

[https://doi.org/10.52326/jss.utm.2021.4\(3\).14](https://doi.org/10.52326/jss.utm.2021.4(3).14)
UDC 004.8:657



THE USE OF MACHINE LEARNING TECHNIQUES IN ACCOUNTING. A SHORT SURVEY

Mariana Vlad, ORCID ID: 0000-0003-1163-5637,
Sorin Vlad*, ORCID ID: 0000-0001-9422-976X

“Ștefan cel Mare” University, 13 University Str., 720229, Suceava, Romania

*Corresponding author: Sorin Vlad, sorinv@usm.ro

Received: 05. 28. 2021

Accepted: 07. 12. 2021

Abstract. Machine learning (ML) is a subset of artificial Intelligence (AI) aiming to develop systems that can learn and continuously improve the abilities through generalization in an autonomous manner. ML is presently all around us, almost every facet of our digital and real life is embedding some ML related content. Customer recommendation systems, customer behavior prediction, fraud detection, speech recognition, image recognition, black & white movies colorization, accounting fraud detection are just some examples of the vast range of applications in which ML is involved. The techniques that this paper investigates are mainly focused on the use of neural networks in accounting and finance research fields. An artificial neural network is modelling the brain ability of learning intricate patterns from the information presented at its inputs using elementary interconnected units, named neurons, grouped in layers and trained by means of a learning algorithm. The performance of the network depends on many factors like the number of layers, the number of each neurons in each layer, the learning algorithm, activation functions, to name just a few of them. Machine learning algorithms have already started to replace humans in jobs that require document's processing and decision making.

Keywords: *artificial intelligence, accounting, connectionist models, expert systems, neural networks.*

Rezumat. Machine learning (ML) este un subdomeniu al Inteligenței Artificiale (AI) care are ca scop realizarea sistemelor care pot învăța și își pot îmbunătăți în mod continuu și autonom abilitățile prin generalizare. ML este prezentă peste tot în mediul înconjurător, aproape orice aspect al vieții noastre digitale și reale, de zi cu zi, include părți care pot fi asimilate de ML. Sistemele de recomandare a produselor, predicția comportamentului clienților, detectarea fraudelor, recunoașterea imaginilor, recunoașterea vorbirii, colorarea automată a filmelor alb-negru, detectarea fraudelor în contabilitate, sunt doar câteva exemple din vastul domeniul al aplicațiilor care încorporează tehnici specifice domeniului ML. Tehnicile pe care le investighează acest articol se concentrează în principal pe utilizarea rețelelor neuronale în domeniile contabilitate și finanțe. O rețea neuronală artificială modelează capacitatea creierului de a învăța tipare complicate pe baza informației prezentate la intrările acesteia,

utilizând unități elementare interconectate, numite neuroni, grupați în straturi și antrenați aplicând un algoritm de antrenare. Performanța rețelei depinde de numeroși factori: numărul de straturi, numărul de neuroni din fiecare strat, algoritmul de învățare, funcțiile de activare fiind doar unii dintre aceștia. Algoritmii ML au început deja să înlocuiască operatorii umani în activități care necesită prelucrarea documentelor și luarea deciziilor.

Cuvinte-cheie: *inteligență artificială, contabilitate, modele conexioniste, sisteme expert, rețele neuronale.*

Introduction

The amount of information available today in any domain is huge and continuously expanding. Processing such immense amount of data would be impossible without the use of computers and algorithms designed to detect and understand relevant patterns. The knowledge extracted is usually used to undertake data classification or to predict the future evolution of some indicators. Classification problems may consist for instance in detection of fraud by considering a large number of documents. The examination of those documents will not be possible without using a significant part of the specialized personnel. AI based software will reduce the time needed to analyze the documents and will detect only the potentially fraudulent documents that will have to be further verified.

The importance of Information Technology was emphasized by some pioneering papers [1, 2]. First applications of Artificial Intelligence (AI) methods and techniques in accounting are dating back in 80s of the past century and were used for auditing and financial reporting. Since then, the process of adopting AI as a result of researches has grown and spread in multiple directions leading to the adoption, in 2016, by one of the largest accounting firms in the world, Deloitte Touche Tohmatsu, of AI in taxation and auditing.

ML is only a part of AI that uses learning algorithms to absorb and process data from the environment in order to create systems that can mimic the human intelligence and inference ability. There are three paradigms of learning: supervised, unsupervised and reinforced learning. For each one of those three categories of algorithms there are corresponding technologies. For instance, connectionist methods like neural networks are trained using all types of learning algorithms: Kohonen or SOM neural networks uses unsupervised learning, multilayer perceptron uses backpropagation, a supervised learning algorithm, etc.

Artificial neural networks

An artificial neural network (ANN) can be broadly defined as being an algorithm designed to model the way in which the human brain accomplishes a certain task. A more specific definition would be the following: an ANN is a parallel processor being able to store the knowledge derived from experiments in order to subsequently use it.

All neuronal models are based on simple interconnected units called artificial neurons which are placed in layers (the input, hidden and output layer). Each connection is assigned a numerical value called weight. The magnitude of the weights is, in fact the strength of the connection. Each neuron will fire a response based on its activation function and the values at the inputs.

A single neuron can do simple calculations, but the whole network has outstanding performances when solving complex problems. The ability of the ANN to learn by examples,

to detect patterns and complex relationships when working with huge amount of data transforms them in powerful tools in difficult problem solving.

ANNs are converting values places on the network inputs into actions. The input dataset (the input pattern dataset) is converted at the output using a function f according to a rule $y=f(x)$. The basic principle of ANNs is to improve this rule by adjusting the weights using a learning algorithm.

Hidden neurons give the accuracy degree of the learning process. If there are too many neurons in the hidden layer the network will memorize the problem and the generalization will be poor, if the hidden neurons are too few, the generalization is good, but the network will not be able to learn correctly all the data in the training set. A MLP with one hidden layer has it's an universal approximator, i.e. has the ability to model anu kind of complex problem. However, if the degree of complexity is very high, the ANN may have more than one hidden layer of neurons.

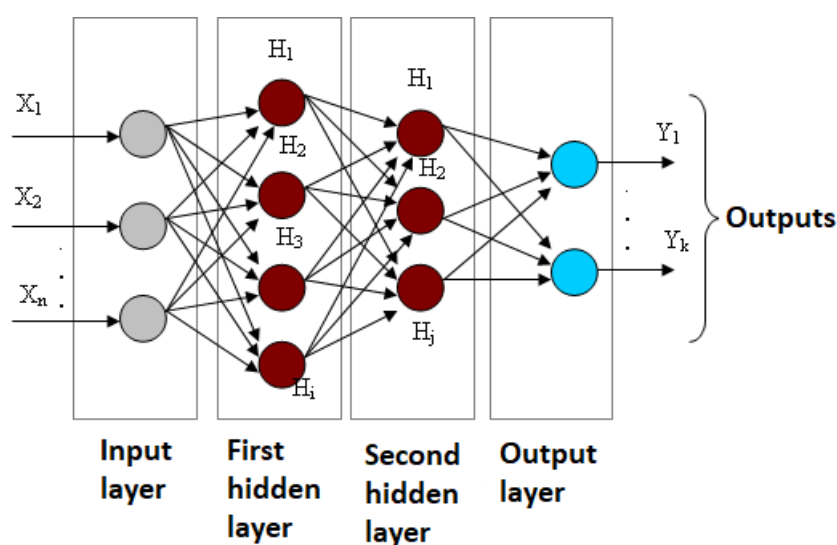


Figure 1. A multilayer perceptron (MLP) with two hidden layers.

The most important features of the ANNs are the same as those of the biological networks and can be summarized as follows:

- Learning ability – ANNs are trained with a data set and the learning algorithm (usually error backpropagation) will try to minimize the difference between the values that should be obtained at the network outputs and the values that are actually obtained. The network will be finally able to “learn” the training set and to construct by itself the algorithm for solving the problem. The learning algorithm will be increasingly successful if the training data is representative for the problem to be solved.
- Generalizing ability – measures the quality of learning and depends on multiple factors (the network structure, the adequacy of the training set, etc.)
- Fault tolerance – the knowledge is distributed through the entire network; thus, some faulty neurons will not cause a significant decrease of performance:
- Adaptability – the weights permanently modified in order to adapt to the environment:
- Flexibility – ANNs can model a broad range of statistical methods simply by modifying the activation functions and / or the network structure (i.e. the number of neurons on each layer).
- Nonlinear relationship modeling – the majority of the natural phenomena are nonlinear, linear models will not perform well.

ANNs are used in two large classes of problems: classification and prediction. There are advantages and drawbacks of using them. The main drawback is that you don't have a

ready-to-use network architecture for a specific problem. The process of obtaining a viable network to solve the problem is time consuming and consists in fine tuning of multiple parameters. The main advantages are that, once successfully trained, the network will be able to model very accurately any complex non-linear system processing a very large amount of (redundant and even noisy) data.

ANN in accounting

The papers researching the use of ANN in accounting and finance are centered on bankruptcy prediction, market segmentation, auditing and credit evaluation. For bankruptcy prediction, the preferred network architecture used to create neuronal model for prediction is generally the multilayer perceptron. However, some authors tried to apply other types of ANNs [3] or combining several techniques together with ANN into hybrid models [4]. The performance of ANN is compared with nonparametric multiple discriminant analysis, multistate logit [5, 7], logistic regression models [7] or with previous used models [8]. The results obtained using ANNs outcomes those obtained using other techniques.

Market segmentation is the process that allows dividing a large heterogenous market in many smaller homogenous components. Smaller homogenous markets are easier to target with effective marketing campaigns and thus, market segmentation can be understood as an alternate method of customer relationship management [9]. Neural networks, especially Kohonen neural networks or self-organizing map (SOM) were used to study different markets (industrial, touristic, telecommunication, etc.) [10 - 12]. A number of 1000 papers indexed in major databases, written between 2000 and 2010 were identified and 100 were examined in [13]. The papers were classified according to the type of ANN used for segmentation.

Auditing

The first AI tool used in audit were the expert systems. The experience of the expert is encapsulated in a so-called knowledge base and used in intelligent decision making based on inference rules managed by an inference engine. Expert systems are well suited for auditing, because the auditor has many solutions to a specific problem but he must choose the best one. The "choosing" part is the duty of expert system.

According to [14] there are 400 audit tasks that can be identified. These tasks can be grouped in the following categories: analytical review procedures, risk assessment, classification, internal control evaluation, materiality assessment and going-concern decision.

An exact definition of materiality assessment cannot be given. Materiality refers to a threshold over which the misstatements become financially relevant for the users of the financial statements.

Papers investigating expert systems application in materiality assessment are rather sparse. A new model using fuzzy logic for constructing rules is described and discussed in [15].

Risk assessment involves pattern deviation detection in order to detect some irregularities i.e., fraud. Fraud detection studies are centered on data mining techniques and neural networks models trained over large training data sets to detect suspect patterns. trained A radial basis neural network with one neuron on the output layer is trained and the output value represents a fraud score [16]. The transaction below the score is considered fraudulent. Other approaches considered different models: a model using SOM neural

networks and user accounts is used in [17]. User activities are stores in matrices that are classified using a threshold classification algorithm.

The ability of neural networks to detect patterns in monthly balances and to recognize the dynamics between different accounts is tested in [18] and the results are found encouraging. The neural networks used seemed to detect and learn relationships among the accounts used in 72 monthly balances of a manufacturing firm.

Conclusions

AI tools have the potential to ease the efforts in any domain. In accounting, machine learning researches have not reached at their full potential yet. Interdisciplinary efforts are still to be done in order to adapt AI methods to the growing amount of accounting data.

References

1. Coopers Lybrand. *The chartered accountant in the IT age*, London, Coopers & Lybrand and ICAEW, 1985
2. Barras R., Swann J. *The adoption and impact of IT in the UK accountancy profession*, London: The Technology Change Centre, 1984
3. Yang R. Z., Platt M., Platt H. *Probabilistic Neural Networks in Bankruptcy Prediction*, Journal of Business Research, vol. 44, Issue 2, 1999
4. Azayite F.Z., Achchab S. *Hybrid Discriminant Neural Networks for Bankruptcy Prediction and Risk Scoring*, Procedia Computer Science, vol. 83, 2016
5. Barniv R., Agarwal A. and Leach R. *Predicting the outcome following bankruptcy filing: a three-state classification using neural networks*, International Journal of Intelligent Systems in Accounting, Finance and Management, 6, No. 3, 1997
6. Wilson R. L., Sharda R. *Bankruptcy Prediction*, Decision support Systems, Volume 11, Issue 5, North Holland, 1994
7. Zhang G., Hu M. Y. *Artificial neural networks in bankruptcy prediction: General framework and cross-validation analysis*, European Journal of Operational Research, vol. 116, Issue 1, 1999
8. Atyia A., *Bankruptcy prediction for credit risk using neural networks: A survey and new results*, IEEE transactions on Neural Networks, vol. 2, Issue 4, 2001
9. Hung F., Tsai F.-C. *Market segmentation based on hierarchical self-organizing map for markets of multimedia on demand*, Expert Systems with Applications 34, 2008.
10. Kiang M. Y., Hu M. Y., Fisher M.D. *An extended self-organizing map network for market segmentation—a telecommunication example*, Decision Support, vol. 42, Issue 1, 2006
11. Fish K. E., BARNES J. H., *Artificial neural networks: A new methodology for industrial market segmentation*, Industrial Marketing Management, vol. 24, Issue 5, 1995
12. Bloom J. Z., *Market segmentation: A Neural Network Application*, Volume 32, Issue 1, 2005
13. Chattopadhyay M., Dan P.K., Majumdar S. *Application of neural networks in market segmentation: a review of recent trends*, Management Science Letters, 2(2), 2012
14. Baldwin A., Brown C., Trinkle B. *Opportunities for artificial intelligence in the accounting domain: the case for auditing*, 2006.
15. Comunale B.L., Sexton T. *A fuzzy logic approach to assessing materiality*, Journal of Emerging Technologies in Accounting 2 (1), 2005.
16. Ghos S., Reilly D. *Credit Card fraud detection using a neural network*, Proceedings of the Twenty-Seventh Hawaii International Conference on System Sciences, 1994
17. Olsewski D. *Fraud detection using self-organizing map visualizing the user profiles*, Knowledge Based Systems, vol. 70, 2014
18. Koshivara E. *Artificial neural models for predicting patterns in auditing monthly balances*, The Journal of Operational Research Society, Vol. 51, No. 9, 2000