International Journal of Neural Systems, Vol. 21, No. 4 (2011) 311–317

DOI: 10.1142/S0129065711002857

ARTIFICIAL METAPLASTICITY NEURAL NETWORK APPLIED TO CREDIT SCORING

ALEXIS MARCANO-CEDEÑO*, A. MARIN-DE-LA-BARCENA†,
J. JIMENEZ-TRILLO‡, J. A. PIÑUELA§ and D. ANDINA¶

Group for Automation in Signals and Communications
Technical University of Madrid, ETSI Telecomunciación
Ciudad Universitaria, Madrid 28040, Spain
**a.marcano@gc.ssr.upm.es*
†amparo.mbarcenagrau@gc.ssr.upm.es
‡j.jimenez@gc.ssr.upm.es
§j.pinuela@gc.ssr.upm.es
¶andina@gc.ssr.upm.es

The assessment of the risk of default on credit is important for financial institutions. Different Artificial Neural Networks (ANN) have been suggested to tackle the credit scoring problem, however, the obtained error rates are often high. In the search for the best ANN algorithm for credit scoring, this paper contributes with the application of an ANN Training Algorithm inspired by the neurons' biological property of metaplasticity. This algorithm is especially efficient when few patterns of a class are available, or when information inherent to low probability events is crucial for a successful application, as weight updating is overemphasized in the less frequent activations than in the more frequent ones. Two well-known and readily available such as: Australia and German data sets has been used to test the algorithm. The results obtained by AMMLP shown have been superior to state-of-the-art classification algorithms in credit scoring.

Keywords: Artificial metaplasticity; neural plasticity; ANN training; learning algorithms; credit scoring.

1. Introduction

Credit scoring is a basic binary classification task in finance. The advantages of credit scoring include reducing the cost of credit analysis, enabling faster credit decisions, closer monitoring of existing accounts, and reducing possible risk.

Many studies have contributed to increasing the accuracy of the classification model with several kinds of statistical tools. A range of statistical techniques has been used to build-up credit scoring models, including linear probability models, logits, probits, and discriminant analysis. The first three techniques use historical data on credit performance and the characteristics of the borrower to estimate the probability of default. These results are then used to predict the probability of default for each new applicant by combining the estimated coefficients from the probability of default regression with the applicant's values given for the explanatory variables. Discriminant analysis differs in that instead of estimating a borrower's probability of default, it divides borrowers into high and low default-risk classes.

In recent years, Artificial Neural Networks (ANNs) have been one the most successful classifiers for credit scoring. The literature reveals that many machine learning methods have been suggested to tackle the credit scoring problem (e.g. Fuzzy Theory, Support Vector Machine, Perceptron Multilayer and ANNs),[1–8] and other related challenging problems, as prediction.[9]

There has been a lot of research in the area of credit scoring that, based on the Australian and

German credit data set, have applied ANNs. In 2000, West,[8] investigated credit scoring accuracy for five neural network models: multilayer perceptron, mixture-of-experts, radial basis function, learning vector quantization, and fuzzy adaptive resonance, which resulted in 85.84%, 86.68%, 87.14%, 82.97% and 75.39% respectively. In 2005, Ong *et al.*,[10] applied genetic programming (GP) for Building credit scoring models achieving a classification accuracy of 88.27%. In 2006, Huang *et al.*,[11] obtained an accuracy of 89.17% with two-stage genetic programming (2SGP) incorporating the advantages of the IFTHEN rules and the discriminant function. In 2007, Martens *et al.*,[12] obtained an accuracy of 85.7% by applying several methods based on rule extraction techniques for SVMs and introducing two methods (Trepan and G-REX), taken from the ANN domain. Hoffmann[13] proposed two evolutionary fuzzy rule learners for generating approximate and descriptive fuzzy rules, and reached an accuracy of 87.7%. Huang *et al.*,[14] presented a hybrid SVM-based credit scoring method to evaluate the applicants credit score; it took the applicants features as input, it is called GA-SVM, and resulted in an accuracy of 86.90% with a standard deviation of 4.22. In 2008, Peng *et al.*,[15] achieved a classification accuracy of 86.36% using Multicriteria Convex Quadric Programming (MCQP). Tsai and Wu[16] investigated the performance of a single baseline classifier and compared it with that of multiple classifiers obtaining an average accuracy of 87.25. In 2009, Khashman[6] applied a neural network model under the seven learning schemes and compared the performance of two neural networks; with one and two hidden layers following the ideal learning scheme (SHNN, DHNN). An accuracy of 81.03% was obtained. Nanni and Lumini[17] investigated the performance of several systems based on ensemble bankruptcy prediction and credit scoring classifiers. "Random Subspace" obtained the best result with an accuracy of 87.05%. Xu *et al.*,[18] reported an accuracy of 89.28% derived from the application of two algorithms which are called input weighted adjustor and class by SVM-based models. Luo *et al.*,[19] applied a new classifier clustering-launched classification (CLC) and SVM in order improve accuracy in credit scoring, achieving on average 86.52%. Tsai[20] used five well-known feature selection methods in bankruptcy prediction, which are t-test, correlation matrix, stepwise regression, principle component analysis (PCA) and factor analysis (FA) to examine their prediction performance. They obtained an accuracy of 89.27%, 84.74%, 89.31%, 86.08%, and 89.93% respectively. Ping[21] proposed a hybrid SVM-based Classifier. He used the neighborhood rough set to select input features, a grid search to optimize the RBF kernel parameters. The optimal input features were modeling parameters to solve the credit scoring problem and an accuracy of 87.52% was obtained. Chen and Li[22] applied a SVM classifier combined with conventional statistical LDA, Decision tree, Rough sets and F-score, thus obtaining an accuracy of 88.52%.

In this paper, we propose to apply an Artificial Metaplasticity implementation (AMP) on Multilayer Perceptron (AMMLP)[23] trained with Backpropagation to improve accuracy in credit scoring problems. This interpretation is modeled in the NNs training phase and to test our proposed AMMLP, the widely known Australian and German credit approval datasets[24] were used. To illustrate the advantages and drawbacks of our AMMLP algorithm its results are compared with Classic Backpropagation and other algorithms that have been recently successfully applied by other researchers on the same database.

This paper is organized as follows. Section 2 is an introduction to Biological Metaplasticity. Section 3 describes the proposed implementation of AMP in a MLP, whose learning process is based on error minimization. In Sec. 4, the real data sets used are described and referenced. Section 5 presents the application of the AMMLP algorithm to the real credit data sets. Section 6 refers to experimental results. Section 7 summarizes a comparison between ours and other results obtained by the latest researchers. Finally, Sec. 8 present our main conclusions.

2. Biological Metaplasticity

In neuroscience and other fields "metaplasticity" indicates a higher level of plasticity, expressed as a change or transformation in the way synaptic efficacy is modified. Metaplasticity is defined as *the induction of synaptic changes, that depends on prior synaptic activity*[23]. The induction of synaptic changes in the levels of neural activity is illustrated in Fig. 1.[25] This graphic shows a family of curves in which each

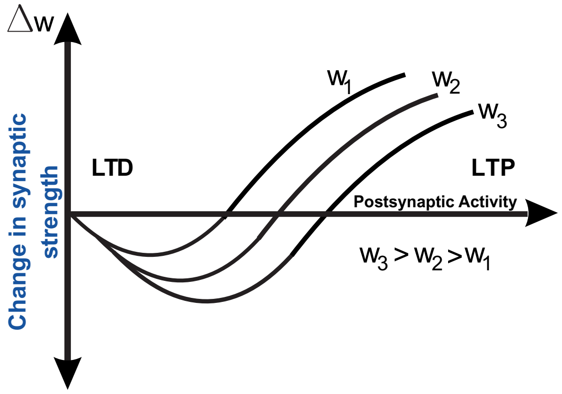

Fig. 1. Changes in synaptic strength due to postsynaptic activity in biological neurons.

curve indicates the variation in synaptic weight, $\Delta\omega$, respective of the neurons activation. Metaplasticity can be represented as variations in curve elongation with respect to the level of activity according to the weight strength of the synapse. The challenging problem of improving machine learning techniques[26–29] can benefit to a significant degree from understanding Metaplasticity.[25]

3. AMP in MLP Training: AMMLP

The proposed model defines artificial metaplasticity as a probabilistic learning procedure[29–34] that produces greater modifications in the synaptic weights with less frequent patterns than frequent patterns, as a way of extracting more information from the former than from the latter. As Biological metaplasticity, AMP then favors synaptic strengthening for low-level synaptic activity, while the opposite occurs for high level activity.[25] The model is applicable to general ANNs. Andina *et al.* proposed general AMP concepts for ANNs, and demonstrate them on Radar detection data.[35] In this paper it has been implemented and tested for an MLP over the Australian and German credit approval data sets. The AMP implementation applied tries to improve results in learning convergence and performance by capturing information associated with significant rare events. *It is based on the idea of modifying the ANN learning procedure such that infrequent patterns which can contribute heavily to the performance, are considered with greater relevance during learning without changing the convergence of the error minimization algorithm.* It is has been proposed in the hypothesis that the biological metaplasticity property maybe significantly due to an adaptation of nature to extract more information from infrequent patterns (low synaptic activity) that, according to Shannon's Theorem, implicitly carry more information.[35]

AMP is analytically introduced in an arbitrary MLP training, by the application of a weighting function $w_X^*(x)$[23]:

$$f_X^*(x) = \frac{A}{\sqrt{(2\pi)^N} \cdot e^{B\sum_{i=1}^N x_i^2}} = \frac{1}{w_X^*(x)} \qquad (1)$$

where $w_X^*(x)$ is defined as $1/f_X^*(x)$, being $f_X^*(x)$ an approximation of the input patterns probability density function (*pdf*), N is the number of neurons in the MLP input layer, and parameters A and $B \in R^+$ are algorithm optimization values which depend on the specific application of the AMLP algorithm. Values for A and B have been empirically determined. Equation (1) is a gaussian distribution. Then, $w_X^*(x)$ has high values for infrequent x values and close to 1 for the frequent ones and can therefore be straightforwardly applied in weights updating procedure for the AMP model of biological metaplasticity during learning.

As the *pdf* weighting function proposed is the distribution of the input patterns that does not depend on the network parameters, the AMMLP algorithm can then be summarized as a weighting operation for updating each weight in each MLP learning iteration as:

$$\Delta^* w = w^*(x)\Delta w \qquad (2)$$

being $\Delta w = w(t+l) - w(t)$ the weight updating value obtained by usual BPA and $w^*(x)$ the realization of the described weighting function $w^*(x)$ for each input training pattern x. Note that for Eqs. (8–10) the proposed implementation of metaplasticity in MLPs is analytically an implementation of importance sampling.[36]

4. Real Data Sets from UCI Repository

Credit data sets in the real world include various attributes. Two real-world data sets were selected for this research, the Australian and German credit data sets derived from the UCI Repository of Machine Learning Databases.[24] The Australian data set consists of 307 good applicants and 383 bad ones. Each applicant contains 15 features, including 6 nominal, 8 numeric attributes and the final one is class label (good or bad credit). These attributes names have been changed to meaningless symbolic data for the confidential reason. The second accuracy evaluation data set is the German credit scoring data set which is composed of 24 numeric features, including credit history, account balance, loan purpose, loan amount, employment status, personal information, age, housing and job. Additionally, 700 cases are creditworthy and the rest of 300 applicants are not.

5. Materials and Methods

To comparatively evaluate classifiers performance, all the classifiers presented in this particular case were trained with the same training data set and tested with the same evaluation data set. Each data set was divided into training and testing data randomly, in which there are 70–30% training and testing sets per data set (The data split was made according to recent researches applying metaplasticity algorithms).[37,38]

For this research, 50 AMMLPs were generated, with different weights and random values following a normal distribution (mean 0 and variance 1). In each experiment 50 networks were trained in order to achieve an average result that did not depend on the initial random value of the weights of the ANN. Two different criteria were applied to stop the training: in one case it was stopped when the error reached 0.01 (the error diminishes but cannot converge to 0), and in the other the training was conducted with a fixed number of 2.000 epochs.

Different network structures and parameters A and B of metaplasticity were tested empirically, in order to obtain the parameters and network structures appropriate for each case. Structure with 8 neurons in the hidden layer and parameter values $A = 8$ and $B = 0.5$ where found valid for all cases.

Table 1 shows the network structure, metaplasticity parameters, epochs, MSE and number of patterns used in the training and testing phases of the classifiers AMMLP and backpropagation neural network (BPNN) applied to each data set.

6. Experimental Results

In this section we present results of experiments to test the behavior of the AMMLP proposed method. It is well known that the most common and straightforward approach to evaluate the performance of a classifier is to use a test set of unseen instances that were not used during the training phase. For every instance in the test set, we compared the actual class to the class that was assigned by the trained classifier. Definitions related to the following tables are:

Accuracy: is the most common measure for evaluating classifiers.

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{3}$$

where: *True positive* (TP) is a correctly accepted credit record. *True negative* (TN) is a correctly rejected credit record. *False positive* (FP) is a rejectable credit record classified as accepted. *False negative* (FN) is an acceptable credit record classified as rejected. Other functions of true positives and false negatives are:

$$\begin{aligned} precision &= \frac{TP}{TP + FN}, \\ recall &= \frac{TN}{FP + TN}(\%) \end{aligned} \tag{4}$$

Table 1. Network structure and parameters of metaplasticity selected for each data set.

Classifiers	Network structure			Metaplas. parameters		MSE	Epoch	Numbers patterns	
	I	HL	O	A	B			Training	Testing
Australia									
AMMLPs	14	8	1	38	0.5	0.01	2000	483	207
BPNNs	14	8	1	—	—	0.01	2000	483	207
German									
AMMLPs	24	8	1	38	0.5	0.01	2000	700	300
BPNNs	24	8	1	—	—	0.01	2000	700	300

Table 2. Confusion matrices obtained in the best simulation by classifiers in Australia and German data set.

Classifiers	Desired results	Output results	
		Accepted	Rejected
Australia			
AMMLPs	Accepted credit	92	0
	Rejected credit	4	111
BPNNs	Accepted credit	84	8
	Rejected credit	12	103
German			
AMMLPs	Accepted credit	80	10
	Rejected credit	27	183
BPNNs	Accepted credit	70	20
	Rejected credit	50	160

Table 3. Average obtained for 50 simulations for each classifier.

Classifiers	Precision	Recall	Accuracy
Australia			
AMMLPs	93.48 ± 2.50	92.17 ± 2.08	92.75 ± 1.55
BPNNs	86.96 ± 3.63	81.74 ± 3.87	84.06 ± 2.29
German			
AMMLPs	86.67 ± 1.98	83.81 ± 1.72	84.67 ± 1.70
BPNNs	74.44 ± 2.98	72.38 ± 2.40	72.67 ± 2.50

Confusion matrix: A confusion matrix contains information about actual and predicted classifications performed by a classifier.

Tables 2 and 3 show the classification results.

As can be observed, AMMLP is superior to classic Backpropagation MLP training in all cases.

7. State-of-the-Art Comparison

To verify the performance of our algorithm we compared our results with those obtained by other researchers using the same database. In Table 4 the names of these researchers are shown, together with the accuracy obtained.

Table 4. Comparison classification accuracies obtained with our method and other classifiers from literature.

Author (year)	Australia	German
West (2.000)[8]	86.68	75.66
Ong *et al.* (2005)[10]	88.27	77.34
Huang *et al.* (2006)[11]	89.17	79.49
Martens *et al.* (2007)[12]	85.70	NA
Martens *et al.* (2007)[12]	86.70 ± 1.50	76.30 ± 1.60
Hoffman *et al.* (2007)[13]	86.90 ± 4.22	77.92 ± 3.97
Huang *et al.* (2007)[14]	86.70 ± 1.50	76.30 ± 1.60
Peng *et al.* (2008)[15]	86.38	94.00
Tsai and Wu (2008)[16]	87.25	87.29
Khasman (2009)[6]	89.28	NA
Nanni and Lumini (2009)[17]	87.05	74.67
Xu *et al.* (2009)[18]	89.28	NA
Luo *et al.* (2009)[19]	86.52	84.80
Tsai (2009)[20]	89.93	78.76
Ping (2009)[21]	87.52	76.60
Chen and Li (2010)[22]	86.52	76.70
This study (2011)*	92.75 ± 1.70	84.67 ± 1.50

NA: no applied.

*Average obtained for 50 simulations.

8. Conclusion

This paper presents an approach to credit risk evaluation using the novel AMMLP, an improvement over the classical MLP trained with the BP algorithm. The AMMLP is based on the biological Metaplasticity property of neurons, and when applied to the well-known Australian and German data sets, it has proven superior results compared to usual MLP. In the case of Australia data set, the results obtained by AMMLP were superior to all the results obtained by other researchers applied on the same database. In the German data set, the results obtained by AMMLP were only exceeded by two algorithms, but note that in these two algorithms, best-case instead of medium results are probably presented. Therefore, it is concluded that the proposed algorithm should be considered to support credit lending decisions.

Acknowledgments

This research has been supported by the Technical University of Madrid under 2010 funds aimed at Consolidated Research Groups. The first author wishes to thank the National Foundation of Science Technology and Innovation (FONACIT) of the Bolivariana Republic of Venezuela for its contribution in the development of their doctoral studies.

References

1. L. C. Thomas, A survey of credit and behavioral scoring: Forecasting financial, *International Journal of Forecasting* **16** (2000) 149–172.

2. C. M. Wang and Y. F. Huang, Evolutionary-based feature selection approaches with new criteria for data mining: A case study of credit approval data, *Expert Syst. Appl.* **36** (2009) 5900–5908.
3. S. T. Li, W. Shiue and M. H. Huang, The evaluation of consumer loans using support vector machines, *Expert Syst. Appl.* **30** (2006) 772–782.
4. A. Laha, Building contextual classifiers by integrating fuzzy rule based classification technique and k-nn method for credit scoring, *Adv. Eng. Inform.* **21** (2007) 281–291.
5. A. Vellido, P. J. Lisboa and J. Vaughan, Neural networks in business: A survey of applications (1992–1998), *Expert Syst. Appl.* **17** (1999) 51–70.
6. A. Khashman, A neural network model for credit risk evaluation, *Int. J. Neural Syst.* **19** (2009) 285–294.
7. T. S. Lee, C. C. Chiu, C. J. Lu and I. F. Chen, Credit scoring using the hybrid neural discriminant technique, *Expert Syst. Appl.* **23** (2002) 245–254.
8. D. West, Neural network credit scoring models, *Comput. Oper. Res.* **27** (2000) 1131–1152.
9. D. Wu, K. Warwick, Z. Ma, M. N. Gasson, J. G. Burgess, S. Pan and T. Z. Aziz, Prediction of Parkinson's disease tremor onset using radial basis function neural network based on particle swarm optimization, *Int. J. Neural Syst.* **20**(2) (2010) 109–116.
10. C. S. Ong, J. J. Huang and G. H. Tzeng, Building credit scoring models using genetic programming, *Expert Syst. Appl.* **29** (2005) 41–47.
11. J. J. Huang, G. H. Tzeng and C. S. Ong, Two-stage genetic programming (2SGP) for the credit scoring model, *Appl. Math. Comput.* **174** (2006) 1039–1053.
12. D. Martens, B. Baesens, T. Van Gestel and J. Vanthienen, Comprehensible credit scoring models using rule extraction from support vector machines, *Eur. J. Oper. Res.* **183** (2007) 1466–1476.
13. F. Hoffmann, B. Baesens, C. Mues, T. Van Gestel and J. Vanthienen, Inferring descriptive and approximate fuzzy rules for credit scoring using evolutionary algorithms, *Eur. J. Oper. Res.* **177** (2007) 540–555.
14. C. L. Huang, M. C. Chen and C. J. Wang, Credit scoring with a data mining approach based on support vector machines, *Expert Syst. Appl.* **33** (2007) 847–856.
15. Y. Peng, G. Kou, Y. Shi and Z. Chen, A multicriteria convex quadratic programming model for credit data analysis, *Decis. Support Syst.* **44** (2008) 1016–1030.
16. C. F. Tsai and J. W. Wu, Using neural network ensembles for bankruptcy prediction and credit scoring, *Expert Syst. Appl.* **34** (2008) 2639–2649.
17. L. Nanni and A. Lumini, An experimental comparison of ensemble of classifiers for bankruptcy prediction and credit scoring, *Expert Syst. Appl.* **36** (2009) 3028–3033.
18. X. Xu, C. Zhou and Z. Wang, Credit scoring algorithm based on link analysis ranking with support vector machine, *Expert Syst. Appl.* **36** (2009) 2625–2632.
19. S. T. Luo, B. W. Cheng and C. H. Hsieh, Prediction model building with clustering-launched classification and support vector machines in credit scoring, *Expert Syst. Appl.* **36** (2009) 7562–7566.
20. C. F. Tsai, Feature selection in bankruptcy prediction, *Knowledge-Based Syst.* **32** (2009) 120–127.
21. Y. Ping, Hybrid classifier using neighborhood rough set and SVM for credit scoring, in *Proc. 2nd Int. Conf. on Business Intelligence and Financial Engineering* (2009) 138–142.
22. F. L. Chen and F. C. Li, Combination of feature selection approaches with SVM in credit scoring, *Expert Syst. Appl.* **37** (2010) 4902–4909.
23. A. Marcano-Cedeño, J Quintanilla-Dominguez and D Andina, Breast cancer classification applying artificial metaplasticity algorithm, *Neurocomputing* **74**(8) (2011) 1243–1250. doi: 10.1016/j.neucom.2010.07.019.
24. UCI Machine Learning Repository. http://archive.ics.uci.edu/ml/datasets.html.
25. D. Andina and J. Ropero-Pelaez, On the biological plausibility of artificial metaplasticity, in *LNCS*, Springer, **6686** (2011) 119–128.
26. J. Menke and T. Martinez, Improving supervised learning by adapting the problem to the learner, *Int. J. Neural Syst.* **19**(1) (2009) 1–9.
27. F. Rasheed, M. Alshalalfa and R. Alhajj, Adapting machine learning technique for periodicity detection in biological sequences, *Int. J. Neural Syst.* **19**(1) (2009) 11–24.
28. R. Savitha, S. Suresh and N. Sundararajan, A fully complex-valued radial basis function network and its learning algorithm, *Int. J. Neural Syst.* **19**(4) (2009) 253–267.
29. S. Schliebs, N. Kasabv and M. Defoin-Platel, On the probabilistic optimization of spiking neural networks, *Int. J. Neural Syst.* **20**(6) (2010) 481–500.
30. F. J. Ropero-Pelaez and D. Andina, Probabilistic versus incremental presynaptic learning in biologically plausible synapses, in *LNCS*, Springer **6686** (2011) 80–89.
31. H. Adeli and A. Panakkat, A probabilistic neural network for earthquake magnitude prediction, *Neural Networks* **22** (2009) 1018–1024.
32. E. Lopez-Rubio and J. M. Ortiz-de-Lazcano-Lobato, Dynamic competitive probabilistic principal components analysis, *Int. J. Neural Syst.* **19**(2) (2009) 91–103.
33. E. Lopez-Rubio, Robust location and spread measures for nonparametric probability density function estimation, *Int. J. Neural Syst.* **19**(5) (2009) 345–357.
34. M. Ahmadlou and H. Adeli, Enhanced probabilistic neural network with local decision circles: A

robust classifier, *Integrated Computer-Aided Engineering* **17**(3) (2010).

35. D. Andina, A. Álvarez-Vellisco, A. Jevtić and J. Fombellida, Artificial metaplasticity can improve artificial neural networks learning, *Intell. Autom. Soft Co.* **15**(4) (2009) 683–696.
36. D. Andina, J. Martinez-Antorrena and I. Melgar, Importance sampling in neural detector training phase, in *Proc. of the World Automation Congress* 43–48 (2004). IEEE Catalog No. 04EX832C.
37. A. Marcano-Cedeño, A. Álvarez-Vellisco and D. Andina, Artificial metaplasticity MLP applied to image classification, in *Proc. 7th Int. Conf. on Industrial Informatics* 650–653 (2009). doi:10.1109/INDIN.2009.5195879.
38. A. Marcano-Cedeño, J. Quintanilla-Domínguez and D. Andina, Wood defects classification using artificial metaplasticity neural network, in *Proc. 35th Annual Conf. on of the IEEE Industrial Electronics Society*, 3422–3427. doi:10.1109/IECON.2009.5415189.