ESANN'2004 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 28-30 April 2004, d-side publi., ISBN 2-930307-04-8, pp. 263-268

Comparison of different classification methods on castabilty data coming from steelmaking practice

Marco Vannucci, Valentina Colla

PERCRO - Scuola Superiore Sant'Anna
Polo Sant'Anna Valdera, Viale R. Piaggio 34, Pontedera, Pisa, Italy
e-mail: mvannucci{colla}@sssup.it

Abstract. The problem of the prediction of a critical situation during continuous casting in common steelmaking practice is faced through different traditional soft–computing techniques: the task is to divide the data in two classes corresponding to good and bad casting behaviour respectively. Moreover, a novel algorithm, the Model–based method, is presented. The performance obtained by the different techniques on the real data coming from an important steelmaking industry are compared.

1 Introduction

In common steelmaking practice, during continuous casting, the liquid material produced in the blast furnace is cast, after some manufacture, into the ladle and, subsequently, into the tundish. On the bottom of the tundish, the submerged entry of some nozzles is located, through which the liquid steel passes into the mould or strip casters. As it is shown by Fig. 1, the section of such nozzles is far smaller with respect to the tundish dimensions. When Aluminium killed steels are produced, alumina precipitation on the entry and on the lateral surface of the nozzles can partially or even totally block the flow of the liquid steel. This phenomenon is commonly know as *clogging* and is highly detrimental to casting reliability and quality of the cast products. The capability of a cast of avoiding clogging is named *castability.*

The clogging phenomenon is still not deeply understood [1], due to the very high number of chemical and process factors affecting the occurrence of the precipitation of the materials on the nozzle internal surface as well as to the impossibility of installing complex systems of probes and sensors in order to closely observe the phenomenon itself. The most common way to contrast clogging is to add Calcium [2], as it reacts with solid Al_2O_3 inclusions by generating Calcium aluminates of lower melting point that does not generate clogging and improve the castability. Anyway, some Calcium may also react with dissolved Sulphur [3], by generating Calcium or Ca-Mn Sulphides inclusions, and nozzle blockage can also be produced when solid Calcium Sulphide inclusions are

ESANN'2004 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 28-30 April 2004, d-side publi., ISBN 2-930307-04-8, pp. 263-268

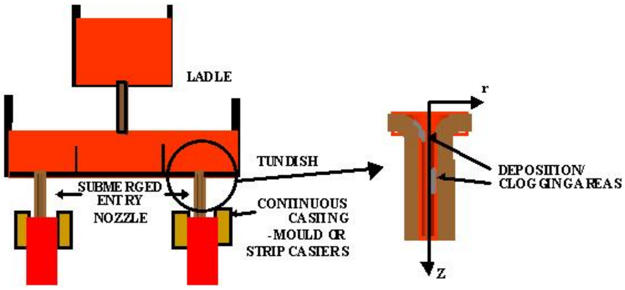

Figure 1: Scheme of the nozzle position and of the location of the precipitated material that can partially clog the nozzle.

present in the steel, whose formation is influenced by the steel composition and by the temperature, thus Calcium addition must be performed in the right intermediate amount, in order to form Calcium aluminates without encouraging the formation of Calcium sulphides

Calcium is added as soon as the first clogging signal occurs, but, as the optimal Calcium amount to be added is not know, successive additions are made by considering the current appearance of the cast and its tendency to clogging. It would obviously be very useful to be able to predict the clogging *before* its actual occurrence. The paper presents some attempts that have been made to accomplish this task by exploiting different neural–network based techniques that should compensate the lack of physical knowledge on the phenomenon by extracting information from the available experimental data.

The paper is organized as follows: Sec. 2 describes some features of the industrial database that has been used in the work, Sec. 3 illustrates the traditional approaches to the classification task that have been attempted before developing the novel method described in detail in Sec. 4. The obtained results are discussed in Sec. 5, while some comments and perspectives for the future work are presented in 6.

2 The available database

The experimental data that have been used for the experiments has been provided by an important Italian steelmaking industry and is formed by 667 records describing the chemical and physical characteristics of as many casts by means of about 100 quantities related to the chemical analysis of the steel performed in different phases of the manufacture as well as to many process variables. During the discussion with the industrial technical personnel it has been pointed out that only a subset of such quantities is supposed to affect the castability of each cast and it is considered during the normal steelmaking practice. Although neural networks–based techniques can discover unknown

ESANN'2004 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 28-30 April 2004, d-side publi., ISBN 2-930307-04-8, pp. 263-268

correlations between some chemical or process variables and the castability, the limited number of the available data forced us to make a reasonable choice among all the available measures. The quantities that have been considered are: Tundish temperature, stiring time, Sulphur content (in wt %, i.e. percentage of the weight of the whole cast) , Aluminum content (in wt %), Soluble Aluminum content (in wt %), Calcium content (in wt %). The aim of the work is to be able to distinguish which data give rise to good casting behaviour (castability index value 1) and which data correspond to bad casting behaviour (castability index value 2).

Since the order of magnitude of the Tundish temperature values are four times greater than those of the other quantities, a normalization stage is required before any further processing stage. Moreover the database has been divided in a training set, that contains the 60% of the available data, and a validation set including the remaining 40% of the data.

The reduced number of the available data obviously limits the number of neurons of the neural–network based classifiers to develop.

3 Traditional neural approaches to the problem

The first approach that has been attempted in order to cluster the casts is based on Learning Vector Quantization (LVQ) method [4] that is an algorithm for learning classifiers from labelled data samples.

The second approach that has been attempted is based on a two–layers Perceptrons [5] trained by means of the Levemberg–Marquardt method [6].

The design of a classifier based on standard Radial Basis Function Networks (RBFN) [7] was also attepted.

Finally, a Self Organizing Map (SOM) was designed [8], by following the basic idea to cluster training samples in an unsupervised manner by using the SOM and then to use the information provided by the castability index of samples to label each cluster of the SOM. More in detail, after the classical unsupervised training of the SOM, each training pattern is classified and each winning cluster stores the information concerning the belonging class of the sample. At the end of this process, each neuron of the SOM contains information about the castability index of all samples for which it resulted the winning neuron. During the prediction, when a new pattern is presented, it is firstly classified by the SOM, then its predicted castability is calculated on the basis of the one corresponding to the winning neuron.

4 The Model–based approach

A novel approach has been applied to the present clustering problem. The main idea of the proposed method is to create a set of reference vectors (that will be named *models* in the following) for each class, in which the training and target sets are divided. Once this set is created and tuned, a new pattern can be

ESANN'2004 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 28-30 April 2004, d-side publi., ISBN 2-930307-04-8, pp. 263-268

classified on the basis of its resemblance to the models. At the beginning, the models are initialized just randomly splitting the training set in as many groups as the number of classes used for the classification. Each model will contain all those samples belonging to a specific class. During the subsequent *training* phase, the patterns are presented one by one to the system and are classified on the basis of the following distance function that evaluates the similarity between the sample and each model:

$$f(m, s) = \alpha \; mean(\|p - s\|_{p \in m}) + \beta \; min(\|p - s\|_{p \in m}) \qquad (1)$$

where m is a model and s is the sample pattern. This distance function considers both the similarity of the sample with all elements in the model and the most similar element within the model. The classification of the sample is made on the basis of the less *distant* model.

Afterwards, depending on the correctness of the classification of pattern u, the models are updated as follows:

1. prediction=i & target=i: each element $v \in model_i$ is updated: $v = v + \rho(u - v)$
2. prediction$\neq i$ & target$\neq i$: each element $v \in model_j$ $i \neq j$ is updated: $v = v - \rho(u - v)$
3. prediction$\neq i$ & target=i: u is added to $model_i$
4. prediction=i & target$\neq i$: u is added to all $model_j$ $i \neq j$

In the first two cases the classification is correct and models are updated in order to reinforce this behaviour; in the remaining cases, the classification is wrong: in case 3 the sample is added to the right model for approaching the model to the sample by decreasing the mean distance between them and by annihilating the distance depending on the minimum similarity between elements in the model and the sample (see equation 1); in case 4 an analogous and opposite thing is done. At the end of the training process new patterns are classified simply by using the same method exploiting equation 1.

5 Results

The results obtained by the all the adopted approach to the problem of the classification of steel castability are summarized in table 1, as well as the time required for training each system. Generally, due to the nature of the problem, it would be strongly preferable to correctly classify as much *class–2* pattern as possible, because it corresponds to the observation of a bad situation and a false alarm is surely preferable to the missed detection of a bad cast so errors in prediction will be weighted according to this criteria.

The results obtained by LVQ method are unsatisfactory; the percentage of success in classification of class–2 casts was only 32%. As far as the MLP

ESANN'2004 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 28-30 April 2004, d-side publi., ISBN 2-930307-04-8, pp. 263-268

approach is concerned, the best performances are obtained with 5 and 8 hidden neurons, but in both cases too much errors (respectively 82.2% and 73.3%) occur when classifying patterns corresponding to the castability class 2. Also the results obtained with the RBF network–based classifier are unsatisfactory, even in the best case, i.e. with 6 hidden neurons.

The SOM-based clustering methods lead to better results. In particular best results are achieved by the 8x8 SOM trained through 100 epochs. This method correctly identifies more than 75% of bad casts and about 70% of good ones.

The performance of the Model–based approach are quite satisfactory, in facts it can correctly identify 54% of bad casts and 97% of good ones; moreover it is the best globally performing method.

Actually the fact that most of the errors were made when classifying patterns belonging to the bad castability class is highly undesirable because it corresponds to an unsuccessful survey of a critical situation and, consequently, to the neglection of suitable countermeasures that could help in avoid nozzle clogging; the opposite behaviour would clearly be preferable, as the Calcium addition does not heavily affect the quality of the produced steel even if it is not really needed. The error in classifying the patterns belonging to the bad castability class could be mainly due to two interacting causes:

1. the number of records referring to bad castability is smaller with respect to the number of cast that show a good casting behaviour;

2. the patterns corresponding to castability index 1 and 2 are not separable in a simple way: a model capable to describe such clustering would need many parameters (and neurons) but, since the available database is not very vast, only few neurons can be used.

The little number of available samples also affects the achievable complexity of the classifier to be designed as well as the generalization capabilities of the examined methods: a higher number of samples should improve generalization capabilities and performances.

6 Conclusions and future work

In this paper we introduced the industrial problem of clustering steel casts in order to predict their castability by exploiting some important chemical and physical properties of the casts. The problem has been faced by using some different well known techniques such as RBFN, SOM, LVQ, MLP as well as a novel method called *Model–based* classifier. All these methods generally achieved satisfactory results by correctly classifying great part of validation samples presented, but most errors are done when coping with patterns corresponding to casts that showed a bad casting behaviour; for this restricted set of samples performance is not satisfactory but hopefully it is due to the reduced number of samples available for the training of presented methods.

Algorithm	Training error (%)			Validation error (%)			Training time (sec.)
	Class 1 85.1%	Class 2 14.9%	Total	Class 1 84.9%	Class 2 85.1%	Total	
LVQ	1.3	56.1	9.5	5.6	68.9	15.1	19.2
MLP [1]	0.26	22.7	1.5	5.2	82.2	16.8	3.7
MLP [2]	0	9.1	0.5	5.2	73.3	15.5	6.9
RBFN [3]	0	0	0	12.3	51.1	20.2	2.1
SOM [4]	18.5	0	17.5	31.8	24.4	34.1	81
SOM [5]	19.0	0	18.0	27.0	28.9	30.3	43
Model–based	0	0	0	3.2	46.7	10.9	24

Table 1: Comparison of the performance achieved by used algorithms. ([1]) 5 hidden neurons ([2]) 8 hidden neurons ([3]) 6 hidden neurons ([4]) 8x8 SOM, 100 epochs ([5]) 8x8 SOM, 50 epochs

In the future much more data will be available and will be used for the training and testing of these methods in order to achieve better results. In particular we will focus our attention on SOM, MLP, RBFN and the Model–based algorithm which, on the basis of obtained results, seem the most promising ones. Moreover it could be interesting to test other methods like ARTMAP and to perfect the presented Model–based method, in order to improve its performance ad reduce the time required for its training.

References

[1] M.J. Heesom: "Physical and chemical aspects of nozzle blockage during continuous casting," *Proc. 1st Int. Calcium Treatment Sym.,* London, 1988.

[2] G.M. Faulring, J.W. Farrel, D.C. Hilty: "Steel flow through nozzles: influence of Calcium," *Ironmaking and Steelmaking*, February 1980.

[3] R.D. Pehlke, T. Fuwa: "Control of Sulphur in liquid iron and steel," *International Metal Review,* Vol. 30, No. 3, 1985.

[4] T. Kohonen: "Improved versions of learning vector quantization", *Proc. of the International Joint Conference on Neural Networks*, pages I 545–550, June 1990.

[5] D.E.Rumelhart, G.E.Hinton, R.J.Williams: " Learning internal representations by error propagation", Rumelhart and McClelland, eds. *Parallel data processing*, Vol. 1, Ch. 8, MIT press, 1986.

[6] D.Marquardt: "An algorithm for least–square estimation of nonlinear parameters", *SIAM Journal of applied mathematics*, No. 11, pp. 164-168.

[7] J. Park and I. W. Sandberg: "Universal approximation using radial-basis-function networks',' *Neural Computation*, Vol. 3, No. 2, pp. 246–257, 1991.

[8] T. Kohonen: "The self-organizing map", *Proceedings of the IEEE*, Vol. 78, No. 9, pp. 1464-1480, 1990.