A Neural Network Approach for Credit Risk Evaluation

Eliana Angelini [a] Giacomo di Tollo [b] Andrea Roli [c]

[a] *DASTA*
Università degli Studi "G.D'Annunzio" Chieti-Pescara
Viale Pindaro 42, I-65127 Pescara (Italia)

[b] *Dipartimento di Scienze*
Università degli Studi "G.D'Annunzio" Chieti-Pescara,
Viale Pindaro 42, I-65127 Pescara (Italia)

[c] *DEIS, Campus of Cesena, Università di Bologna*
Via Venezia 52, I-47023 Cesena (Italia)

Abstract

The Basel Committee on Banking Supervision proposes a capital adequacy framework that allows banks to calculate capital requirement for their banking books using internal assessments of key risk drivers. Hence the need for systems to assess credit risk. Among the new methods, artificial neural networks have shown promising results. In this work, we describe the case of a successful application of neural networks to credit risk assessment. We developed two neural network systems, one with a standard feedforward network, while the other with a special purpose architecture. The application is tested on real-world data, related to Italian small businesses. We show that neural networks can be very successful in learning and estimating the *in bonis/default* tendency of a borrower, provided that careful data analysis, data pre-processing and training are performed.

Key words: Credit risk, Basel II, Neural networks

1 Introduction

The Basel Committee on Banking Supervision, with its revised capital adequacy framework "International Convergence of Capital Measurement and Capital Standards" (Basel Committee on Banking Supervision, 2005) - commonly known as Basel II - proposes a more flexible capital adequacy framework to encourage banks to make ongoing improvements in their risk assessment capabilities . The text reflects the results of extensive consultations with supervisors and bankers worldwide [1] . Regulators allow banks the discretion to calculate capital requirement for their banking books using "internal assessments" of key risk drivers, rather than the alternative regulatory standardized model: the risk weights and thus capital charge are determined through the combination of quantitative inputs provided by bank and formulas specified by the Committee. For the first time, banks will be permitted to rely on their own assessments of a borrower's credit risk.

Credit risk has long been an important and widely studied topic in bank lending decisions and profitability. For all banks, credit remains the single largest risk, difficult to offset, despite advances in credit measurement techniques and the diversification of portfolio. Continuing increases in the scale and complexity of financial institutions and in pace of their transactions demand that they employ sophisticated risk management techniques and monitor rapidly changing credit risk exposures. At the same time, fortunately, advances in information technology have lowered the cost of acquiring, managing and analysing data, in an effort to build more robust and sound financial systems.

In recent years, a number of the largest banks have developed sophisticated systems in an attempt to asses credit risk arising from important aspects of their business lines. What have been the benefits of the new model-based approach to risk measurement and management? The most important is that better risk measurement and management contribute to a more efficient capital allocation [2] . When risk is better evaluated, it can be more accurately priced and it can be more easily spread among a larger number of market participants. The improvement in credit risk modelling has led to the development of new markets for credit risk transfer, such as credit derivatives and collateralised debt obligations (CDOs). These new markets have expanded the ways that market participants can share credit risk and have led to more efficient pricing of that risk (R.W.Ferguson, 2001).

The aim of this paper is to investigate the possibility of employing neural networks to tackle the problem of estimating the probability of default, that

[1] The Committee subsequently released the first round of proposals for revising the capital adequacy framework in June 1999; than in January 2001 and April 2003 and furthermore conducted three quantitative impact studies related to its proposals. As a results of these labours, many valuable improvements have been made to the original proposals.

[2] "...More accurate risk measurement and better management do not mean the absence of loss. Those outcomes in the tails of distributions that with some small probability can occur, on occasion do occur; but improved risk management has meant that lender and investors can more thoroughly understand their risk exposure and intentionally match it to their risk appetite, and they can more easily hedge unwanted risk..." (R.W.Ferguson, 2001)

measures the likelihood that the borrower will default over a given time horizon. As a first step into this direction, we apply the neural network approach in the small-business lending analysis to assess credit risk of Italian companies. The output of the network can be used as a rating value for classifying the company and it may also be incorporated into a probability of default prediction model. The focus of this article is on the empirical approach: We present and discuss two neural network-based models and report the results achieved by performing an extensive experimental analysis, training and validating the networks on real-world data. We emphasize that, since the methodology for developing neural network systems is inherently empirical, the aim of this work is to show that a careful and systematic experimental approach can lead to the design of an effective system.

The paper is organized as follows. The paper begins (section 2) by stating the overall objectives of the Basel Committee in addressing the topic of sound practices to the credit risk management process. Moreover, the paper presents important elements of regulatory function. Section 3 presents an analysis of the conceptual methodologies to credit risk modelling and focuses on the various techniques used for parameter estimation. We have tried to simplify the technical details and analytics surrounding these models. Finally, we emphasize on neural network models, extremely useful for analysing small-business lending problems. In section 4 we first briefly overview the main principles and characteristics of neural networks; then we describe the models we developed and tested. In section 5 the data set used is illustrated: We select 30 companies as samples and assign each of them to one of the two groups: a “good” one, which means that the economic and financial situation is good, and a “bad” one, which means that the company is close to default. Section 6 discusses the experimental settings and the results we obtained, leading to considerable accuracy in prediction. The paper concludes with a discussion of advantages and limitations of the solution achieved and the future work for improvement.

2 Key Elements of the New Basel Capital Accord

2.1 Why new capital requirements?

The existing rules, based on relatively simple 1988 Basel Accord, represented an important step in answering the age-old question of how much capital is enough for banks to hedge economic downturns. Specifically, under the current regulatory structure, virtually all private- sector loans are subject to the same 8 percent capital ratio with no account of the size of the loan, its maturity and, most importantly, the credit quality of the borrower. Thus, loans to a firm near bankruptcy are treated (in capital requirement terms) in the same fashion as loans to a AAA borrowers. Moreover, the current capital requirement is additive across all loans; there is no allowance for lower capital requirements because of a grater degree of diversification in the loan portfolio (Saunders, 1999).

By the late 1990s, it became clear that the original Accord was becoming outdated. Its nature has had a tendency to discourage certain types of bank lending. It has also tended to encourage transactions whose sole benefits is regulatory capital relief (W.J.McDonough, 2003). Further, the international competition, the globalisation and the improvements in the risk management tools, changed the way that banks monitor and measure risk in manner that the 1988 Accord could not anticipate. To response to these challenges, the Committee began a few years ago to develop a more flexible capital adequacy framework. The New Accord consists of three pillars: minimum capital requirements, supervisory review of capital adequacy and public disclosure. The Committee believes that all banks should be subject to a capital adequacy framework comprising minimum capital requirements, supervisory review, and market discipline.

The objective is reached by giving banks a range of increasingly sophisticated options for calculating capital charges. Banks will be expected to employ the capital adequacy method most appropriate to the complexity of their transactions and risk profiles. For credit risk, the range of options begins with the standardized approach and extends to the internal rating-based (IRB) approach. The standardized approach is similar to the current Accord: banks will be expected to allocate capital to their assets based on the risk weights assigned to various exposures. It improved on the original Accord by weighting those exposures based on each borrower's external credit risk rating. Clearly, the IRB approach is a major innovation of the New Accord: bank internal assessments of key risk drivers are primary inputs to the capital requirements. For the first time, banks will be permitted to rely on their own assessments of a borrower's credit risk. The close relationship between the inputs to the regulatory capital calculations and banks' internal risk assessments will facilitate a more risk sensitive approach to minimum capital. Changes in a client's credit quality will be directly reflected in the amount of capital held by banks.

How will the New Basel Accord promote better corporate governance and improve risk management techniques?

First, the Basel Committee has expressly designed the New Accord to provide tangible economic incentives for banks to adopt increasingly sophisticated risk management practices. Banks with better measure of their economic risks will be able to allocate capital more efficiently and more closely in line with their actual sensitivity to the underlying risks. Second, to achieve those capital benefits, the more advanced approaches to credit and operational risk require banks to meet strong process control requirements. Again, the increasing focus on a bank's control environment gives greater weight to the management disciplines of measuring, monitoring and controlling risk (W.J.McDonough, 2003).

2.2 "Core components" in the IRB Approach

In the IRB Approach to credit risk there are two variants: a foundation version and an advanced version. In the first version, banks must provide internal estimates of probability of default (PD)- which measures the likelihood that the borrower will default over a given time horizon. In addition, in the advanced approach, banks, subject to certain minimum conditions and disclosure requirements, can determine other elements needed to calculate their own capital requirements. They are: (a) Loss given default (LGD), which measures the proportion of the exposure that will be lost if a default occurs; (b) Exposure at default (EAD), which for loan commitments measures the amount of the facility that is likely to be drawn if a default occurs; (c) Maturity (M), which measures the remaining economic maturity of the exposure (Basel Committee on Banking Supervision, 2005).In others words, the two approaches differ primarily in terms of the inputs that are provided by the bank based on its own estimates and those that have been specified by the Committee.The risk weights and thus capital charges are determined through the combination of quantitative inputs provided by banks and formulas specified by the supervisor. The risk weight functions have been developed for separate asset classes. For corporate, sovereign and banks formulas are:

$$\textit{Correlation(R)} = 0.12 \times \frac{1-e^{-50\times}\mathrm{PD}}{1-e^{-50}} + 0.24 \times \left[1 - \frac{1-e^{-50\times}\mathrm{PD}}{1-e^{-50}}\right];$$

$$\textit{Maturity adjustment(b)} = (0.11852 - 0.05478\ timesln(\mathrm{PD}))^2;$$

$$\textit{Capital requirement(K)} =$$

$$\left[\mathrm{LGD} \times N\left[(1-R)^{-0.5} \times \mathrm{G(PD)} + \left(\frac{R}{1-R}\right)^{0.5} \times G(0.999)\right] - \mathrm{PD} \times \mathrm{LGD}\right] \times$$
$$\times (1 - 1.5 \times b)^{-1} \times (1 + (M - 2.5) \times b);$$

$$\textit{Risk-weighted assets (RWA)} = K \times 12.5 \times \mathrm{EAD}.$$

Where: $N(x)$ denotes the cumulative distribution function for a standard normal random variable (i.e. the probability that a normal random variable with mean zero and variance of one is less than or equal to x). $G(z)$ denotes the inverse cumulative distribution function for a standard normal random variable (i.e. the value of x such that $N(x) = z$). PD and LGD are measured as decimals, and EAD is measured as currency, except where explicitly noted otherwise.This risk weight function, based on modern risk management techniques, translate a bank's inputs into a specific capital requirement. Under the IRB approach for corporate credits, banks will be permitted to separately distinguish exposures to small and medium sized entities (SME). They are defined as corporate exposures where the reported sales for the consolidates group of which the firm is part is less than €50 million from those to large firms. A firm size adjustment $[0.04 \times 1 - (S - 5)/45]$ is made to the corporate risk weight formula for exposures to SME borrowers (Basel Committee on Banking Supervision, 2005).

$$\textit{Correlation(R)} = 0.12 \times \frac{1-e^{-50\times \mathrm{PD}}}{1-e^{-50}} + 0.24 \times \left[1 - \frac{1-e^{-50\times \mathrm{PD}}}{1-e^{-50}}\right] - 0.04\times$$

$$\times (1 - (S-5)/45).$$

S is expressed as total annual sales in millions of euros with value of S falling in the range of equal to or less than €50 million or greater than or equal to €5 million. Another major element of the IRB approach is the treatment of credit risk mitigants, namely, collateral, guarantees and credit derivatives; in particularly the LGD parameter provides a great deal of flexibility to assess the potential value of credit risk mitigation techniques.In these formulas we can see that the most important risk components is the probability of default. PD estimates must be a long-run average of one year realised default rates for borrowers in the grade. For corporate and bank exposures, the PD is the greater of the one year PD associated with the internal borrower grade to which that exposure is assigned, or 0.03%. Banks may use one or more of the three specific methods (internal default experience, mapping to external data and statistical default models) as well as other information and techniques as appropriate to estimate the average PD for each rating grade.Improvements in the rigour and consistency of credit risk measurement, the flexibility of models in responding to changes in the economic environment and innovations in financial products may produce estimates of credit risk that better reflect the credit risk of exposure. However, before a modelling approach could be used in the formal process of setting regulatory capital requirements for credit risk, regulators would have to be confident not only that models are being used to actively manage risk, but also that they are conceptually sound and empirically validated. Additionally, problems concerning data limitations and model validation must be cleared before models may be implemented in the process of setting regulatory capital requirement. At present, there is no commonly accepted framework for periodically verifying the accuracy of credit risk models; it is important to note that the internal environment in which a model operate – including the amount of management oversight, the quality of internal controls, the rigour of stress testing, the reporting process and others traditional features of credit culture – will also continue to play a key part in the evaluation of bank's risk management framework (Basel Committee on Banking Supervision, 1999).

3 Conceptual methodologies to credit risk modelling

Over the last time, enormous strides have been made in the art and science of credit risk measurement. Banks have devoted increased attention to measuring credit risk and have made important gains, both by employing innovative and sophisticated risk modelling techniques and also by strengthening their more traditional practices (L.H.Meyer, 2000). Measuring credit risk accurately allows banks to engineer future lending transactions, so as to achieve targeted return/risk characteristics. However, credit risk models are not a simple extension of their market risk counterparts for two key reasons (Basel Committee on Banking Supervision, 1999).

- The specification of the process of default and rating migration is severely constrained by a lack of data on the historical performance of loans and other modelled variables; most credit operations are not market to market and the predictive nature of a credit risk model does not derive from a statistical projection of future prices based on a comprehensive record of historical prices. The difficulties in specification are exacerbated by the longer term time horizons used in measuring credit risk, which suggest that many years of data, spanning multiple credit cycles, may be needed to estimate the process of default. Even if individual default probabilities could be modelled accurately, the process of combining these for a portfolio might still be hampered by the scarcity of data with which to estimate reliably the correlations between numerous variables. Hence, in specifying models parameters, credit risk models require the use of simplifying assumptions and proxy data.
- The validation of credit risk models is fundamentally more difficult than the back testing of market risk models. Where market risk models typically employ a horizon of a few days, credit risk models generally rely on a time frame of one year or more; the longer holding period, coupled with the higher confidence intervals used in credit risk models, presents problems to model-builders in assessing the accuracy of their models; the effect of modelling assumptions on estimates of the extreme tails of the distributions is not well understood.

It is difficult to draw the line between traditional and new approaches, especially because many of the better ideas of traditional models are used in the new models. We consider three broad classes of traditional models: 1) expert systems, 2) rating systems, and 3) credit scoring systems [3].

In an expert system, the credit decision is left to the local or branch lending officer. The persons expertise, subjective judgment and weighting of certain key factors are the most important determinants in the decision to grant credit. Bankers have relied on the so called '5 Cs' of expert systems to assess credit quality: the expert analyzes these five key factors and performes a credit decision. These factors are: character (reputation of the firm), capital (leverage), capacity (volatility of the borrower's earnings), collateral and cycle (macroeconomic) conditions. These systems face two main problems: *i* human experts may be inconsistent and subjective in their assessments; *ii* traditional expert systems specify no weighting scheme that would consistently order the '5 Cs' in terms of their relative importance in forecasting the default: what are the important common factors to analyze across different types of borrower?

External credit ratings provided by firms specializing in credit analysis were first offered in the U.S. by Moody's in 1909. These companies offer bond investors access to low cost information about the creditworthiness of bond issuers. The usefulness of this information is not limited to bond investors. However, the rating system has been rather crude, with most loans rated as Pass-performing and only a minority of loans differentiated according to the four non-performing classifications (listed in order of declining credit quality): other assets especially mentioned, substandard, doubtful, and loss (Allen et al., 2003). The most Commonly used traditional credit risk measurement model is the multiple discriminant credit scoring analysis pioneered by Altman (Altman, 1968). The model identifies financial variables that have statistical explanatory power in differentiating "bad firms" from "good firms". Once the model's parameters are obtained, loan applicants are assigned a Z-score. In some circumstances, the score can be interpreted as a probability of default; in others, the score can be used as a classification system: it placed a potential borrower into either a good or bad group, based on a score and a cut-off point.

The traditional scoring models can be subdivided in three main categories:

- models of linear discriminant analysis, in particular, Altman Z-score model.The multivariate analysis, through a discriminating function, permit to synthesize

[3] For a more complete discussion of these models, see (Allen et al., 2003) and (Caouette et al., 1998).

the value of more variables in a single Z value that, compared with a Z_c value (cut-off point) concurs to classify the loan applications into the groups of acceptance or rejection for the loan. The linear discriminant function is following:

$$Z = \lambda_1 x_1 + \lambda_2 x_2 + \ldots + \lambda_n x_n \tag{1}$$

where x_i represents indicators used as independent variables and λ_i indicates discrimination coefficients.

- models of linear regression: they identify, in a selected sample, some random variables ($X_{i,j}$); these variables reflect important information and they are used as independent variables in a linear regression in which the dependent variable is represented by the variable Z (that can assume 0 or 1 value alternatively). In this way, it identifies variables statistically meaningful in insolvency evaluating and it also estimate regression coefficients. In analytical terms:

$$Z_i = \sum_{j=1}^{n} \beta_j X_{i,j} + \epsilon_i \tag{2}$$

where β_j represents the importance of X_j variable in evaluating the past insolvency. This approach suffers from an important problem whenever the probability of default of one new borrower assumes external values to the interval (0 1); such problem is faced from the models logit and probit;

- logit and probit model: the problem of linear model regarding the output not limited in interval (0, 1) is solved by the model of logistic regression (logit); it uses an exponential transformation and results of the regression analysis are included within this interval. We have the following equation:

$$S(z_i) = \frac{1}{1 + e^{-Z}} \tag{3}$$

The expression provides the conditional probability of finding the borrower i in the group of insolvent customers. The probit model only differs from the logit model as far as concerns the relative hypothesis to the distribution: it assumes that the distribution is the standardized normal and therefore $F(Z)$ represents the accumulated function of the normal distribution. In the logit model $F(Z)$ indicates the accumulated function of the logistic distribution, characterized from thicker tails. In the application, it does not determine important differences between the two models if not there are numerous extreme cases in the reference sample. Credit scoring models are relatively inexpensive to use and do not suffer from the subjectivity and inconsistency of expert system. Anyway, these methodologies suffers from important statistical restrictions. First, the distributions of discriminating variables must be jointly normally distributed. Second, the disturbance term of the discriminant function is assumed to be normally distributed with homoscedastic variance. These assumptions may cause problems for interpreting the results of empirical estimation and generating reliable forecasts. Moreover, the explanatory variables are mainly limited to balance sheet data.These data are updated infrequently and are determinated by accounting methods that rely on book, rather than market valuation. The recent application of non-linear methods such as neural networks to credit risk analysis shows improvements on the traditional credit scoring models.

3.2 New Approaches to Credit Risk Measurement

The new models – some publicly available and some partially proprietary – try to offer "internal model" approaches to measure the credit risk of a loan or a portfolio of loans. In this section, we do not propose to make a taxonomy of these approaches, but aim to discuss key elements of the different methodologies.

First, within the current generation of credit risk models, banks employ either of two conceptual definitions of credit loss, the *default mode* (DM) paradigm or the *mark to market* (MTM) paradigm. In the first paradigm a credit loss arises only if a borrower defaults within the planning horizon [4]. In the absence of a default event, no credit loss would be incurred. In the case that a client defaults, the credit loss would reflect the difference between the bank's credit exposure and the present value of future net recoveries. In contrast to the DM paradigm, in the MTM models a credit loss can arise in response to the deterioration in an asset's credit quality. Given the rating transition matrix associated with each client, Monte Carlo methods are generally used to simulate migration paths for each credit position in the portfolio [5].Second, there are different methodologies for the unconditional and conditional models. Unconditional approaches typically reflect customer or facility-specific information [6]. Such models are currently not designed to capture business cycle

[4] The choice of a modelling horizon of one year reflects the typical interval over which: (a) new capital could be raised; (b) new customer information could be revealed; (c) loss mitigation actions could be undertaken; (d) internal budgeting, capital planning and accounting statements are prepared; (e) credits are normally reviewed for renewal.

[5] See J.P.Morgan's Credit MetricsTM framework (1997) for MTM approach; Credit Risk PlusTM of Credit Suisse Financial Product (1997) for DM approach.

[6] Examples are CreditMetricsTM and Credit Risk +TM; these modelling frameworks derive correlation effects on relationship between historical defaults and borrower-specific information such as internal risk ratings. The data is estimated over (ideally) many credit cycles.

effects, such as the tendency for internal ratings to improve (deteriorate) more during cyclical upturns (downturns). Instead, conditional models incorporate information on the state of the economy, such as levels and trends in indicators of economic and financial health, in domestic and international employment, in stock prices and interest rates, ecc. In these models, rating transition matrices are increased likelihood of an upgrade during an upswing in a credit cycle and vice versa[7].

Finally, there are different techniques for measuring the interdependence of factors that contribute to credit losses. In measuring credit risk, the calculation of a measure of the dispersion of credit risk requires consideration of the dependencies between the factors determining credit- related losses, such as correlations among defaults or rating migrations, LGDs and exposures, both for the same borrower and among different borrowers (Basel Committee on Banking Supervision, 2000).

In closing, the fundamental elements of credit risk measurement are easy to describe in the abstract but are far more difficult to apply case by case. Each situation is unique, built around the roles and capabilities of individuals, the technology system, activities and objectives of the institutions. Anyway, to remain competitive, institutions must adapt and constantly improve their credit risk measurement techniques.

4 The neural network approach

Neural networks have recently emerged as an effective method for credit scoring (Wu and Wang, 2000; Rong-Zhou et al., 4-5 Novembre 2002; Pang et al., 4-5 Novembre 2002; Piramuthu, 1999; Atiya, 2001). They differ from classical credit scoring systems, such as the *Z–score model* (Altman, 1968), mainly in their *blackbox* nature and because of they assume a non-linear relation among variables. Neural networks are learning systems which can model the relation between a set of inputs and a set of outputs, under the assumption that the relation is nonlinear. They are considered blackboxes, since, in general, it is not possible to extract symbolic information from their internal configuration.

In this section, we introduce neural networks with the aim of defining the background needed to discuss the model we propose and the experimental results.

[7] One example is McKinsey and Company's Credit Portfolio ViewTM.

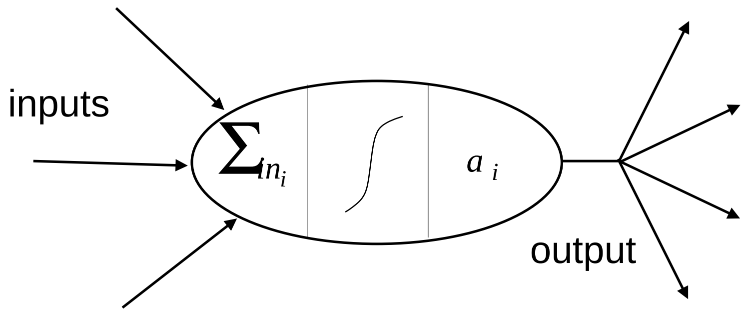

Fig. 1. Artificial neuron basic structure.

4.1 Neural networks: an introduction

Neural networks are machine learning systems based on a simplified model of the biological neuron (Hykin, 1999). In the same way as the biological neural network changes itself in order to perform some cognitive task (such as recognizing faces or learning a concept), artificial neural networks modify their internal parameters in order to perform a given computational task.

Typical tasks neural networks perform efficiently and effectively are: Classification (i.e., deciding which category a given example belongs to), recognizing patterns in data, prediction (such as the identification of a disease from some symptoms, or the identification of causes, once effects are given).

The two main issues to be defined in a neural network application are the network typology and structure and the learning algorithm (i.e., the procedure used to adapt the network so as to make it able to solve the computational task at hand).

An artificial neural network[8] is composed of a bunch of neurons, connected in a predefined topology. There are some possible topologies, usually depending on the task the network has to learn. Usually, the network topology is kept constant, but in some applications (for instance in robotics) the topology itself can be considered as a parameter and can dynamically change. The connections (links) among neurons have associated a *weight* which determines type and intensity of the information exchanged. The set of weights represents, in essence, the information the network uses to perform the task, i.e., given a topology, the weights represent the *functions* that defines the network behavior.

The artificial neuron is an extremely simplified model of the biological neuron and it is depicted in Fig.1. Neurons are the elementary computational units of the network. A neuron receives inputs from other neurons and produces an

[8] In the following, we will skip the adjective “artificial”, since we will not deal with biological neural networks.

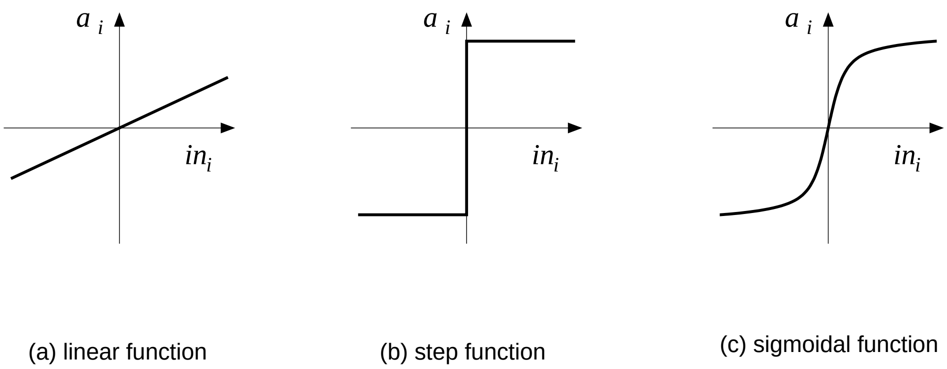

Fig. 2. Examples of activation functions.

output which is transmitted to other destination neurons.

The generation of the output is divided in two steps. In the first step, the weighted sum of inputs is evaluated, i.e., every single input is multiplied by the weight on the corresponding link and all these values are summed up. Then, the *activation* is evaluated by applying a particular *activation function* to the weighted sum of inputs. In formulas, for neuron i receiving inputs from neurons in the set I,

$$
\begin{array}{ll}
y_i = \sum_{j \in I} W_{j,i} a_j & \text{input evaluation} \\
a_i = g(y) & \text{the activation function } g \text{ is applied}
\end{array}
$$

where $W_{j,i}$ is the weight of the link connecting neuron j with neuron i, and a_j is the activation of neuron j.

Some kinds of activation function are used, for instance: linear, step function and sigmoid (see Fig.2).

The most used network topologies are the following:

- Layered
- Completely connected

Networks of the first category have neurons subdivided in layers. If the connections are only in one direction (i.e., each neuron receives inputs from the previous layer and sends output to the following layer), they are called *feed-forward networks*. Otherwise, if also 'loops' are allowed, the network is called *recurrent network*. Completely connected networks, on the other hand, have neurons which are all connected with each other [9]. In Fig.3, examples of the three main kinds of topologies are depicted.

[9] In principle, a completely connected network can be seen as a special case of a recurrent network with one level of n neurons and n^2 connections.

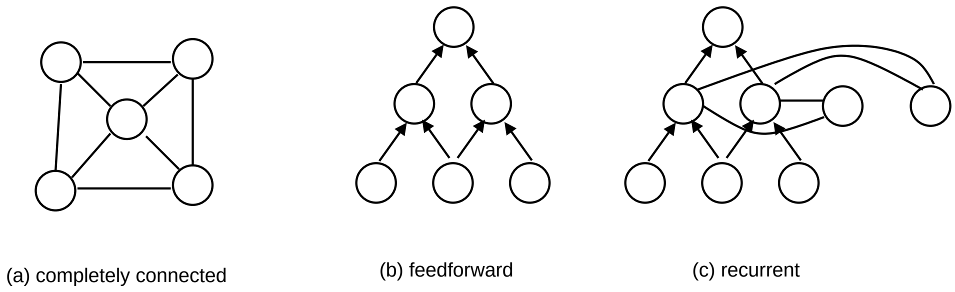

Fig. 3. Examples of network topologies.

Before the neural network can be applied to the problem at hand, a specific tuning of its weights has to be done. This task is accomplished by the learning algorithm which trains the network and iteratively modifies the weights until a specific condition is verified. In most applications, the learning algorithm stops as soon as the discrepancy (*error*) between desired output and the output produced by the network falls below a predefined threshold. There are three typologies of learning mechanisms for neural networks:

- Supervised learning
- Unsupervised learning
- Reinforced learning

Supervised learning is characterized by a *training set* which is a set of correct examples used to train the network. The training set is composed of pairs of inputs and corresponding desired outputs. The error produced by the network then is used to change the weights. This kind of learning is applied in cases in which the network has to learn to generalize the given examples. A typical application is classification: A given input has to be inserted in one of the defined categories.

In *unsupervised learning* algorithms, the network is only provided with a set of inputs and no desired output is given. The algorithm guides the network to self-organize and adapt its weights. This kind of learning is used for tasks such as data mining and clustering, where some regularities in a large amount of data have to be found.

Finally, *reinforced learning* trains the network by introducing prizes and penalties as a function of the network response. Prizes and penalties are then used to modify the weights. Reinforced learning algorithms are applied, for instance, to train adaptive systems which perform a task composed of a sequence of actions. The final outcome is the result of this sequence, therefore the contribution of each action has to be evaluated in the context of the action chain produced. [10]

[10] In these cases it is difficult, if not impossible, to produce an effective training set, hence the need of the use of prizes and penalties, instead of the definition of an

Diverse algorithms to train neural networks have been presented in the literature. There are algorithms specifically designed for a particular kind of neural networks, such as the *backpropagation algorithm* (Werbos, 1988) or general purpose algorithms, such as *genetic algorithms* (Mitchell, 1998) and *simulated annealing* (Kirkpatrick et al., 1983).

At the end of this introductory section, we would like to remark advantages and limits of systems based on neural networks. The main advantages have to be found in their learning capabilities and in the fact the derived model does not make any assumpion on the relations among input variables. Conversely, a theoretical limit of neural networks is that they are *black-box* systems and the extraction of symbolic knowledge is awkward. Moreover, design and optimization neural network methodologies are almost all empirical, thus the experience and sensibility of the designer have a strong contribution in the final success. Nevertheless, with this work we show that some useful general design and parameters optimization guidelines exist.

In the next section we will describe in more detail the neural network models we adopted. We used a classical feedforward network and a variation of feedforward network with *ad hoc* connections. Both networks are trained with the backpropagation algorithm.

4.2 Our models

In our experiments, we used a feedforward neural network –in the classical topology– and with a feedforward neural network with *ad hoc* connections.

The feedforward network architecture is composed of an input layer, two *hidden* layers and an output layer (composed of only one neuron). In Fig.4 an example of a classical two hidden layers feedforward network is represented. In the following, we will indicate with $W_{k,j}$ the weights between input and hidden neurons and with $W_{j,i}$ the weights between hidden and output neurons. Moreover, the activation will be I_k, H_j and O_i for input, hidden and output neurons respectively.

The feedforward network with *ad hoc* connections (thereinafter referred to as ad hoc network) is a four layers feedforward network with the input neurons grouped by three. Each group is connected to one neuron of the following layer. The reasons for choosing this topology are to be found in the actual date we used and will be described in the next section. The ad hoc network topology is depicted in Fig.5.

---

explicit error as in the case of supervised learning.

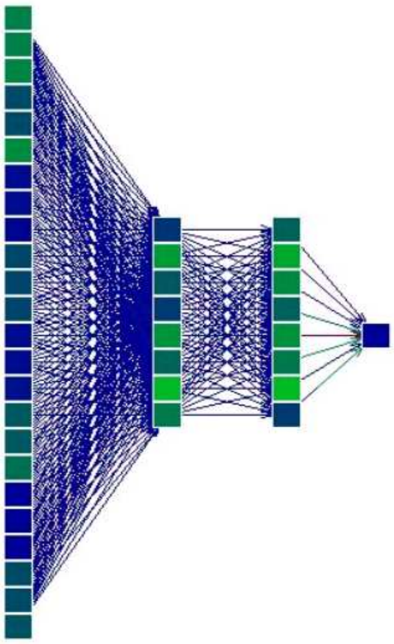

Fig. 4. A four layers feedforward network.

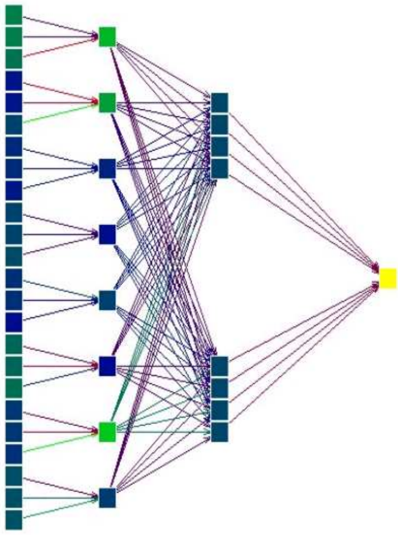

Fig. 5. The ad hoc network used in our experiments.

Both the networks have been trained by means of a supervised algorithm, namely the backpropagation algorithm. This algorithm performs an optimisation of the network weights trying to minimize the error between desired and actual output. For each output neuron i, the error is: $Err_i = T_i - O_i$, where T_i is the desired output. The weight update formula used to change the weights $W_{j,i}$ is the following:

$$W_{j,i} \leftarrow W_{j,i} + \eta H_j Err_i g'(in_i),$$

where η is a coefficient which controls the amount of change (the *learning rate*), g' is the first derivative of the activation function g and in_i is equal to $\sum_j W_{j,i} H_j$.

The formula for updating weights $W_{k,j}$ is similar:

$$W_{k,j} \leftarrow W_{k,j} + \eta I_k \Delta_j,$$

where I_k is the k-th input [11] and $\Delta_j = g'(in_j) \sum_i W_{j,i} Err_i g'(in_i)$.

The algorithm is iterated until a stopping criterion is satisfied [12].

The definition of both the network structure and the algorithm needs the careful choice of parameters, such as the number of neurons at each layer and the value of the learning rate. Moreover, both the data used to train the network (the training set) and the data on which the performance of the network is validated (the test set) need to be carefully chosen. In Sec.6, we will describe our choices along with the benchmark used. In the next section we will describe the data set used in the experiments.

4.3 Neural networks applications to economics and finance

Neural networks have been widely applied to tackle problems arising from the financial and economical areas (see (Vellido et al., 1999) for an overview). We can classify neural networks applications to these areas as works tackling the following problems.

Classification and discrimination In this kind of problems the net must determine which class an input pattern belongs to. Classes are either user-defined (in classification problems) or run-time grasped by the net itself (similarly to the *cluster-analysis* (Wooldridge, 2002)). The main problem to be tackled is the credit risk assessment, but other problems exist: Bank failure prediction (Tam and Kiang, 1992), stocks classification (Kryzanowsky et al., 1993) etc.

Series prediction In this class of problems the net must forecast the next element belonging to a temporal series (Kaastra and Boyd, 1996). Representatives of this problem class are forecasting stocks prices (White, 1988) and indexes (Chen et al., 2003), currencies prices (Shazly and Shazly, 1999), options (Hamid and Iqbal, 2004), as well as macro-economics indexes (Moshiri

[11] The activation of input neurons is simply the input value.
[12] E.g., the overall network error is below a predefined threshold or the maximum amount of allotted time is reached.

et al., 1999; Tkacz, 2001). Other problems have been tackled, such as predicting financial performances of equities w.r.t benchmarks (Lam, 2004), financial crashes forecast (Rotundo, 2004), economic crisis warnings (Kim et al., 2004) and pupil expenditure in public school prediction (Baker and Richards, 1999).

Function approximation and optimization This class consists of problems that can not be said to belong to the previous two classes. Examples are given by portfolio selection and re–balancing (Fernandez and Gomez, 2005; Steiner and Wittkemper, 1997; Zimmermann and Neuneier, 1999) and investment project return prediction (Badiru and Sieger, 1998).

Neural networks application to the first two classes have proved to be very effective and often to be able to outperform traditional methods, especially when the task consists in modelling unknown relationships amongst variables. This is the case of *options volatility forecasting*, in which neural networks have been shown to accurately forecast S & P Index future options, outperforming the classical BAW futures options pricing (Hamid and Iqbal, 2004). Similar results have been obtained in predicting excess returns on large stocks (Desai and Bharati, 1998), showing that neural network forecasting is conditionally efficient w.r.t regression linear models [13]. Some works address the issue of comparing neural networks with standard linear methods. For example, neural networks and linear models have been compared to forecast industrial production series (Heravi et al., 2004), showing that linear models outperform neural networks when the forecast horizon is smaller than one year, whilst neural networks better forecast the trend of the variation. There are works comparing neural networks and linear models for the credit risk assessment too, and the outcome of the comparison is still unclear. Neural Networks outperform Linear Methods (Logistic and Probit as well) in most works (Galindo and Tamayo, 2000; Odon and Sharda, 1990; Salchenberger et al., 1992; Fan and Palaniswami, 2000), but the superiority of these latter methods is claimed in (Altman et al., 1994; Yang et al., 1999), while other works lead to comparable outcomes (Coats and Fant, 1993; Boritz and Kennedy, 1995). We observe that the cited works employ different benchmarks and problem definitions, therefore a direct comparison is not possible. As a concluding comment, it is worth to remark that the choice of a method over another one is usually based on several, not homogeneous criteria, such as the actual goals and data features available. Thus, it is difficult, if not meaningless, try to assess the general supseriority of a technique over another one.

[13] One set of forecast is said to be conditionally efficient w.r.t another if combining the second set with the first produces no gains in the forecast error compared with using the first set alone.

5 Data set

For our experiments, we used data of 76 small businesses from a bank in Italy. The sample is rather diversificated and distributed across some industries. Small businesses characterize the Italian economy and therefore we apply the neural network method to small business lending decisions. Loans to small businesses differ from loans to large businesses (particularly widespread in US markets) in terms of credit analysis. We suggest three major differences. First, since banks (and other lending institutions) face fixed costs in lending, lending to small firms is by definition more expensive. The loan decision making is a labor-intensive and time-consuming process that is very costly for small business lending. Second, the relationship between the owner/manager of a small firm and a bank is often very close. Finally, small firms are more informationally opaque. Generally, it's much more difficult to collect small firm data, and even when data are available they often contain substantial noise. These problems are particularly troublesome for Italian banks with a high proportion of small business loans. Because of these structural features, more banks have turned to neural networks for assistance in making lending decisions better, faster and cheaper. For each business we have annual data across three years (2001-2003) [14]. The sample period covers 3 years; the period may happen to be a down turn (or up turn) in a cycle. Anyway, external factors such as business effects and the macroeconomic environment can be neglected for the low correlation between small and medium exposures (SME) and economic cycle (Basel Committee on Banking Supervision, 2005).

For each business we have 11 fields: 8 of them are financial ratios drawn from the balance sheet of the firms: The ratios (fields) on the sample set are the following:

(1) $\frac{\text{Cash flow}}{\text{Total debt}}$, where cash flow is given by summing earnings,amortizations and depreciations, less bad debt provisions and working capital devaluation;

(2) $\frac{\text{Turnover}}{\text{Inventory}}$, this ratio can be misleading because sales can be of different magnitude depending on the economic sector of firms;

(3) $\frac{\text{Current Liability}}{\text{Turnover}}$, this ratio measures the skill to create cash-flow to extinguish short-term-debts;

(4) $\frac{\text{Equity}}{\text{Total assets}}$, where equity represents the internal way of financing;

(5) $\frac{\text{Financial costs}}{\text{Total debts}}$, indicating the burden of passive interests (originated by

[14] Data are drawn from the annual balance sheet of businesses.

external financing sources) on total debts, measuring the relevance of financial operations;

(6) $\frac{\text{Net working capital}}{\text{Total assets}}$, where net working capital is given by the difference between current assets and current liability;

(7) $\frac{\text{Trade accounts receivables}}{\text{Turnover}}$, indicating if sale operations have the effect to generate a cash flow or entail credit position arising. Values higher then 1 show firms not able to collect their credits;

(8) $\frac{\text{Value added}}{\text{Total assets}}$, where value added represents the new wealth produced by firms as their contribution to the valorization of external bought resources.

The remaining 7 are calculated analysing the credit positions with the supplying bank (*"Andamentale"* ratios):

(9) $\frac{\text{Utilized credit line}}{\text{Accorded credit line}}$;

(10) $\frac{\text{Unsolved effects (quantity)}}{\text{Under usual reserve effects (quantity)}}$ [15];

(11) $\frac{\text{Unsolved effects (value)}}{\text{Under usual reserve effects (value)}}$ [16].

and with the overall Italian Banking System (*"Centrale dei Rischi"*[17] ratios[18]).

(12) $\frac{\text{Transpassing short-term}}{\text{Accorded credit line short-term}}$;

(13) $\frac{\text{Transpassing medium-long term}}{\text{Accorded credit line medium-long term}}$;

(14) $\frac{\text{Utilized credit line short-term}}{\text{Accorded credit line short-term}}$;

(15) $\frac{\text{Utilized credit line medium-long term}}{\text{Accorded credit line medium-long term}}$.

Herebelow we report the main summary statistics for indices at hand:

The sample businesses are categorized in two groups: the "in bonis" group (composed of firms repaying the loan obligation at the end of the analysing period) and the "default" group (composed of firms not repaying the loan obligation at the end of the analysing period). "

[15] Represents the number of unsolved effects divided for their total number.

[16] Represents the value of unsolved effects divided for their total value.

[17] This data are drawn from the "Centrale Dei Rischi", a "Banca D'Italia" kept data-base.

[18] Utilized credit line consists on the actual credit used by customer; accorded credit line consists on the maximum threshold of credit that the bank decides to grant. When the utilized credit line is higher then accorded one, transpassing is given by the difference between these two terms, otherwise it is equal to 0.

Table 1
Main statistics of indices at hand.

	Min	Max	Correct entries	Mean	Median	std.dev.
$\frac{\text{Cash flow}}{\text{Total debt}}$	-1.07	0.52	315	0.04	0.03	0.18
$\frac{\text{Turnover}}{\text{Inventory}}$	0	1877	233	77.95	8.53	173.84
$\frac{\text{Current Liability}}{\text{Turnover}}$	0	1277.5	312	11.35	0.72	102.21
$\frac{\text{Equity}}{\text{Total assets}}$	-10.69	1.6	305	0.03	0.08	1.08
$\frac{\text{Financial costs}}{\text{Total debts}}$	0	0.72	315	0.05	0.04	0.07
$\frac{\text{Net working capital}}{\text{Total assets}}$	-78.75	1.01	305	-0.82	0	7.65
$\frac{\text{Trade accounts receivables}}{\text{Turnover}}$	0	13.84	312	0.51	0.17	1.77
$\frac{\text{Value added}}{\text{Total assets}}$	-0.22	1.13	305	0.24	0.21	0.24
$\frac{\text{Utilized credit line}}{\text{Accorded credit line}}$	0	14.54	219	0.81	0.68	1.47
$\frac{\text{Unsolved effects (quantity)}}{\text{Under usual reserve effects (quantity)}}$	0	4.48	39	0.48	0.09	1.05
$\frac{\text{Unsolved effects (value)}}{\text{Under usual reserve effects (value)}}$	0	5.15	39	0.52	0.07	0.31
$\frac{\text{Transpassing short-term}}{\text{Accorded credit line short-term}}$	-2.74	18.95	156	0.21	0	1.61
$\frac{\text{Transpassing medium-long term}}{\text{Accorded credit line medium-long term}}$	0	3.99	207	0.11	0	0.39
$\frac{\text{Utilized credit line short-term}}{\text{Accorded credit line short-term}}$	0	19.95	157	0.8	0.63	1.8
$\frac{\text{Utilized credit line medium-long term}}{\text{Accorded credit line medium-long term}}$	0	4.99	208	0.85	1	0.57

Before feeding the neural net with our data we decided to operate some pre-processing operations. Here below we describe the issues about data-pre-processing.

Missing and wrong values Some values were missing from the data about firms. This occurrence can be due to two different reasons:

(1) the value is missing because in that year there is no value in the data-base for that firm in that particular field, or because the value is not in the theoretical allowed range [19];
(2) the value is missing because of a computation-error. This happens because the variables we are working on are mathematical ratios, and in our sample some ratios can derive by a division by zero.

The usual way of overcoming this problem is to discard from the data set the all the entries of the corresponding firm (this is the approach pursued on several works about neural nets and credit risk), but operating in that way we would loss a significant amount of information. In order to preserve this information we will perform the replacement of missing values with other values. We decide to handle differently these two situations.

(1) In the first case (missing because of none or wrong value) we decide to

[19] Such a kind of errors may occur because of typos and transcription errors accidentally introduced by bank employees. This situation is quite common and constitutes one of the sources of noise in real-world data.

substitute the empty space with the arithmetical mean of the field, being the mean calculated as the mean of the existing values belonging to that field for all the businesses in the overall period of collecting (column substitution);

(2) In the latter case we decide to replace the missing value with the upper limit of the Normalization interval (see below). This choice can easily understood: as this occurrence is given by a division by zero, we can imagine the result of this ratio being ∞. If we normalize the belonging field, this occurrence will be replaced by the maximum value in the range. We mentioned earlier in our experiment this upper limit being 1.

Erasing useless fields We discussed about replacing missing and wrong values with suitable ones in order to allow the net using those data and preserving useful information. This operation can be harmful if the processed field show too many missing and wrong values: replacing them would convey the net to draw wrong conclusions about the variables-dynamic. Furthermore, as we will see, the overall-sample splitting into *training* and *test* set is made using some percentage ratios. For this purpose we will use the ratios [70% 30%]. In the worst case, if for a specific field there are more than 30% of missing and wrong values (on the overall observations) they can be all included in the test set. This will result in an incredible loss of generalization capability of the net. For this reason we decided not to use the fields containing more than 30% of missing and wrong values and precisely the following fields belonging to "Andamentale" and "Centrale dei Rischi" ratios:

- $\frac{\text{Unsolved effects (quantity)}}{\text{Under usual reserve effects (quantity)}}$;
- $\frac{\text{Unsolved effects (value)}}{\text{Under usual reserve effects (value)}}$;
- $\frac{\text{Transpassing short-term}}{\text{Accorded credit line short-term}}$;
- $\frac{\text{Utilized credit line short-term}}{\text{Accorded credit line short-term}}$.

Data Normalization Data normalization must be performed in order to feed the net with data ranging in the same interval for each input node. We choose to use the interval [0, 1] for each input node. The most common way for normalizing data is the *Min-Max linear transformation to* [0, 1] [20], but we cannot use this formula because there are "outliers" [21] in the data-base we are using. Using the Min-Max formula we would lose a lot of useful information

[20] The general formula of this kind of normalization is $\overline{x}_{ia} = \frac{x_{ia} - min_a}{max_a - min_a}$, where $\overline{x}_{ia}$ is the after-normalizing value and x_{ia} is the actual value.

[21] Outlier means that the value is too different from the other values in the field we are analysing.

and several fields would have almost all data close to one of the limit of normalization. For this reason we decided to use the logarithmic formula to normalize data. This formula is the most flexible because can be defined by the user, provided that the argument of the formula being < 1. For our ratios we used the following formula:

$$\overline{x} = log_m(x + 1)$$

Where m is near to the actual maximum of the field we are analysing. We add 1 to the argument to avoid the value being > 1.

5.0.0.1 Correlation analysis In order to decide which ratios to use as input variables we performed a correlation analysis: with this operation we want to find out the most strongly correlated variables and remove them from our further experiments. The results of this analysis show that there is no strong correlation between couples of variables, so we use the 11 fields examined so far.

5.0.0.2 Training and test set We select 53 firms to consist of training set and 23 firms to consist of test set. We want to reserve, defining both sets, the ratio between in bonis and default firms existing in the overall sample set. For this reason we use 33 in bonis examples and 20 default examples to define the training set. For the same purpose we use 15 in bonis examples and 8 default examples to define the test set.

6 Experimental Results

The two network architectures introduced in Sec.4.2 have been trained with backpropagation using training and test sets as described in the previous section.

The inputs of the networks are the eleven (normalized) attributes. In the classical feedforward network, they are simply given as an ordered array, while in the ad hoc network they are first grouped by three, each group corresponding to the values of an attribute over three years.

The output y of the network, a real value in the range $[0, 1]$ is interpreted as follows:

- If $y < 0.5$ then the input is classified as **in bonis**;
- otherwise, the input is classified as **default**

A large number of tests has been performed trying to optimally tune parameters of the network and the algorithm. Despite many years of research in neural networks, parameter tuning is usually still lacking a theoretical framework, therefore this design phase is still performed empirically. Nevertheless, some general methodological guidelines are possible and useful. For instance, our approach is based on a simple –but effective– systematic optimisation procedure. More formally, we applied a procedure similar to a gradient ascent algorithm in the space of parameters (Blum and Roli, 2003). First, we selected for each parameter p_h (e.g., number of hidden neurons or backpropagation parameters such as learning rate and momentum) a set of values V_h ($h = 1, \ldots, N_{param}$). Thus, the problem is to find an assignment $\{(p_1, v_1), \ldots, (p_{N_{param}}, v_{N_{param}})\}$, $v_h \in V_h$, that leads to a network with optimal performance. We start from a heuristically generated assignment $\{(p_1, v_1^0), \ldots, (p_{N_{param}}, v_{N_{param}}^0)\}$ and we iterate the following procedure: at step h the current assignment is $\{(p_1, v_1^*), \ldots,$ $\ldots, (p_{h-1}, v_{h-1}^*), \ldots, (p_h, v_h^0), \ldots, (p_{N_{param}}, v_{N_{param}}^0)\}$ (only parameters $p_1, \ldots,$ $\ldots, p_{h-1}$ have been assigned) and the network is trained on each instance for every possible value of parameter p_h, while the other parameters are kept constant. Then, the optimal value v_h^* is chosen. At the end, we obtain the assignment $\{(p_1, v_1^*), \ldots, (p_h, v_h^*), \ldots, (p_{N_{param}}, v_{N_{param}}^*)\}$ that is, at least, not worse than any other partial assignment $\{(p_1, v_1^*), (p_{h-1}, v_{h-1}^*), \ldots, (p_h, v_h^0), \ldots,$ $\ldots, (p_{N_{param}}, v_{N_{param}}^0)\}$.

We performed a long series of experiments, of which we report only the ones corresponding to the networks achieving the best results in terms of classification. We obtained a very effective tuning for the classical feedforward network by applying a slight variation of the standard backpropagation algorithm that does not propagate errors if they are below a given threshold. With this technique we were able to produce a network with null error on the training set and an error of 8.6%, corresponding to only wrong in bonis classifications (i.e., an input that should be classified as in bonis is classified as default). We remark the fact that this network was able to correctly classifying all the default cases, that are considerably riskier than in bonis ones. As in machine learning and data mining techniques is often discussed (Han and Kamber, 2000), false negatives are usually much more dramatically important in real world cases (e.g., in diagnosis). In table 2, we report the number of neurons in the hidden layer, the wrong in bonis (*misbo*) and default (*misdef*) classifications and the global error (i.e., the overall error on the considered set), in the training set and the test set respectively. All errors are reported in percentage. For completeness, we report also the values of the algorithm parameters: $\eta = 0.2$, $\beta = 0$, $\delta = 0.1$; initial weights are randomly chosen in the range $[-1, 1]$.

With the ad hoc network we could also achieve very good results, even with a lower error on the test set that was equal to 4.3% in our best configuration. Nevertheless, the errors are all related to false negative cases, i.e., the network classifies as in bonis inputs that should be classified as default. Table 3

Table 2
Best results achieved with a classical feedforward network. we report the number of neurons in the hidden layer, the wrong in bonis (*misbo*) and default (*misdef*) classifications and the global error (i.e., the overall error on the training/test set), in the training set and the test set respectively. All errors are reported in percentage.

# of hidden neurons	tr. set misbo	tr. set misdef	tr. set error	test set misbo	test set misdef	test set error
25	0%	0%	0%	13.3%	0%	8.6%
26	0%	0%	0%	13.3%	0%	8.6%
27	0%	0%	0%	13.3%	0%	8.6%
28	0%	0%	0%	13.3%	0%	8.6%
29	0%	0%	0%	13.3%	0%	8.6%
33	0%	0%	0%	13.3%	0%	8.6%

summarizes the performance of the best found configuration. For this reason, we may consider the two networks as complementary: The one returns *safe* answers, while having a – rather very small – error; the second has a very high performance in terms of overall error on the test set, but it wrongly classifies positive cases. The parameters used for training the ad hoc network are the following: $\eta = 0.8$, $\beta = 0.2$, $\delta = 0$; initial weights are randomly chosen in the range $[-1, 1]$.

Table 3
Best results achieved with the ad hoc feedforward network. We report the wrong in bonis (*misbo*) and default (*misdef*) classifications and the global error, in the training set and the test set respectively. All errors are reported in percentage.

# of hidden neurons	tr. set misbo	tr. set misdef	tr. set error	test set misbo	test set misdef	test set error
$11 + 11$ (2 layers)	0%	5%	1.8%	0%	12.5%	4.3%

In tables 2 and 3 we reported the best results achieved during the parameter optimization phase and for a given training/test set combination of the available data. In order to assess the significance of these results, we analyzed the network behavior with respect to different partitions of the data set into training and test set. Tables 4 and 5 reports the statistics of the overall error, and error in bonis and default over 30 different partitions of the data set, respectively for the feedforward and ad hoc networks. We can observe that the networks are very robust with respect to the choice of the training/test sets and these results confirm the effectiveness of this approach. Indeed, we can note that the average total errors are between 11% and 14% for the feedforward networks and the ad hoc network has a total average error of about 7%.

Table 4
Feedforward networks: statistics over 30 different partitions of the data set into training and test set.

# of hidden neurons	test set misbo		test set misdef		test set error	
	mean	std.dev.	mean	std.dev.	mean	std.dev.
25	10.53%	11.59	14.00%	11.19	11.74%	8.96
26	12.13%	11.87	13.00%	11.83	12.43%	9.50
27	12.80%	12.83	14.00%	11.47	13.22%	9.89
28	10.27%	9.84	14.00%	9.98	11.57%	7.93
29	10.93%	10.41	12.75%	9.61	11.57%	8.58
30	12.13%	11.87	13.25%	11.68	12.52%	8.94
31	13.20%	10.89	14.75%	11.76	13.74%	8.84
32	11.60%	12.03	12.25%	11.14	11.83%	9.58
33	13.60%	12.34	16.00%	11.58	14.43%	9.43

Table 5
Ad hoc networks: statistics over 30 different partitions of the data set into training and test set.

# of hidden neurons	test set misbo		test set misdef		test set error	
	mean	std.dev.	mean	std.dev.	mean	std.dev.
11 + 11 (2 layers)	6.133%	7.59	9.5%	10.58	7.30%	6.59

7 Conclusion and future work

In this paper, we have presented an application of artificial neural network to credit risk assessment. We have discussed two neural architecture for the classification of borrowers into two distinct classes: in bonis and default. The system has been trained and tested on data related to Italian small businesses. One of the system we developed is based on a classical feedforward neural network, while the other one has a special purpose feedforward architecture. Results in both cases show that the approach is very effective and leads to a system able to correctly classify the inputs with a very low error. The overall performance of the networks we developed can be considered as a state-of-the-art result. One of the reasons for this performance is the careful analysis of the data at hand. In fact, real-world data are often noisy and incomplete; therefore, our analysis was aimed at eliminating wrong values and replace empty entries with meaningful values. Moreover, since also data normalization plays an important role in the final performance, we investigated some normalization procedures in order to keep as much information as possible in the inputs used to feed the network.

This empirical work proves, on the one hand, evidence for the actual applicability of neural networks in credit risk applications, especially as black-box non-linear systems to be used in conjunction with classical rating and classification systems. On the other hand, this research also shows that the critical issues for developing such a kind of systems are data analysing and processing.

Future work is focused on both methodological and application issues. As to methodology, we are currently working on the design of procedural techniques for data analysis and processing and for parameter optimisation. On the side of applications, we plan to assess the generalization capabilities of the networks by testing them on wider data bases and to investigate on the applicability of recurrent networks.

References

Allen, L., DeLong, G., Saunders, A., April 2003. Issues in the credit risk modelling of retail markets. In: Conference on Retail Credit Risk Management and Measurement. Philadelphia (USA).

Altman, E., 1968. Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. J. Finance 13.

Altman, E., Marco, G., Varetto, F., 1994. Corporate distress diagnosis: Comparison using linear discriminant analysis and neural networks. Journal of Banking and Finance 18.

Atiya, A., 2001. Bankruptcy prediction for credit risk using neural networks: A survey and new results. IEEE Transactions on Neural Networks 12 (4).

Badiru, A. B., Sieger, D., 1998. Neural networks as a simulation metamodel in economic analysis of risky projects. European Journal of Operational Research 105.

Baker, B., Richards, C., 1999. A comparison of conventional linear regression methods and neural networks for forecasting educational spending. Economics of education review 18.

Basel Committee on Banking Supervision, April 1999. Credit risk modelling: Current practice and applications. Mimeo.

Basel Committee on Banking Supervision, 2000. Overview of conceptual approaches to credit risk modelling, Part II. mimeo.

Basel Committee on Banking Supervision, November 2005. International convergence of capital measurement and capital standards. A revised framework. Bank for International Settlements, Basel.

Blum, C., Roli, A., September 2003. Metaheuristics in combinatorial optimization: Overview and conceptual comparison. ACM Computing Surveys 35 (3), 268–308.

Boritz, J., Kennedy, D., 1995. Effectiveness of neural networks types for prediction of business failure. Expert Syst.Appl. 9.

Caouette, J., Altman, E., Narayanan, P., 1998. Managing Credit Risk: The Next Great Financial Challenge. John Wiley & Sons.

Chen, A., Leung, M., Daouk, H., 2003. Application of neural networks to an emerging financial market: forecasting and trading the taiwan stock index. Computers & Operations Research 30.

Coats, P., Fant, L., 1993. Recognizing financial distress patterns using a neural network tool. Financial Management 22.

Desai, V., Bharati, R., 1998. A comparison of linear regression and neural network methods for predicting excess return on large stocks. Annals of Operational Research 78.

Fan, A., Palaniswami, M., 2000. A new approach to corporate loan default prediction from financial statements. In: Proocedings of CF–FFM.

Fernandez, A., Gomez, S., 2005. Portfolio selection using neural networks. Computers & Operations Research.

Galindo, J., Tamayo, P., 2000. Credit risk assessment using statistical and machine learning: Basic methodology and risk modeling applications. Computational economics 15.

Hamid, S., Iqbal, Z., 2004. Using neural networks for forecasting volatility of s & p 500 index futures prices. Journal of Business Research 57.

Han, J., Kamber, M., 2000. Data Mining: Concepts and Techniques. The Morgan Kaufmann Series in Data Management Systems. Morgan Kaufmann Publishers.

Heravi, S., Osborn, D., Birchenall, C., 2004. Linear versus neural network forecast for european industrial production series. International Journal of Forecasting 20.

Hykin, S., 1999. Neural Networks: A Comprehensive Foundation, 2nd Edition. Prentice Hall International, Inc.

Kaastra, I., Boyd, M., 1996. Designing a neural network for forecasting financial and economic time series. Neurocomputing 10.

Kim, T., Oh, K., Sohn, I., Hwang, C., 2004. Usefulness of artificial neural networks for early warning system of economic crisis. Expert System with Applications 26.

Kirkpatrick, S., Gelatt, C., Vecchi, M., 1983. Optimization by simulated annealing. Science, 13 May 1983 220 (4598), 671–680.

Kryzanowsky, L., Galler, M., Wright, D., 1993. Using artificial neural network to pick stocks. Financial analysts Journal 17.

Lam, M., 2004. Neural network techniques for financial performance prediction: integrating fundamental and technical analysis. Decision Support System 37.

L.H.Meyer, 2000. Why risk management is important for global financial institutions? BIS Review (68).

Mitchell, M., 1998. An introduction to genetic algorithms. MIT press, Cambridge, MA.

Moshiri, S., Cameron, N., Scuse, D., 1999. Static, dynamic and hybrid neural networks in forecasting inflaction. Computational Economics 14.

Odon, M., Sharda, R., 1990. A neural network for bankrupcy prediction. In: Proceedings of the International joint Conference on Neural Networks, San Diego.

Pang, S., Wang, Y., Bai, Y., 4-5 Novembre 2002. Credit scoring model based on neural network. In: Proc. of the First International Conference on Machine Learning and Cybernetics. Beijing.

Piramuthu, S., 1999. Financial credit-risk evaluation with neural and neurofuzzy systems. European Journal of Operational Research 112.

Rong-Zhou, L., Su-Lin, P., Jian-Min, X., 4-5 Novembre 2002. Neural network credit-risk evaluation model based on back-propagation algorithm. In: Proceedings of the First International Conference on Machine Learning and Cybernetics. Beijing.

Rotundo, G., 2004. Neural networks for large financial crashes forecast. Physica 344.

R.W.Ferguson, 2001. Credit risk management: models and judgment. BIS Review (85).

Salchenberger, L., Cinar, E., Lash, N., 1992. Neural networks: a new tool for predictiong thrift failures. Decision Science 23.

Saunders, A., 1999. Credit Risk Measurement. John Wiley & Sons.

Shazly, M. E., Shazly, H. E., 1999. Forecasting currency prices using a genetically evolved neural network architecture. International review of financial analysis 8 (1).

Steiner, M., Wittkemper, H. G., 1997. Portfolio optimization with a neural network implementation of the coherent market hypothesis. European Journal of Operational Research 1 (3), 27–40.

Tam, K., Kiang, M., 1992. Managerial applications of neural networks: the case of bank failure predictions. Management Science 38.

Tkacz, G., 2001. Neural network forecasting of canadian gpd growth. International journal of forecasting 17.

Vellido, A., Lisboa, P., Vaughan, J., 1999. Neural networks in business: A survey of applications (1992-1998). Expert Syst. Appl. 17, 51–70.

Werbos, P., 1988. Backpropagation, past and future. In: Proceedings of the IEEE International Conference on Neural Networks. IEEE Press.

White, H., 1988. Economic prediction using neural networks: the case of ibm daily stock returns. In: Proceedings of the Second IEEE International Conference on Neural Networks.

W.J.McDonough, 2003. Risk management, supervision and the new basel accord. BIS Review (5).

Wooldridge, J., 2002. Introductory Econometrics: A Modern Approach (second edition). South-Western College Pub.

Wu, C., Wang, X., 2000. A neural network approach for analyzing small business lending decisions. Review of Quantitative Finance and Accounting 15.

Yang, Z., Platt, M., Platt, H., 1999. Probabilistic neural networks in bankrupcy prediction. Journal of Business Research 44.

Zimmermann, H., Neuneier, R., 1999. Active portfolio-management with neural networks. In: Proceedings of Computational Finance, CF1999. Springer Verlag.