# Computational Intelligence Applied to Audits: Pattern Classification Using Artificial Neural Networks.

**Renan Martins de Sousa** is a public servant at the Federal Court of Accounts in Brazil, with a degree in Electrical Engineering from Ceará State University and has a specialist degree in regulation of telecommunications services from the National Institute of Telecommunications (Inatel) and from the International Union of Telecommunications (UIT)..

## ABSTRACT

The rise in the demand for greater transparency of information held by the public bodies has been leading to an increased availability of various databases. This fact, allied with the advances on the processing capacity, has been promoting interest in the use of computational intelligence in less explored areas, such as the audit activities related to public administration control. The ability of the artificial neural networks to classify patterns may help control agencies to perform their duties more efficiently. Typical applications of classification standards in auditing are related to fraud detection, auditing of financial statements and risk assessment, among others. The Federal Court of Accounts of Brazil (TCU), aware of this reality, has been carrying out several actions to develop skills associated with data analysis.

**Keywords**: Applied Computational Intelligence. Neural Networks. Pattern Classification. Audit. Federal Court of Accounts.

## 1. INTRODUCTION

The technological advances made in recent decades, both regarding computer processing and to data storage capacity, combined with the increasing availability of information, pose a great challenge for those who need to treat them and issue an opinion based on such treatment.

According to Byrnes *et al* (2014), data science, likewise, has advanced enormously, incorporating theories, techniques and software applications from many disciplines, including data analysis, business intelligence, mathematics and probability, statistical learning (including pattern recognition), data visualization and analysis and treatment of large data sets, such as data mining and visualization.

The application of these theories may be used by control bodies, so that they may present new types of evidence and conduct more focused audits, and may result in more reliable opinions on the audit objects, even when subjected to high performance requirements, such as time, accuracy and cost.

This article conceptualizes pattern classification as an applied computational intelligence tool and briefly outlines the origins, features and training of artificial neural networks, especially the multilayer perceptron (MLP) network, illustrating its use in auditing. Moreover, it shows how the Federal Court of Accounts (TCU) has encouraged the treatment of information databases to render its performance more effective, timely and intelligent.

## 2. WHAT DOES PATTERN CLASSIFICATION MEAN?

Automatic recognition, description, grouping and pattern classification are very important tools for a wide range of engineering and sciences disciplines, such as biology, psychology, medicine, computer vision, artificial intelligence and remote sensing, among others.

However, what are patterns? Jain, Duin and Mao (2000 apud Watanabe, 1985) define pattern as "the opposite of chaos; an entity to which a name may be assigned". For example, a pattern may be a voice signal, a DNA sample, a text document, a video clip, a fingerprint, a handwritten word etc.

Once there is a pattern, its recognition (or classification) may be carried out with or without supervision. Without supervision, the pattern is associated with a class unknown until then, a technique known as clustering. In this case, the problem lies in categorization, in which classes are defined by the system designer or are learned based on the similarity of patterns. With supervision, of greater interest to this paper, the pattern is identified as part of a predefined class. The separation of patterns among classes may be carried out by a discriminant analysis.

Examples of applications in this field include data mining, document classification, financial forecasts, organization and search in multimedia databases and biometrics, among many others.

In short, pattern recognition consists in studying how machines can observe the environment, learn to distinguish patterns of interest and make reliable and reasonable decisions about the categories of these patterns.

The design of a pattern recognition system involves three basic macro steps: (i) data acquisition and preprocessing; (ii) data representation and (iii) decision-making. In general, the problem domain determines the choice of the method applied in each of these steps. The models most commonly used to make decisions on pattern recognition are: template matching, structural matching, statistical classification and artificial neural networks.

## 3. ARTIFICIAL NEURAL NETWORKS

Artificial neural networks are numerical processing systems, and consist in connecting a large number of simple processors. These interconnections have biological origin, namely, the nervous system of living things. The concept underlying these systems is that complex processing may be obtained when many simple, highly interconnected processors are combined, and such concept is referred to in literature as connectionism.

Connectionism, which employs distributed and parallel processing, as opposed to centralized processing, represents with certain ease the features of cognitive processes, such as the ability to simultaneously consider a number of restrictions or combine various sources of knowledge. It is also able to represent the ability to generalize.

In spite of their biological inspiration, the current models of neural networks do not represent the aspects and structures widely known to brain physiology, such as the spatial organization of neurons and interconnections and the existence of various types of signals between these “processors”. This is due to the search for balance between the biological plausibility of such models and their mathematical treatment.

The figures below illustrate the schematic view of a typical neuron (Figure 1) and an abstraction for computational purposes (Figure 2):

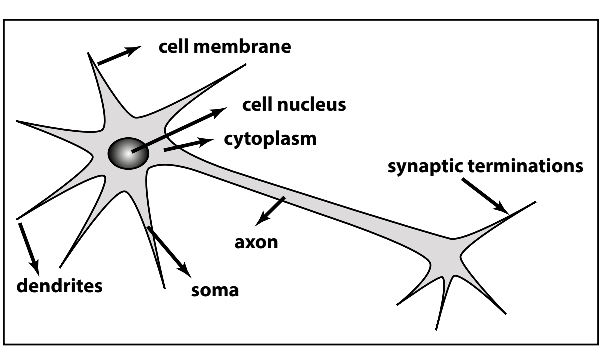
The fundamentals of the current neuron model and the principle of association were set out in the late nineteenth century, in the works of James (1890). However, the first major studies about the mathematical ability of networks of neuron-inspired processing elements were only outlined in the first half of the twentieth century. McCulloch and Pitts (1943) showed that associations of such artificial neurons could implement any finite logical function, and this may be considered the first theoretical success on connectionism.

However, the most important step on the subject was taken when Rosenblatt (1958) introduced the first concrete neural model, called the perceptron, which initially had only two layers: input and output. This model was able to classify patterns from examples, but its use was hampered by some limitations exposed in the work of Minsky and Papert (1969), which led the research to be discontinued. The interest in connectionism was revived with the use of the error backpropagation training algorithm, presented by Rumelhart and McCelland (1986),which extended Rosenblatt’s perceptron to multiple layers (multi-layer perceptron - MLP), overcoming the limitations of the original model and enabling the development of applications in various branches of knowledge.

After this brief history, we find it important to place the artificial neural networks in the universe of methods applicable to pattern recognition and describe some of their main features.

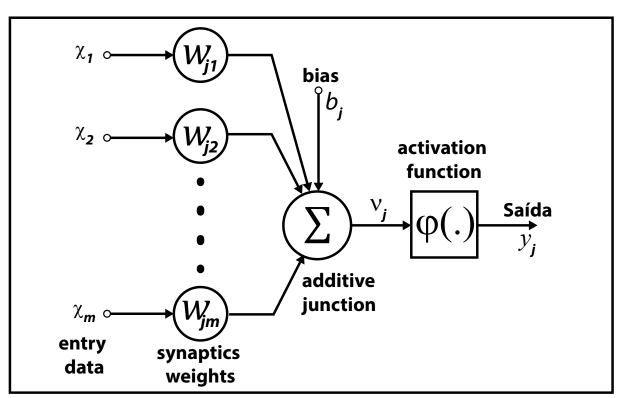
While there are several approaches to pattern recognition, this paper is only interested in the statistical and neural approaches. The latter has different opera-

**Figure 1:**
A typical neuron

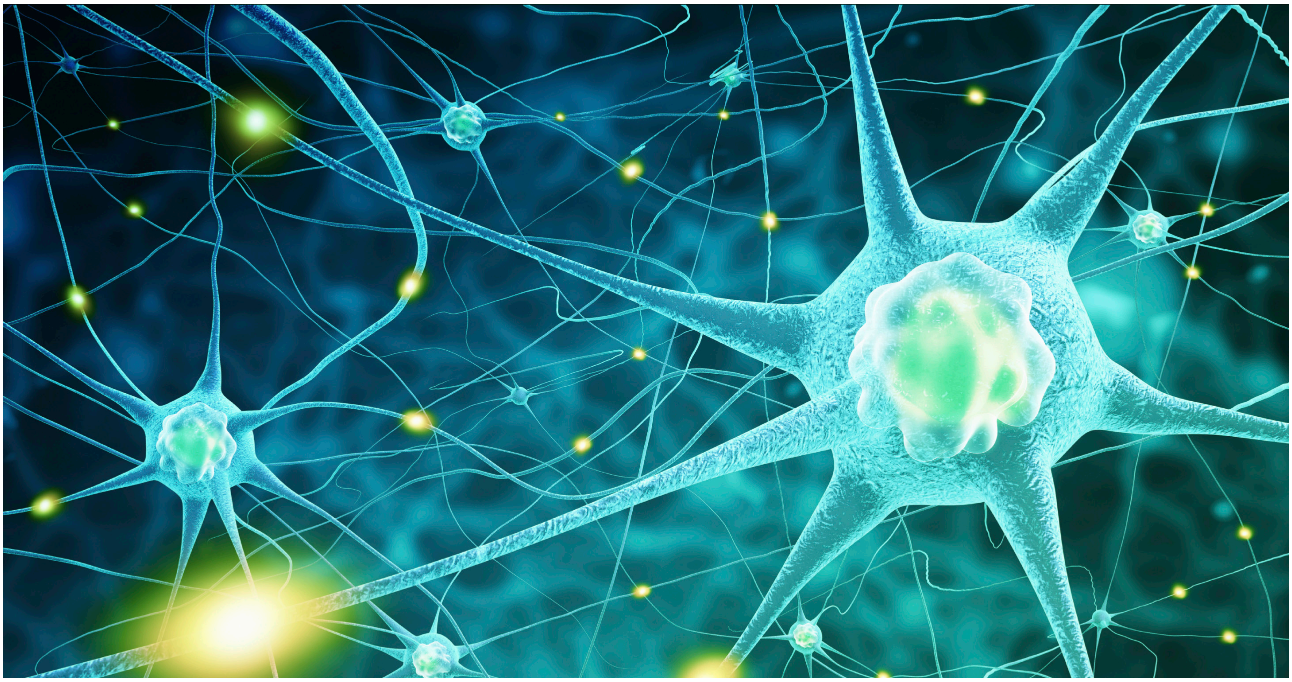



Source: http://blogdopetcivil.com/2013/07/05/redes-neurais-artificiais/

**Figure 2:**
Nonlinear model of a neural network node




Source: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1806-11172011000100009



ting principles, although it uses models that are equivalent to those of the latter.

In the statistical approach, each pattern is represented by a set of "n" features–or measures–and viewed as a point in an *n*-dimensional space. According to Jain, Duin and Mao (2000), the purpose of this representation is to choose the features so that the patterns (feature vectors) belonging to different classes occupy compact, well-ordered regions of then-dimensional space of features. Thus, the effectiveness of this choice is greater the easier it is to separate the patterns belonging to different classes.

Based on a subset of patterns of various classes (training set), the purpose of the statistical approach is to define decision boundaries in the *n*-dimensional space of features capable of separating patterns belonging to different classes. In the statistical approach, decision boundaries are determined by the pattern probability distributions, which must be known or learned *a priori*.

The neural approach, on the other hand, uses a nonlinear discriminant analysis, namely, a geometric analysis. The discriminant functions are constructed by a linear combination of basic nonlinear functions, and have the following form:

$$g(\boldsymbol{x}) = \sum_{i=1}^{m} w_i \varphi_i(\boldsymbol{x}, \boldsymbol{u_i}) \ (1)$$

The combination of x and parameter vector $\boldsymbol{u}$ results in a scalar product, i.e, $\phi_i\ (\boldsymbol{x},\boldsymbol{u}_i)=\phi_i\ (\boldsymbol{x}^T\ \boldsymbol{u}_i)$. The form of the nonlinear function $\phi$, referred to as the activation function, is chosen *a priori* and the optimization procedure determines, simultaneously, the $w_{ij}$ and $\boldsymbol{u}_i$ parameters. In other words, the basic functions are selected previously, but their parameters are adaptable during the optimization phase.

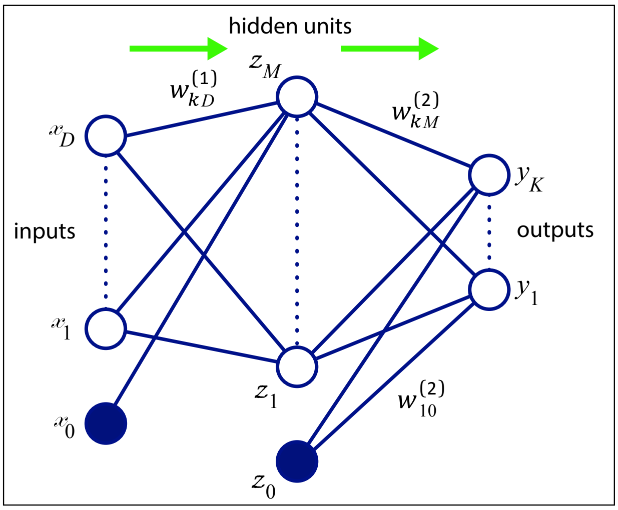
In sum, the main features of neural networks are: (i) the ability to learn nonlinear relationships between inputs and outputs; (ii) the use of sequential training procedures and (iii) the ability to adapt to the given data.

## 4. MULTILAYER PERCEPTRON NETWORK, LEARNING AND THE ERRORBACKPROPAGATION ALGORITHM

According to Bishop (2006), the most successful model using the neural approach, within the scope of pattern recognition, is the multilayer perceptron (MLP) neural network, of interest to this paper. A neural network is mainly specified by its topology, the features of its nodes and training rules.

MLP networks are organized in layers, linked by unidirectional connections as shown in Figure 3.

**Figure 3:**
Architecture of an MLP network with one hidden layer and an output layer.




Source: Bishop (2006)

The mathematical model of an MLP network, such as the one shown in Figure 3, may be represented by the following formula:

$$y_k(\boldsymbol{x},\boldsymbol{w}) = \sigma\left(\sum_{j=0}^{M} w_{kj}^{(2)}\, h\left(\sum_{i=0}^{D} w_{ji}^{(1)} x_i\right)\right) \quad (2)$$

In equation (2), the activation functions σ(*)and *h*(*) usually take the shape of the sigmoid function, since they represent a nonlinear mapping between input and output and may be differentiated, although other shapes are admitted, depending on the application.

The linear combination $\sum_{i=0}^{D} w_{ji}^{(1)} x_i$ , after undergoing nonlinear *h*(*) transformation, serves as input to the neurons of the hidden layer. This result is linearly combined with weights $w_{kj}^{(2)}$ and subject to further nonlinear σ(*) transformation, resulting in output $y_k$.

In summary, the MLP network is a nonlinear mapping from a set of input variables $\{x_i\}$ to a set of output variables $\{y_k\}$ controlled by a vector of adjustable parameters, known as synaptic weights.

The learning procedure (optimization) involves updating parameters $w_{kj}^{(2)}$ and $w_{ji}^{(1)}$, so that the neural network is able to efficiently perform the classification task. Regarding the MLP network, the most used and widespread optimization rule is the one defined by the error backpropagation algorithm.

During training with this algorithm, the network operates in a two-stage sequence. First, a pattern – a set of variables $\{x_i\}$ – is introduced to the network's input layer. The processing is performed through the network, layer by layer, until the output layer produces the response in an iterative process.

Thus, the initial step of the first stage comprises calculating the activation level and the outputs of all neurons in the hidden layer and output layer. The outputs of the hidden layer neurons are the input for the output layer neurons. Subsequently, the outputs of the output layer neurons are computed.

In the second stage, the obtained output is compared to the desired output for the presented pattern, since the desired output is known *a priori*. If the output obtained is not correct, the error (the difference between the desired and obtained outputs) is computed, and such error is propagated from the output layer to the input layer in the opposite direction. The weights of the hidden layer connections are modified as the error is backpropagated using the generalized delta rule, which is not explained in details herein since it is out of scope of this paper.

This step involves calculating the local gradients of the output layer neurons – $\delta_k(t)$– and hidden layer neurons– $\delta_i(t)$ – and adjusting the weights of all neurons.

The second step involves updating the synaptic weights of the MLP network. Thus, regarding the hidden layer, the rule to update weights $w_{ji}^{(1)}$ for the next iteration is:

$$w_{ji}^{(1)}(t+1) = w_{ji}^{(1)}(t) + \nabla w_{ji}^{(1)}(t) = w_{ji}^{(1)}(t) + \alpha\delta_i(t)x_i(t) \quad (3)$$

Regarding the output layer, the rule to update weights $w_{kj}^{(2)}$ is:

$$w_{kj}^{(2)}(t+1) = w_{kj}^{(2)}(t) + \nabla w_{kj}^{(2)}(t) = w_{kj}^{(2)}(t) + \alpha\delta_k(t)y_k(t) \quad (4)$$

In equations (3) and (4), "α" is one of the input parameters of the algorithm, and is known as the learning rate.

To summarize, when a pattern is initially introduced to the network, it produces an output. After measuring the distance between the current and desired response (error), the appropriate adjustments are made to the weights of the connections to reduce this distance. After the network is trained and the error achieves a satisfactory level, it may be used as a tool to classify new data.

The design of a neural network and the operation of this algorithm involve specifying a number of parameters, which decisively influence their performance, their convergence and the network generalization capability. These considerations do not fall within the scope of this paper.

The following example illustrates, in a very intuitive way, what problems MLP artificial neural networks, such as pattern classifiers, seek to solve. Consider a database containing information on a particular flower genus called Iris. This database consists of the following characteristics (features) of this flower genus: petal width, petal length, sepal width and sepal length. Depending on the values of these features, the flower is classified into one of three species (classes): *iris virginica, iris setosa* or *iris versicolor*. The database has various entries and each one of them associates the set of features of a certain flower to its respective species, as shown in the table below:

The MLP neural network is trained as part of these randomly selected entries (training set), i.e., the input data set (characteristics) and its respective flower class are introduced to the network, with the purpose of having the neural network learn from these data. After training, the neural network is able to classify, among one of the flower species, a new pattern (the set containing four characteristics of the flower), introduced to such network with a certain success rate.

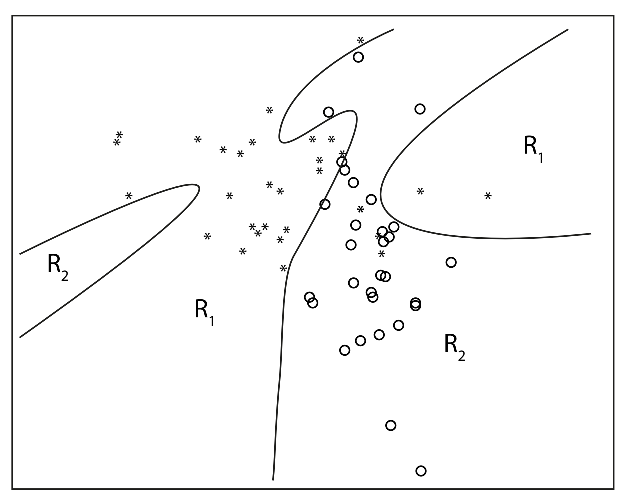
To have a graphical visualization of a MLP neural network classification solution, Figure 4 shows the optimal separation boundaries for two classes of patterns represented by the symbols "*" and "o".

## 5. APPLYING PATTERN CLASSIFICATION THROUGH ARTIFICIAL NEURAL NETWORKS IN THE AUDIT DOMAIN

Although it is not as frequent when compared to other branches of knowledge, the use of neural networks in the audit field is mentioned several times in the scientific literature, although the references cited herein are not the result of a comprehensive review.

Calderon and Cheh (2002) analyze papers published in some categories: risk assessment (three pa-

**Figure 4** – Decision boundaries ($R_1$ and $R_2$) of an MLP neural network to classify patterns from two classes




Source: Jain, Duin e Mao (2000).

**Table 1:** Example of only three entries from an input database used by a MLP neural network

| sepal | | petal | | Species |
|---|---|---|---|---|
| length (cm) | width (cm) | length (cm) | width (cm) | |
| 4,8 | 3,4 | 1,6 | 0,2 | *Iris setosa* |
| 6,2 | 2,2 | 4,5 | 1,5 | *Iris versicolor* |
| 6,9 | 3,1 | 5,4 | 2,1 | *Iris virginica* |

Source: prepared by the author

pers), fraud and error identification (six papers), issue of going concern[I] opinion (three papers), identification of situations where there is too much exposure to financial risks (three papers) and bankruptcy forecast (twelve papers). These authors state that neural networks may be superior to other techniques when data is available in large samples, the scale of values to be analyzed is large and associations among data are poorly defined and barely noticeable.

Garrity, O'Donnell and Sanders (2006), when defending continuous auditing and the use of computational intelligence, also highlight applications of artificial neural networks in the same areas mentioned by Calderon and Cheh (2002) and Koskivaara (2003).

Cerullo and Cerrullo (2006) analyze the use of neural networks to predict fraud in financial statements, by using coefficients and information on resulting from analyses of the accounting statements themselves. The authors state that neural networks process large amounts of data to solve problems by recognizing trends and complex relationships, which are barely noticeable to other computational methods.

Taha (2012) justifies the use of neural networks in auditing and concludes that such neural networks are better than statistical methods for planning and conducting audits. From his perspective, neural networks may indicate which financial statements are most likely to contain substantial errors, guiding the auditor in relation to how in-depth audit tests are and providing further conditions to issue a more accurate opinion on these financial statements.

Pourheydari, Nezamabadi-Pour and Aazami (2012) used four pattern classification techniques to identify modified[II] and unmodified[III] audit opinions on financial statements. Although their study shows other very interesting findings, its results showed that the MLP neural network proved to have high ability to identify different types of audit opinions on the financial statements, achieving a success rate of over 87%, jointly considering modified and unmodified opinions.

Finally, Byrnes *et al* (2014) defend the appropriation of data analysis techniques by auditing standards. From their perspective, technological developments, such as cloud computing, and the advances of data science contribute to enhance the effectiveness and efficiency of the audit work. They argue that the incorporation of computational intelligence enables continuous and predictive audits, more effective fraud detection and the safer issue of opinions.

These authors also point out several opportunities that are enhanced by the use of data analysis in financial audits, such as: the identification of risks associated with audit contracts (risks of bankruptcy and senior management fraud); the identification of risks of material errors and the performance of substantive tests and the identification of non-conformities in financial statements due to fraud.

These applications support, almost entirely, the works carried out by the control bodies. In addition, pattern recognition using artificial neural networks, may be used in many other audit problems. Some examples are the identification of fraud in bidding and procurement public processes and the granting of benefits from government programs, the identification of personnel admission or pension registration deeds not qualified for such, as well as a tool for continuous and predictive audit of the Government's Accounts, part of the annual review of the President's government accounts. Therefore, the use of neural networks may be an important tool to improve the effectiveness, efficiency and even the economy of the works carried out by these institutions.

## 6. FOSTERING THE USE OF COMPUTATIONAL INTELLIGENCE IN AUDITS

It has been well established that it is the duty of the Federal Court of Accounts (TCU) not only to control legality and compliance, but also to control efficiency, economy, efficacy and effectiveness of management actions in relation to individuals who use, collect, keep, manage or administer public monies, values and goods.

To perform an audit in all its dimensions, the TCU has been granted several duties by constitutional and infra-constitutional norms, which, over time, have become quite complex and varied, demanding a timely, focused and intelligent action, in order to optimize the resources made available to it.

Society, by becoming increasingly connected and aware of the need for greater transparency in the use of public resources, has increasingly demanded that government databases be made available. In this context, the TCU has often faced the need to properly handle this information and use it to assist its mission to improve the Government to benefit society.

In this scenario, the Presidency of the TCU, during the administration of Minister Aroldo Cedraz, decided to undertake actions to encourage the use of computational intelligence applied to audits.

Two strategic guidelines, which were outlined for TCU's 2015-2021 Strategic Planning and approved by TCU Order 141 from April 1st, 2015, stand out in this context: (i) using control intelligence to identify on a large scale the risks of nonperformance or inadequate implementation of products and services and inducing such practices to other parties subject to jurisdiction; and (ii) developing comprehensive organizational competence to work with emerging technology resources and analyzing large databases (Big Data).

At the tactical level, the 2015-2017 Audit Plan defines a line of action, which takes part in this movement, i.e., continuously monitoring, from the treatment of information databases, the use of public funds, in order to timely detect and correct possible diversions.

An outstanding initiative associated with this movement was the launch, on September 28th, 2015, of the TCU Center for Research and Innovation (CePI). This unit, which had already begun its activities in January 2015, aims to promote applied research in the TCU and coordinate the Innovation and Co-participation Lab (coLAB-i).

The coLAB-i aims to support innovative projects, ensure the knowledge management of developed solutions, coordinate cooperation activities and promote training activities and events in relation to topics at the frontier of knowledge. In addition, the coLAB-I, the first innovation laboratory in a control body, had the privilege of joining the select group included in Nesta's map of global laboratories (http://www.nesta.org.uk/ ) in its first year operating.

In addition, the TCU has promoted seminars and training courses on data analysis and given awards, through the *Reconhe-Ser* program, to several projects applying data science tools.

## 7. CONCLUSION

The rapidly growing capacity of computer processing and the availability of large databases made it easier to use more sophisticated data analysis and classification methods. In this context, pattern recognition techniques, such as artificial neural networks, gained prominence in applications in various branches of knowledge.

Artificial neural networks are number processing systems formed by highly connected processing units and able to map nonlinear relationships present in large databases. Using the nonlinear discriminant analysis theory and the error backpropagation algorithm for training and optimizing their parameters, the multilayer perceptron neural networks are able to generalize the acquired knowledge and classify patterns with high success rates.

Classification patterns with the use of neural networks may help control bodies at all audit stages, whether in the planning, execution or report stage. Several studies have presented applications for risk assessment, fraud and error identification, continuing assessment (going concern opinion), identification of situations where there is too much exposure to financial risk, bankruptcy forecast, and identification of modified and unmodified audit opinions on financial statements, among others. This is a vast and yet little explored scope of application.

The audit activity, in particular, understands information and knowledge as input and product, and such elements increasingly depend on information technology. Data processing to extract information is, therefore, indispensable to leverage control activities in an increasingly connected society.

The TCU, aware of this reality and the need to conduct increasingly focused, timely and intelligent actions, has encouraged the use of data science applied to audits, through various actions contained in its Strategic Plan and Audit Plan. We highlight the launch, on September 28th, 2015, of the Center for Research and Innovation (CePI).

## NOTES

I Also known as continuity evaluation, it consists in analyzing short and medium term cash flow of an operation, business or company. It sustains short-term strategic decisions based on cash generation and business liquidity, revealing the financial strength of a company for the following months and deciding on fundings, refundings, investments and other strategic, operational and financial elements.

II An opinion with reservations, adverse opinion and abstention of opinion on financial statements, necessary when: (a) the auditor concludes, based on the obtained audit evidence, that the financial statements as a whole present relevant misstatements; or (b) the auditor is able to obtain appropriate and sufficient evidence to conclude that the financial statements as a whole do not present relevant misstatements.

III The opinion given by the auditor when he concludes that the financial statements, in all relevant respects, were prepared in accordance with the applicable financial statement structure.

## REFERENCES

BISHOP, Christopher M. Pattern Recognition and Machine Learning. Springer, October 2006.Availableon<http://www.rmki.kfki.hu/~banmi/elte/Bishop%20-%20Pattern%20Recognition%20and%20Machine%20Learning.pdf>. Accessed on: Aug 10. 2015.

BYRNES, Paul., CRISTE, Tom., TREVOR, Stewart., VASARHELYI, Miklos. Reimagining Auditing in a Wired World. AICPA White Paper, August2014.Available on: <https://www.aicpa.org/InterestAreas/FRC/AssuranceAdvisoryServices/DownloadableDocuments/Whitepaper_Blue_Sky_Scenario-Pinkbook.pdf >. Accessed on:Nov 11. 2015.

CALDERON, Thomas G., CHEH, John J. A roadmap for future neural networks research in auditing and risk assessment. International Journal of Accounting Information Systems, Volume 3, Issue 4, December 2002, Pages 203-236. Available on: <http://www.sciencedirect.com/science/article/pii/S1467089502000684>. Accessed on: Jan 18. 2016.

CERULLO, Michael J., CERULLO, Virginia. Using Neural Network Software as a Forensic Accounting Tool. Information Systems Control Journal, ISACA, 2006. Available on: <https://www.google.com.br/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&uact=8&ved=0ahUKEwiGn6W76OvKAhULGpAKHV6UCGQQFggdMAA&url=http%3A%2F%2Fwww.isaca.org%2FJournal%2Farchives%2F2006%2FVolume-2%2FDocuments%2Fjpdf0601-using-neural-network.pdf&usg=AFQjCNGLi7Omn2qs3hPy4lqcuxU_AYAmoA>. Accessed on: Dec 15. 2015.

GARRITY, Edward J., O'DONNELL, Joseph B., SANDERS, Lawrence. Continuous auditing and data mining. Idea Group Inc., 2006. Available on: <http://www.igi-global.com/chapter/continuous-auditing-data-mining/10596?camid=4v1>. Accessed on: Jul 7. 2015.

JAIN, Anil K., DUIN, Robert P. W., MAO, Jianchang. Statistical Pattern Recognition: A Review. IEEE Transactions on Pattern Anaçysis and Machine Intelligence, Vol. 22, nº 1, January 2000. Available on: <http://www.cs.indiana.edu/~predrag/files/jain_2000.pdf >. Accessed on: Jul 22. 2015.

KOSKIVAARA, Eija. Artificial Neural Networks in Auditing: State of the Art. TUCS TechnicalReport 509, 2003. Available on: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.67.459&rep=rep1&type=pdf >. Accessed on: Oct 10. 2015.

POURHEYDARI, Omid., NEZAMABADI-POUR, Hossein., AAZAMI, Zeinab. Identifying qualified audit opinions by artificial neural networks. African Journal of Business Management, Vol 6, p. 11077-11087, 7 November 2012. Available on: <http://www.academicjournals.org/app/webroot/article/article1380541005_Pourheydari%20et%20al.pdf >. Accessed on: Feb 2. 2016.

TAHA, Reem O. Hussein. The possibility of using Artificial Neural Networks in Auditing - Theoretical Analytical Paper. European Journal of Economics, Finance and Administrative Sciences, Issue 47, 2012. Available on: <http://www.academia.edu/1634776/The_Possibility_of_using_Artificial_Neural_Networks_in_Auditing_-Theoretical_Analytical_Paper >. Accessed on: Jan 30. 2016.