## Tackling business intelligence with bioinspired deep learning

J. Fombellida · I. Martín-Rubio · S. Torres-Alegre · D. Andina

#### **Abstract**

To tackle the complex problem of providing business intelligence solutions based on business data, bioinspired deep learning has to be considered. This paper focuses on the application of artificial metaplasticity learning in business intelligence systems as an alternative paradigm of achieving a deeper information extraction and learning from arbitrary size data sets. As a case study, artificial metaplasticity multilayer perceptron applied to the automation of credit approval decision based on collected client data is analyzed, showing its potential and improvements over the state-of-the-art techniques. This paper successfully introduces the relevant novelty that the artificial neural network itself estimates the pdf of the input data to be used in the metaplasticity learning, so it is much closer to the biologic reality than previous implementations of artificial metaplasticity.

**Keywords** Business intelligence  $\cdot$  Business data  $\cdot$  Bioinspired systems  $\cdot$  Metaplasticity  $\cdot$  Deep learning  $\cdot$  MLP  $\cdot$  AMMLP  $\cdot$  AMP  $\cdot$  Artificial neural network

#### 1 Introduction

Big data (BD) solutions for business services and utilities are one of the key issues in the development of nowadays Industry 4.0 revolution. System networks are in the need of new powerful intelligent systems to successfully process all the available information from the different users to study their behavior and improve service and associated business. Metaplasticity is an inherent property of biological neuron connections that consists in the ability of modifying the

plasticity of learning mechanism depending on the frequency of stimulus. It is a basic property that neuroscientists believe crucial in achieving the biological deep learning that is successful in real-world complex problems. This concept can be applied to artificial learning algorithms using a novel technique called artificial metaplasticity that achieves a deeper and more efficient learning on general data sets, not only from typical patterns in them, but also from atypical ones [1, 2].

This paper is organized as follows: big data and business intelligence issues are first generally stated in point Sect. 2, proposing in Sect. 3 that bioinspired systems can significantly contribute as a priority solution. To illustrate the proposal and correct interpretation of the deep learning concept in this paper, the case study of a paradigmatic business intelligence problem, the credit approval decision is presented after Sect. 3.1, ended with discussion and conclusions extraction.

#### 2 Big data and business intelligence

Although, the term big data, probably originated in lunchtab conversations at Silicon Graphics Inc., in the mid-1990s, it became widespread in 2011. In [3] it can be found various definitions of big data, highlighting the fact that volume is only one dimension of big data. Other dimensions, such as velocity, variety, veracity, variability and value, are equally important. Given that the discourse on big data is contextualized in predictive analytic frameworks, analytic have captured the imaginations of business and government leaders. Big data is characterized by data types considered unstructured not predefined or known—thus with a high degree of variety. Velocity of big data can be measured in less than a few minutes. Big data is used in machine learning and predictive analysis where organizations focus on what will happen versus what has happened [4].

New interactive Information and Communication Technology (ICT) infrastructures are being developed to support the optimal exploitation of the business results in different sectors. Business data processing empowers the associated business and creates opportunities up and down the industry value chain. These new utilities require management of services and technologies that span ICT Systems.

Technology solutions introduce new technical and organizational concepts. Companies are switching into IT-driven efficient management of complex information [5, 6], that interprets the information of the different actors, improving efficiency, reliability and economy. Business processes will need to be established to accommodate these technologies and market evolutions.

These applications require distribution of data and a computing platform [6, 7]. A business model aligns the utility goals by creating rate-based opportunities and reducing expenses and risks [8].

These utilities may inter-operate between them, and this can result in real cost savings [9]. Consider interoperability as the communication infrastructure that allows elements with different origins to interact seamlessly through standardization. This aspires to have three degrees of freedom, meaning that technology can move along any axis without disrupting existing technology along the other axis, including standards for data sharing, data governance and information sharing. Business can bring together data and information to serve analytic and business needs. So, business is about BD and devising new powerful intelligent systems (IS) to successfully process the available information in them.

In an emerging and rapidly changing business environment, organizations that have the competency to leverage information about their customers, marketplace, suppliers and business operations, will be able to benefit from the business opportunities and achieve sustained competitive advantage. For monitoring activities and assessing performance of a firm's business processes, the firm needs access to information and also an effective data management mechanism. Business intelligence (BI) is a way to collect,

correlate, and analyze events from multiple sources, leveraging existing tools while gaining end-to-end management of the entire network. Business intelligence is the practice of interpreting and visualizing data to make useful business-oriented decisions. BI decision support applications facilitate many such multidimensional analysis like online analytical processing, click-stream analysis, balance score-card, preparation, visualization, querying, reporting, charting, data mining for text content and voice, forecasting, geospatial analysis, enterprise portal implementation, knowledge management, digital dashboard access and other cross-functional activities.

Business intelligence (BI) is defined by literature and scholars in similar ways. Noble [10] defines BI as the ability to provide the business an information advantage; business doing what it has always done, but more efficient. BI is described as the value proposition that helps organizations tap into decision-making information that regular reporting does not provide. Singer [11] outlined that BI requires tools, applications, and technologies that focused on enhanced decision-making and is commonly used in supply chains, sales, finances and marketing. BI promotes consistent decisions and reduces risk caused by inconsistencies in the decision-making process [12].

Halpern [13] has extended BI to be an umbrella term which includes applications, tools, infrastructure, and practices to enable access and analysis of information to optimize performance and decision-making. Peters et al. [14] establish BI quality of BI infrastructure, BI functionality, and BI (managerial) self-service. They apply three dimensions of data, information and knowledge, to relate BI quality to support of the planning and reporting activities that underlie performance measurement information.

BI tasks fall into three categories: exploring and analyzing data, monitoring ongoing data flows through dashboards and communicating insights to others; and seven layers: IT and related infrastructure, data acquisition, data integration, data storage, data organizing, data analytics and data presentation [15]. Business intelligence systems (BIS) are well recognized as contributing to decision-making. BIS are most commonly identified as technological solutions holding quality information in well-designed data stores, enabling them to make the right decisions or take the right actions.

# 3 Potential of bioinspired systems and the case of automated credit scoring decision

BD and BI systems complexity makes their evaluation difficult. Due to the complex and multidisciplinary nature of these systems [16], it is very difficult to reach a

successful solution using classical engineering methods. Here is where bioinspired artificial intelligent systems that implement adaptive, reactive, complex and distributed processing may provide technological solutions. They have the potential to provide a more efficient and effective use of explicit and implicit knowledge in BD. In the last few decades, computer scientists have invented some new computing paradigms and algorithms such as evolutionary computing, swarm intelligence and artificial neural networks [17-20], but they have to be improved to capture the nuances of evolutionary theory and behavioral biology. Transformative bioinspirations for building computing and communication models will play an increasingly important role in the efforts to solve some real-world complex and difficult problems, which can be hard to tackle with existing computing algorithms.

This is the case that we are presenting in this paper, a novel learning algorithm inspired in the metaplasticity, a basic property of biological neuron connections that neuroscientists believe crucial in achieving the biological deep learning that allows biological brains to successfully deal with real-world complex problems. This concept is applied to artificial learning algorithms using a novel technique called artificial metaplasticity that achieves a deeper and more efficient learning on general BD, not only from typical patterns in them, but also from atypical ones. It also offers an alternative paradigm of deep learning as the improvement of overall learning does not rely on the increasing complexity of the artificial intelligence system by adding neuron layers, but in the more efficient learning modelization for each one of its components that produces an emergent ability on the intelligent system for deeper information extraction and classify information improvement on the dataset.

In this case study, we progress on previous works [1, 21–23]. In them, metaplasticity, a basic property of biological neuron connections that neuroscientists believe crucial in achieving the biological learning, is modeled as artificial metaplasticity. The learning is faster that classical methods and the resulting systems achieve performance that competes with the most powerful methods of the state of the art, for any multidisciplinary real data sets where it has been tested, and, not only from typical patterns in them, but also from atypical ones.

In the experiments that have been performed in the frame of this investigation, several state-of-the-art systems were compared to classify the Australian Credit Approval Database (ACAD) [24]. We will now show in this paper how to improve and generalize the results obtained in [23] by taking advantage of the artificial metaplasticity on multilayer perceptron (AMMLP) inherent distribution estimations of patterns in the dataset used in the design process. Inside each set of experiments, the first step is to

optimize the parameters used in the nominal backpropagation algorithm (BPA) for well-known multilayer perceptrons (MLPs) [25]. In the following experiment, a posteriori probability estimation of the input distributions is used to implement a metaplasticity learning algorithm and so generalizing the artificial metaplasticity applicability to any class of BD.

#### 3.1 Materials and methods

#### 3.1.1 Australian Credit Approval dataset

The financial crisis has put the focus of investigation in banking issues, specially in the ones related to the approval of credits. Until recently, the decisions related to credit loans were based on individual perceptions and the human capacity to assess the risk. The growing demand for credit has led to the use of a statistician method, known as credit scoring, to decide whether to or not to grant credit.

The method of punctuation credit is widely used for consumer loans, and it is getting more used for commercial loans. The credit score is a binary classification task of basic finance. An advantage of the credit score method is the reduction of the costs of credit analysis: faster credit decisions, greater control and reduced potential risks.

This dataset has been selected for this business intelligence application in order to compare the results of an automatic decision method with the classic one.

#### 3.1.2 Data Preparation

The database contains 690 cases, divided into two classes, 307 applicants "accepted" and 383 applicants "rejected". Each applicant has 15 features, including 6 nominal, 8 numeric attributes and the last one is the label of each class (accepted or rejected). This dataset is interesting because there is a good mix of attributes: continuous and nominal, nominal ratings with small and large values. Another important characteristic is that few values are missing.

To obtain results statistically independent of the distribution of the patterns, a tenfold cross-validation evaluation method has been considered. Using this method, the possible dependence of the results with the distribution of the samples is eliminated: all the samples are used to train the networks, and all the samples are used to evaluate the performance. Mean values are calculated to establish the final performance results.

It has empirically been proved that the classifiers based on neural networks produce better results if the training sets are balanced presenting the same number of patterns belonging to each one of the possible classes. In order to achieve this situation in the creation of the sets used to train and to evaluate the system, some accepted patterns will be repeated instead of eliminating some rejected patterns to get these balanced sets.

Using 10 sets of elements, we will create 10 different folders. Each one of the training sets consists of 9 of the 10 groups. The final evaluation of the performance of the network will use the other element. The 10 folders will be created with the variation of the initial set that is used for evaluation and not for training. The networks are trained from the same initial conditions presenting the data corresponding to each of the 10 folders.

#### 3.1.3 Artificial metaplasticity neural network model

Artificial neural networks (ANNs), widely used in pattern classification, are biologically inspired distributed parallel processing networks based on the neuron organization and decision-making process of the human brain [22].

The concept of biological metaplasticity was defined in 1996 by Abraham [26] and now is widely applied in the fields of biology, finances, neuroscience, physiology, neurology and others [26, 27]. Recently, Andina [21], Ropero-Peláez [22] and Marcano-Cerdeño [28] have introduced and modeled the biological property metaplasticity in the field of ANNs, obtaining excellent results.

The activation function used in all the neurons of the system is sigmoidal, input patterns set normalized, initialization of the weights of the neurons is random but included in a limited interval and parameter value  $\sigma$  is constant and equal to 1. So the range of inputs to the activation function  $\sigma \sum \omega_i x_i$  will be limited to the linear range.

To introduce artificial metaplasticity (AMP) in an arbitrary MLP training, all that has to be done is to introduce a weighting function  $\frac{1}{f_X^*}$  in the MLP learning equation, based on the premise that  $f_X^*$  that has the properties of a probability density function [1]. Then, it is up to the designer to find a function that improves MLP learning. Several have been already proposed [1, 21–23], and we use a derived approach in this paper:

$$w_{ij}^{(s)}(t+1) = w_{ij}^{(s)}(t) - \eta \frac{\partial E^*[W(t)]}{\partial w_{ij}^{(s)}} = w_{ij}^{(s)}(t) - \eta \frac{1}{f_X^*} \frac{\partial E[W(t)]}{\partial w_{ij}^{(s)}}$$
(1)

where  $s, j, i \in N$  are the MLP layer, node and input counters, respectively, for each W(t) component,  $w_{ij}^{(s)}(t) \in R$  and being  $\eta \in R^+$  the learning rate.

The BPA presents some limitations and problems during the MLP training. The AMMLP algorithm aims to improve BPA by including a variable learning rate  $\eta(x) = \frac{\eta}{f_X^*}$  in the training phase affecting the weights in each iteration step based on an estimation of the inherent distribution of training patterns.

# 3.1.4 Artificial metaplasticity using a posteriori probability distribution

Marcano-Cedeño et al. [23] present an implementation of the AMMLP theory that is based on a hypothesis that considers that the probability density function of the dataset under study is known previously by the user; in this case, it has been considered that the pdf of the data is a Gaussian function. Our paper presents a different approach:

$$\widehat{y} \cong P(H_l/x) = f_X^*(x) \tag{2}$$

where  $\hat{y}$  is the output of the neuron of the final layer of the network that estimates the a posteriori probability. Equation 2 takes advantage of the inherent a posteriori probability estimation for each input class of MLP outputs [21, 22]. Note that if this is not the case, as it happens in first steps of BPA training algorithm, the training may not converge. In this first steps, the outputs of the MLP do not provide yet any valid estimation of the probabilities. In these first steps of training, it is better either to apply ordinary BPA training or to use another valid weighting function until BPA starts to minimize the error objective.

For two classes in the training set, we only need one output. If the desired output activation corresponds to input vectors  $x \in H_1$ , thus y = 1, then the AMP is implemented by:

$$f_X^*(x) = \widehat{y} \tag{3}$$

and for the complementary class patterns  $x \in H_0$ , thus y = 0:

$$f_{\mathbf{v}}^*(x) = (1 - \widehat{\mathbf{v}}) \tag{4}$$

This method also allows using AMMLP without any kind of previous knowledge of the pdf of the input dataset.

#### 4 Results

#### 4.1 Network characteristics

Structure of the network:

- Number of inputs: equal to the number of attributes of the pattern (14).
- Number of hidden layers: 1
- Number of neurons included in the hidden layer: 8
  neurons are considered ideal for a tradeoff between the
  flexibility and the complexity of the system.
- Number of neurons in the output layer: 1 to classify in two classes.
- Activation function: Sigmoidal with output included in the interval (0, 1).

Conditions considered to finalize the network training:

 Reach a defined number of inputs presented to the network to have enough iterations to reach a stable output without overspecializing the network.

#### 4.2 Evaluation method

The following hypotheses are defined to build a confusion matrix model:

- True Positive (TP) H(1/1): The pattern is accepted and has been classified as accepted.
- False Positive (FP) H(1/0): The pattern is rejected and has been classified as accepted.
- False Negative (FN) H(0/1): The pattern is accepted and has been classified as rejected.
- True Negative (TN) H(0/0): The pattern is rejected and has been classified as rejected.

To evaluate the performance, two measures are used and defined as: Sensitivity(SE) = TP/(TP+FN) (%) that evaluates the performance of the network identifying the malign patterns, and Accuracy(AC) = (TP+TN)/(TP+TN+FP+FN) (%) that evaluates the performance of the network classifying both malign and benign patterns. TP, TN, FP, and FN stand for true positive, true negative, false positive and false negative, respectively.

Fifty initial networks have been trained and evaluated with the 10-folder cross-validation algorithm. From the results obtained for the same network with each one of the folders, the mean confusion matrix is obtained for each network. Once these 50 mean values are calculated, an additional calculation is made and the final mean value is obtained, so we assure that the results are independent of the initial random value in the creation of the networks.

We have considered sensitivity as the driver figure in these experiments.

## 4.2.1 Nominal backpropagation algorithm

The results obtained for sensitivity correspond to value  $\eta = 1$ :

- % Accuracy = 83.3708
- % Sensitivity = 91.8433

The complete results for this case have been presented in [23], but it is now relevant to add the graphical evolution of learning, shown in Fig. 1.

Confusion matrix for this learning rate is shown in Table 1.

#### 4.2.2 AMP based on the output of the network

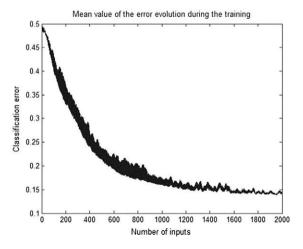
In order to check the theoretical approach of applying the a posteriori estimation of the probability distribution as AMP function, it is necessary to reach a point when the network has started to learn. This experiment will be divided in two parts:

- The initial part of the training will use the backpropagation classic algorithm using learning rate η = 23 (empirically determined, as in usual BPA) until classification error reaches 0.3.
- The second and principal part of the training will use Eqs. 1, 2, 3 and 4 from error 0.3 to the finalization of the training.

The results obtained are:

- % Accuracy = 85.9008
- % Sensitivity = 91.9060

Confusion matrix for these parameters rate is shown in Table 2.



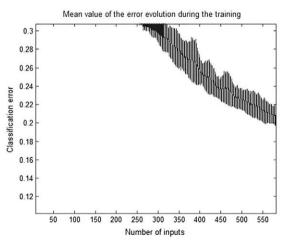


Fig. 1 Global evolution of the classification error of the network output (left sub-figure). Detail of the classification error evolution for the first iterations of the training (right sub-figure)— $\eta = 1$  nominal backpropagation—2000 training iterations

Table 1 Confusion matrix  $\eta = 1$ —nominal backpropagation—2000 training iterations

TP = 91.8434%	FP = 21.1018%
FN = 8.9368%	TN = 78.8982%

**Table 2** Confusion matrix  $\eta=23$ —output as probability estimation—2000 training iterations

TP = 91.9060%	FP = 20.1044%
FN = 8.0940%	TN = 79.8956%

The evolution of the classification error during the training phase is shown in Fig. 2. The ROC of this experiment can be found in Fig. 3, and the area under the associated curve is 0.91.

A comparison between the AMMLP classification results for the ACAD dataset and other classification methods presented in the literature can be found in Table 3.

#### 5 Discussion

New computing bioinspired systems and algorithms can be used to tackle real-world complex problems arisen with the Industry 4.0 revolution that involves the concept of BI based on data collected from clients, as it is the case of automated credit scoring. Considering how the application of artificial metaplasticity in artificial neural networks improves the results, it can be observed that:

 The results show that the theory exposed in [21, 22] is applicable to this dataset and network structure and that

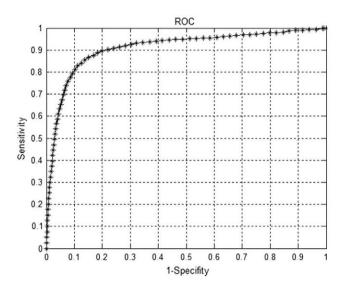
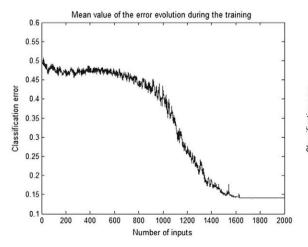


Fig. 3 ROC  $\eta=23$ —AMP based on the output of the network—2000 training iterations

the pdf information provided by the network itself can be used to improve the results compared to the basic BPA.

- With AMP training method, we can observe an improvement in the quality of the performance (for both accuracy and sensitivity).
- The slope of the learning error curve increases when error reaches 0.3 (change of function), showing that AMP performs a considerable quicker learning.
- Observing the detail of the evolution of the classification error, in the first experiment (classic BPA) there are a lot of peaks in the learning evolution; however, using the AMP this evolution is more smooth and natural.



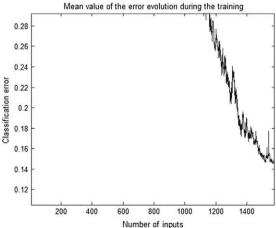


Fig. 2 Global evolution of the classification error of the network output (left sub-figure). Detail of the classification error evolution for the first iterations of the training (right sub-figure)—AMP based on the output of the network—2000 training iterations

**Table 3** Comparison of classification methods for the ACAD dataset

References	Method	Sensitivity	Accuracy (%)
Fombellida	AMMLP Output	91.90%	85.90
West [29]	MOE	86.70%	86.68
Ong et al. [30]	GP	_	88.27
Huang et al. [31]	2SGP	_	89.17
Martens et al. [32]	SVM	_	85.70
Hoffman et al. [33]	Bayes	_	86.70
Huang et al. [34]	GA-SVM	_	86.90
Peng et al. [35]	MCQP	87.00%	86.38
Tsai and Wu [36]	Multiple Classifiers	_	87.25
Nanni and Lumini [37]	LMNC	_	87.44
Xu et al. [38]	HARA	_	89.28
Luo et al. [39]	CLC	_	86.52
Tsai [40]	PCA	_	89.93
Ping [41]	RBF-SVM	_	87.52
Chen and Li [42]	LDA+SVM	_	86.52
Marcano-Cedeño et al. [23]	AMMLP Gaussian	_	84.67
Wang et al. [43]	Boosting	_	86.56
Wang et al. [44]	RS-Bagging DT	_	88.17
Vukovic et al. [45]	Pure CBR-PF-GA	_	88.12
Tsai [46]	SOM+MLP	_	91.61
Semenkina and Semenkin [47]	SelfCEA	_	90.22
Zhao et al. [48]	MLP Scoringt	_	87.00
Liang et al. [49]	Naive Bayes	_	86.09
Xiao et al. [50]	ECSC	_	87.62
Chen et al. [51]	PBIL-AIS	_	86.23

#### **6 Conclusions**

In this paper, the artificial metaplasticity on MLP based on the output of the network has been applied as a case of a business intelligence application on data compiled from clients, concretely to the problem of credit approval database. An artificial metaplasticity learning implementation on MLPs shows improved learning over the classical MLPs and other state-of-the-art systems, not only in final performance, but also in the evolution of learning. Applied to a challenging classification problem, credit approval, the learning has been more natural (regular), providing better performance in accuracy and sensitivity. This allows useful applications of the theory to data sets with a reduced amount of patterns without creating over-specialization in the final network. This method also allows using AMMLP without any kind of previous knowledge of the pdf of the input patterns. Due to the general nature of the metaplasticity concept, that is as wide as plasticity itself, these results have the potential to be generalized in deep Learning and therefore contribute to the development of Industry 4.0 revolution through improvements over state of the art in big data information extraction and classification.

#### Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

#### References

- Andina D, Alvarez-Vellisco A, Jevtic A, Fombellida J (2009) Artificial metaplasticity can improve artificial neural network learning. Intell Autom Soft Comput Special Issue Signal Process Soft Comput 15(4):681–694
- Quintanilla-Dominguez J, Cortina-Januchs MG, Ojeda-magaña B, Jevtić A, Vega-Corona A, Andina D (2010) Microcalcification detection applying artificial neural networks and mathematical morphology in digital mammograms. In: World Automation Congress (WAC), pp 1–6
- 3. Gandomi A, Haider M (2015) Beyond the hype: big data concepts, methods, and analytics. Int J Inf Manag 35(2):137–144
- Davenport TH (2014) Big data @ work: dispelling the myths, uncovering the opportunities. Harvard Business Review Press, Brighton
- Groznik A (2014) Towards multi utility management in Europe. In: ICEE 2014: international conference on energy and environment, vol 2, no 8, p 195
- Cardenas J, Gemoets L, Rosas J, Sarfi R (2014) A literature survey on Smart Grid distribution: an analytical approach. J Clean Prod 65:202–216

- Martín-Rubio I, Florence-Sandoval A, Jiménez-Trillo J, Andina D (2015) From Smart Grids to business intelligence, a challenge for bioinspired systems, bioinspired computation in artificial systems. In: International work-conference on the interplay between natural and artificial computation, IWINAC, proceedings, part II
- Starace F (2008) The utility industry in 2020. In: Bausch A, Schwenker B (eds) Handbook utility management. Springer, Berlin
- López G, Moreno J, Amarís H, Salazar F (2015) Paving the road toward Smart Grids through large-scale advanced metering infrastructures. Electr Power Syst Res 120:194–205
- 10. Noble J (2006) The core of IT. CIO Insight, pp 15-17
- Singer T (2001) Information engineering: the search for business intelligence. Plant Eng 55(11):34–36
- Rubin E, Rubin A (2013) The impact of Business Intelligence systems on stock return volatility. Inf Manag 50:67–75
- Halpern F (2015) Next-generation analytics and platforms for business success: tDWI research report. www.tdwi.org. Accessed Jan 2015
- Peters MD, Wieder B, Sutton SG, Wakefield J (2016) Business intelligence systems use in performance measurement capabilities: implications for enhanced competitive advantage. Int J Account Inf Syst 21:1–17
- Glaser J, Stone J (2008) Effective use of business intelligence. Health Finance Manag 62(2):68–72
- Kezunovic M, Vittal V, Meliopoulos S, Mount T (2014) The big picture: smart research for large-scale integrated smart grid solutions. IEEE Power Energy Mag 10(4):22–34
- 17. Abbott D (2014) Applied predictive analytics: principles and techniques for the professional data analyst. Wiley, New York
- Choy K, Lee WB, Lo V (2003) Design of an intelligent supplier relationship management system: a hybrid case based neural network approach. Expert Syst Appl 24(2):225–237
- Bonabeau E, Meyer C (2001) Swarm intelligence: a whole new way to think about business. Harvard Bus Rev 79(5):106–115
- Jevtić A, Andina D (2010) Adaptive artificial ant colonies for edge detection in digital images. In: IECON 2010-36th annual conference on IEEE industrial electronics society, pp 2813–2816
- Andina D, Ropero-Pelaez J (2012) On the biological plausibility of artificial metaplasticity learning algorithm. Neurocomputing. https://doi.org/10.1016/j.neucom.2012.09.028
- Ropero-Pelaez J, Andina D (2012) Do biological synapses perform probabilistic computations? Neurocomputing. https://doi. org/10.1016/j.neucom.2012.08.042
- Marcano-Cedeño A, Marin-De-La-Barcena A, Jimenez-Trillo J, Piñuela JA, Andina D (2011) Artificial metaplasticity neural network applied to credit scoring. Int J Neural Syst 21:311
- https://archive.ics.uci.edu/ml/datasets/Statlog+(Australian+Credit +Approval). Accessed Feb 2016
- Andina D, Pham D (2007) Computational intelligence for engineering and manufacturing. Springer, Berlin
- Abraham WC (1996) Activity-dependent regulation of synaptic plasticity (metaplasticity) in the hippocampus. In: The hippocampus: functions and clinical relevance. Elsevier Science, Amsterdam, pp 15–26
- Kinto E, Del-Moral-Hernandez E, Marcano-Cedeño A, Ropero-Pelaez J (2007) A preliminary neural model for movement direction recognition based on biologically plausible plasticity rules. Proc IWINAC 45(28):628–636
- Benchaib Y, Marcano-Cedeño A, Torres-Alegre S, Andina D (2013) Application of artificial metaplasticity neural networks to cardiac arrhythmias classification. Lect Notes Comput Sci 79(30):181–190
- West D (2000) Neural network credit scoring models. Comput Oper Res 27:1131–1152

- Ong CS, Huang JJ, Tzeng GH (2005) Building credit scoring models using genetic programming. Expert Syst Appl 29:41–47
- Huang JJ, Tzeng GH, Ong CS (2005) Two-stage genetic programming (2SGP) for the credit scoring model. Appl Math Comput 174(2):1039–1053
- Martens D, Baesens B, Van Gestel T, Vanthienen J (2007)
   Comprehensible credit scoring models using rule extraction from support vector machines. Eur J Oper Res 183(3):1466–1476
- Hoffmann F, Baesens B, Mues C, Van Gestel T, Vanthienen J (2007) Inferring descriptive and approximate fuzzy rules for credit scoring using evolutionary algorithms. Eur J Oper Res 177:540–556
- Huang CL, Chen MC, Wang CJ (2007) Credit scoring with a data mining approach based on support vector machines. Expert Syst Appl 33(4):847–856
- Peng Y, Kou G, Shi Y, Chen Z (2008) A multi-criteria convex quadratic programming model for credit data analysis. Decis Support Syst 44(4):1016–1030
- Tsai CF, Wu JWZ (2008) Using neural network ensembles for bankruptcy prediction and credit scoring. Expert Syst Appl 44(4):2639–2649
- Nanni L, Lumini A (2009) An experimental comparison of ensemble of classifiers for bankruptcy prediction and credit scoring. Expert Syst Appl 36:3028–3033
- Xu X, Zhou C, Wang Z (2009) Credit scoring algorithm based on link analysis ranking with support vector machine. Expert Syst Appl 36(2):2625–2632
- Luo ST, Cheng BW, Hsieh CH (2009) Prediction model building with clustering-launched classification and support vector machines in credit scoring. Expert Syst Appl 36(4):7562–7566
- Tsai CF (2009) Feature selection in bankruptcy prediction. Knowl Based Syst 22(22):120–127
- Ping Y (2009) Hybrid classifier using neighborhood rough set and SVM for credit scoring. In: Proceedings of BIFE 09, international conference on business intelligence and financial engineering, pp 138–142
- Chen FL, Li FC (2009) Combination of feature selection approaches with SVM in credit scoring. Expert Syst Appl 37(7):4902–4909
- Wang G, Hao J, Ma J, Kiang H (2011) A comparative assessment of ensemble learning for credit scoring. Expert Syst Appl 38(1):223–230
- Wang G, Ma J, Huang L, Xu K (2012) Two credit scoring models based on dual strategy ensemble trees. Knowl Based Syst 26:61–68
- Vukovic S, Delibasic B, Uzelac A, Suknovic M (2012) A casebased reasoning model that uses preference theory functions for credit scoring. Expert Syst Appl 39(9):8389–8395
- Tsai CF (2014) Combining cluster analysis with classifier ensembles to predict financial distress. Inf Fusion 16:46–58
- Semenkina M, Semenkin E (2014) Hybrid self-configuring evolutionary algorithm for automated design of fuzzy classifier. Int Conf Swarm Intell ICSI 2014:310–317
- Zhao Z, Xu S, Kang BH, Kabir MMJ, Liu Y, Wasinger R (2015) Investigation and improvement of multi-layer perceptron neural networks for credit scoring. Expert Syst Appl 42:3508–3516
- Liang D, Tsai CF, Wu HT (2015) The effect of feature selection on financial distress prediction. Knowl Based Syst 73:289–297
- Xiao H, Xiao Z, Wang Y (2016) Ensemble classification based on supervised clustering for credit scoring. Appl Soft Comput 437:73–86
- Chen MH, Chang PC, Wu JL (2016) A population-based incremental learning approach with artificial immune system for network intrusion detection. Eng Appl Artif Intell 51:171–181