knowledge about these lessons, but also about offering the proper help to these students.

However, when a teacher has lots of students in the same course, it is very difficult to get to know them well enough to detect who needs more or less help and of which type. In the best scenario, the teacher gives a control test at the beginning of the course aimed to check the students' general level. This is not only insufficient but also leads to unrealistic conclusions in most cases.

Thus arises the need to develop a system which can help the teacher to classify his students according to their needs. Logically, this classification should be done considering the knowledge, ability and motivation estimated for each student. A system of these characteristics could be aimed at forecasting this classification based on academic performance and different opinion surveys conducted during the year. The classifications performed by the system could help the teacher to decide how to treat each group or student before the end of the course.

In the pattern recognition context, a dataset with class labels is needed to classify students in different groups. However, assigning these labels based on the above-mentioned information is not easy, because the distribution of the possible classes that we want to recognize is unknown a priori. So it is very common to apply unsupervised algorithms, also known as clustering algorithms, which are able to construct groups of training patterns using their relative distances. Once these groups are built, a classification algorithm can be used.

This paper introduces a methodology for classifying secondary school students by combining data from two sources: a) the partial student performance results recorded at different moments during the year and b) two opinion surveys. The students were described by 52 variables, which include: sex, age, the previous year's grade, the first evaluation mark, second evaluation mark, final mark and 46 variables related to the two surveys (Initial and Final) that were conducted during the academic year.

To address the construction of the student groups, we applied a cluster analysis algorithm to decide the optimal number of classes. We concluded that there are three well distinguished classes by using the *k*-means algorithm [1] and different values for *k*. Then, linear (such as discriminant analysis) and non-linear (such as artificial neural networks (ANNs) [3]) classification methods were used in order to analyse the suitability of the groups discovered and the possibility of an automatic classification system.

The main objectives that we wanted to achieve in this work were the following:

- To be able to classify the students in groups regarding their needs in order to pay greater attention to certain groups.
- To select the minimum number of relevant features that warranties a correct classification rate closer to 90%.

- To analyse this feature selection in order to redesign the surveys by considering the variables which best discriminate students.

By using the results and conclusions of this paper, the final purpose is to implement a customized expert system [2] that will provide the teacher with the status, possible problems, level and ability of each student. Thereby, the teacher will have the possibility of guiding and recovering the students. The system would construct the classifier by using the opinions and partial results observed in the middle of the course and it could even have the capability of using new samples to automatically correct models year by year.

The results achieved with this methodology could be extended to transnational education for example in the framework of international education programs. Particularly, it should be interesting to study the relevant features (variables) extraction done by the algorithm and compare it by considering different countries.

This paper is organized as follows: Section 2 describes the non-linear statistical technique considered for classifying students (Evolutionary Neural Networks); Section 3 describes the experimental methodology undertaken to achieve the previously defined objectives; a summary of the results is given in Section 4 and, finally, Section 5 includes the conclusions of the work.

2. Non-linear classifier: Evolutionary Neural Networks

In our problem, the student's information is composed of some subjective variables (the results of the opinion surveys) and other objective variables (academic performance results). This fact made us consider the use of multilayer perceptron (MLP) neural networks as an alternative non-linear classifier to discriminant analysis methods. The MLP does not need a priori normality assumptions, or hypotheses about the independence of the variables or about their nature.

In the area of Artificial Neural Network (ANN) design [4], one of the main problems is to find suitable architectures to solve specific problems. The selection of such architecture is very important: a smaller network than needed would be unable to learn and a larger network than needed would end in over-training. The problem of finding a suitable architecture and the corresponding weights of the network is a very complex task (for a very interesting review of the matter the reader can consult [5]).

The automatic design of ANNs involves two basic concepts: parametric learning and structural learning. In parametric/structural learning, both architecture and parametric information must be optimized through the training process. Basically, we can consider three models of structural learning: constructive algorithms, destructive algorithms, and evolutionary computation. Both methods, constructive and destructive, limit the number of available architectures, thus introducing constraints in the search space of possible structures that may not be suitable to the problem.

Evolutionary computation has been widely used in recent years to evolve neural-network architectures and weights. There have been many applications for parametric learning [6] and for both parametric and structural learning [7]-[11]. These works fall into two broad categories of evolutionary computation: genetic algorithms and evolutionary programming. Evolutionary programming [12] is, for many authors, the most suited paradigm of evolutionary computation for evolving ANNs [7]. Evolutionary programming uses a natural representation for the problem. Once the representation scheme has been chosen, mutation operators specific to the representation scheme are defined. The use of evolutionary learning for designing neural networks dates from no more than two decades ago (see [13] for a review). However, a lot of work has been carried out in these two decades, resulting in many different approaches, for instance, [5], [14]. The main advantage of evolutionary computation is that it performs a global exploration of the search space avoiding entrapment in local minima as is common in local search procedures.

In this work we used an Evolutionary Algorithm (EA) to estimate the parameters and the structure of the MLP models. The objective is to design a neural network with optimal structure and weights for the classification problem tackled. The population is subject to the operations of replication and mutation and crossover is not used due to its potential disadvantages in evolving artificial networks [7]. With these features, the algorithm falls into the class of evolutionary programming [10]. The algorithm considered is the Neural Net Evolutionary Programming (NNEP) algorithm proposed in different previous works [15]. NNEP is a software package developed in JAVA by the authors, as an extension of the JCLEC framework [16] and is available in the non-commercial JAVA tool called KEEL [17].

3. Experimental methodology

The student dataset is composed of data collected from 587 students found in samples taken during two consecutive years in 10 non-university educational centres. The features used for classification purposes have been: academic performance, ability variables, motivation and methodological and evaluation preferences for the subject. Academic performance has been obtained by evaluating students' tests. The rest of the variables were obtained via an initial survey, done in November, and a posterior survey in May. These surveys are composed of grouped items in an ordinal scale with 5 levels. The item blocks are related to the previously mentioned features. In the final survey, there is a new item group where the student can evaluate the quality of the teaching received, teacher methodology and the evaluation performed.

As previously mentioned, a cluster analysis was done by using a k-means algorithm to obtain the groups of students. Different values of k were tested and the results pointed out that the optimal number of groups is 3. By using these groups to assign a class label to each student, we applied two supervised classification algorithms: discriminant analysis and evolutionary neural networks. The next section includes an analysis of the classes obtained and the classification of both methods by considering these classes or groups.

4. Analysis of the results

First of all, the characteristics of the different classes, obtained by the *k*-means algorithm, are:

- Class 1 is composed of 229 students with very low academic performance (marks lower than 5 in a 1 to 10 scale), very little study motivation, more prone to practical activities and to the use of multimedia technologies and computers as educational resources. Students in this class prefer a more participative methodology which evaluates practical exercises, their interest and their effort. They find the quality of the teaching received acceptable. The students consider that the evaluation tests have been suitable and feel that the teacher has always been available for solving issues. Finally, they do not have a good study environment.
- Class 2 is made up of 210 students with higher academic performance (with average marks between 7 and 8). They are very motivated to study. Students in this class prefer the classical lectures given by the teacher, although they agree with the use of newer educational resources. They are indifferent to a participative methodology. They would like to be evaluated considering practical activities, effort and interest, but they consider that they are fairly evaluated. They think that the teaching received is suitable and they value their teacher's work positively. The students within this class are the most satisfied about their evaluation and consider that they could have consulted the teacher without any difficulty. They have a good family environment and a good study plan.
- Class 3 is composed of 148 students with acceptable academic results (average marks between 5 and 6). This group is less motivated with respect to study and the subject. They do not prefer lecture lessons but they also do not like multimedia and computer educational resources. These students also refuse to work in teams. They are not comfortable with continuous student assessment, with respect to either effort or interest. On average, they feel fairly evaluated and they agree with the evaluation method. This is a group with an intermediate study environment who also lack skills in study strategies.

For the second part of the experiment, we applied different techniques to classify new students into one of the three groups obtained by *k*-means. The experimental design consisted of a hold-out cross-validation procedure where 80% of the patterns was for composing the training data set and the other 20%, for the generalization set.

A stepwise lineal discriminant analysis was applied using the 52 features pertaining to each of the students. The correct classification ratios using this discriminant analysis were: 91.7%, 86.7% and 90.5%, for Class 1, 2 and 3, respectively. The total number of incorrectly classified students was 27 for Class 1, 14 for Class 2 and 13 for Class 3; this leads us to conclude that the discriminant model was well adjusted since it presented 13% of total errors. The two obtained discriminant functions were composed of the following variables (sorted according to their discriminant capability):

1. Final mark.
2. Question #14 of the Initial Survey: I think class work and activities should be highly considered in the evaluation.
3. Question #10 of the Final Survey: The teacher must also evaluate the student's interest and effort.
4. Question #7 of the Final Survey: It is difficult for me to pass this subject.
5. Question #16 of the Initial Survey: The teacher must also evaluate the student's interest and effort.
6. Question #1 of the Final Survey: The teacher explains the lessons clearly.
7. Question #15 of the Final Survey: I prefer to study the subject by working in teams with other classmates.
8. Question #17 of the Final Survey: The marks obtained are generally in accord with one's knowledge.
9. Question #1 of the Initial Survey: I am studying because I like it very much.
10. Average mark in the last year.
11. Question #11 of the Final Survey: I think that student work in class and other activities should be highly evaluated.
12. Question #11 of the Initial Survey: I like the teacher's evaluation method.
13. Question #17 of the Initial Survey: I have a good place to study and suitable study material.
14. Question #7 of the Initial Survey: Multimedia and computer resources should be used in the subject methodology.
15. Sex of the student.
16. First evaluation mark.
17. Question #16 of the Final Survey: More activities and practical work should be included.
18. Question #18 of the Final Survey: I consider that I have been fairly evaluated to date.
19. Question #8 of the Final Survey: I find the subject difficult to study.
20. Question #4 of the Final Survey: The teacher makes it "easy" to study the subject.
21. Question #3 of the Initial Survey: I am studying because I would like to go to University.

By using Evolutionary Neural Networks, we obtained a model with three layers. A data pre-processing technique eliminated 7 variables from the initial 52. Then, the input layer included 45 neurons. The second layer, i.e. the hidden layer, was composed of 24 neurons with sigmoidal functions as basis functions. The output layer was made up of 2 neurons, this is, the number of classes obtained by the cluster analysis minus one. Input data was scaled between -1 and 1. 92% of the generalization set, 117 students, were well classified by this network model.

Analysing the best neural network model, the variables that finally compose the discriminant functions are significantly lower than the original 52. Thus, the most relevant features regarding academic performance are: the final subject mark, closely followed by the mean mark from the previous year and the first evaluation mark. It is worthwhile to mention the low discriminant power of the second evaluation mark. Regarding the opinions collected in the first survey, the most discriminant variables are those related to the evaluation methods used by the teacher. On the other hand, the discriminant capability of the items in the final survey is greater than the items in the initial survey. This fact assigns more importance to the opinions in the final survey, which makes sense, since the students are more aware of the complexity of the subject and their own academic performance.

5. Conclusions

Considering the results, we observe that both classification models are highly robust, with correct classification ratios of nearly 90%. The advantage of the statistical methods is the simplicity of the programming of decision functions whereas the artificial neural networks methods achieve greater accuracy in classifying students with inconsistent features and, therefore, their generalization ability is greater. The analysis of the models attributes more importance to the features related to the Final Survey and to the Final Mark obtained in the subject.

It is important to point out that this is a preliminary study and that our final objective is to build an expert system which can be used as a help and resource for the teacher in order to detect students with special needs. The construction of this system could be achieved by applying a similar methodology (*k*-means plus a classification technique) but not including those variables related to the final part of the course (Final Mark and Final Survey) in the classifier. In this way, the classifier could be applied in subsequent academic courses for predicting the group of a student before the end of the course.

As future work, the proposed methodology could be applied to higher education centres. Since the new trends in higher education, i.e. the Bologna process, remarks the need of properly tracking students learning process, this methodology may be especially suitable in this case. Note that the methodology is flexible enough so that it could be extended by adding more surveys or other data sources.

Acknowledgements

This work was partially subsidised by the Spanish Inter-Ministerial Commission of Science and Technology (MICYT) under Project TIN 2008-06681-C06-03, the FEDER funds, and the P08-TIC-3745 Project of the "Junta de Andalucía" (Spain).

6. References

[1] A. JAIN AND R. DUBES, *Algorithms for clustering data*, Englewood-Cliffs. N. Jersey, Prentice-Hall, 1988.

[2] C. HERVAS, J. FEDRIANI, N. GARCIA AND E. ROMERO, *Sistema experto en evaluación docente*, Congreso Iberoamericano de informática educativa, Santo Domingo (República Dominicana), 2, (1992) 282-289.

[3] C. HERVAS, E. ROMERO, N. GARCIA AND R. MEDINA, *Comparison between artificial neural network and classical statistical methods in pattern recognition*, IPMU, Palma de Mallorca, (1992) 6-10.

[4] S. HAYKIN, *Neural Networks—A Comprehensive Foundation*, New York: Macmillan, 1994.

[5] X. YAO AND Y. LIU, *A new evolutionary system for evolving artificial neural networks*, IEEE Trans. Neural Networks, 8, (1997), 694-713.

[6] A. J. F. VAN ROOIJ, L. C. JAIN, AND R. P. JOHNSON, *Neural Networks Training Using Genetic Algorithms*, Singapore: World Scientific, 1996, vol. 26, Series in Machine Perception and Artificial Intelligence.

[7] P. J. ANGELINE, G. M. SAUNDERS, AND J. B. POLLACK, *An evolutionary al gorithm that constructs recurrent neural networks*, IEEE Trans. Neural Networks, vol. 5, pp. 54-65, Jan. 1994.

[8] V. MANIEZZO, *Genetic evolution of the topology and weight distribution of neural networks*, IEEE Trans. Neural Networks, vol. 5, pp. 39-53,. 1994.

[9] B. A. WHITEHEAD AND T. D. CHOATE, *Cooperative-competitive genetic evolution of radial basis function centers and widths for time series prediction*, IEEE Trans. Neural Networks, 7, 1996.

[10]X. YAO AND Y. LIU, *Making use of population information in evolutionary artificial neural networks*, IEEE Trans. Syst. Man, Cybern. B, vol. 28, pp. 417-425, June 1998.

[11]D. E. MORIARTY AND R. MIIKKULAINEN, *Forming neural networks through efficient and adaptive coevolution*, Evol. Comput., vol. 4, no. 5, 1998.

[12]D. B. FOGEL, *Evolving artificial intelligence*, Ph.D. dissertation, Univ. California, San Diego, CA, 1992.

[13]J. D. SCHAFFER, L. D. WHITLEY, AND L. J. ESHELMAN, *Combinations of genetic algorithms and neural networks: A survey of the state of the art*, in Proc. Int. Workshop Combinations Genetic Algorithms Neural Networks, L. D. Whitley and J. D. Schaffer, Eds., Los Alamitos, CA, 1992, pp. 1-37.

[14]S. V. ODRI, D. P. PETROVACKI, AND G. A. KRSTONOSIC, *Evolutional development of a multilevel neural network*, Neural Networks, vol. 6, pp. 583-595, 1993.

[15]F. J. MARTÍNEZ-ESTUDILLO, C. HERVÁS-MARTÍNEZ, P. A. GUTIÉRREZ-PEÑA, A. C. MARTÍNEZ-ESTUDILLO, *Evolutionary Product-Unit Neural Networks for Classification*, Neurocomputing. 72, 549-562. (2008).

[16]S. VENTURA, C. ROMERO, A. ZAFRA, J.A. DELGADO, AND C. HERVÁS-MARTÍNEZ, *JCLEC: a JAVA framework for evolutionary computation*, Soft Computing 12(4) (2007) 381-392.

[17]J. ALCALA-FDEZ, L. SÁNCHEZ, S. GARCÍA, M.J.D. JESUS, S. VENTURA, J.M. GARRELL, J. OTERO, C. ROMERO, J. BACARDIT, V.M. RIVAS, J.C. FERNÁNDEZ, AND F. HERRERA, *KEEL: A Software Tool to Assess Evolutionary Algorithms for Data Mining Problems*, Soft Computing, Volume 13, Issue 3, 2009, Pages 307-318