**Data mining methods – Artificial Neural Networks: A Literature Review from 1991 - 2017**

**CA1**

**Magda L.**

Table of Contents

[Abstract 2](#_Toc507507082)

[Introduction 3](#_Toc507507083)

[1. Definition of ANN 3](#_Toc507507084)

[1. Neural Networks in Data mining 4](#_Toc507507085)

[2. Learning in ANN 6](#_Toc507507086)

[2.1 Supervised learning 6](#_Toc507507087)

[2.2 Unsupervised – reinforcement learning 7](#_Toc507507088)

[2.3 Semi – Supervised Learning 7](#_Toc507507089)

[2.4 Back Propagation algorithm 7](#_Toc507507090)

[3. Network Architecture 8](#_Toc507507091)

[3.1 Single Layer Feedforward Networks 8](#_Toc507507092)

[3.2 Multilayer Feedforward Networks 9](#_Toc507507093)

[3.3 Recurrent Networks 10](#_Toc507507094)

[1. ANN contra traditional PC 11](#_Toc507507095)

[Conclusion 12](#_Toc507507096)

[Bibliography 12](#_Toc507507097)

# Abstract

Artificial Neural Network is quite young definition and literature available is limited but there are many very good positions from ‘90. We manage to gather literature from 1991 to 2017. Although many articles start to show up from 1986. That was a renaissance for ANN.

Keywords: Data mining; neural networks, ANN, Artificial Neural Network

# Introduction

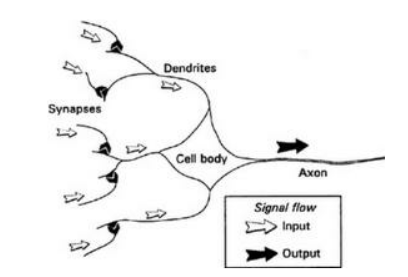
Nowadays developments of tools which can check big data, scan, analyse and interpret started develop very quickly. Also the amount of big data increase every day and stored are calculated now in exabytes. This work is about one of tool called Neural Networks or sometimes called Artificial Neural Networks.

Artificial Neural Networks has been widely used in many aspects of life. We will be covering few aspects of ANN:

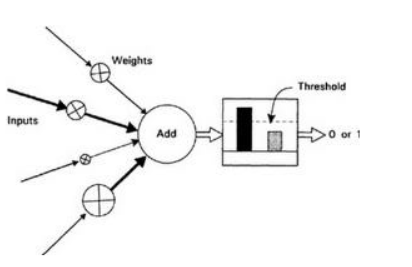
* Definition of ANN
* ANN in Data mining – techniques and application
* Learning processes – supervised versus unsupervised
* Network Architecture – common types
* Critical evaluation of Neural Networks - ANN Versus traditional PC

# Definition of ANN

The most common definition of Artificial Neural Network (ANN) is based on human and animal brain and nervous system. The programmers were inspired by nervous system and how it works. Brain has large number of neurons working together. The Neural Networks it is not simple model that you can create and it’s ready to use. You have to train you Network first to be able to do your tasks, just like an animal.



Picture1. Kevin Gurney,’ Essential components of a neuron shown in stylized form’ 1997, p.13, ‘An introduction to neural networks’[[1]](#footnote-1).



Picture2. Kevin Gurney,’ Simple artificial neuron’ 1997, p.14, ‘An introduction to neural networks’[[2]](#footnote-2).

On picture 1 we can see elements of a natural neuron from human or animal brain. On picture 2 in other hand we can see artificial neuron. We can see that the artificial was inspired on natural one.

In the other side we can read in very popular work of David Kriesel published online "’The neural network’ does not exist. There are different paradigms for neural networks, how they are trained and where they are used.”[[3]](#footnote-3) It is interesting point of view but of course we call it neural networks because they have been inspired by natural neural networks not because they are on of them.

# Neural Networks in Data mining

In ‘Introduction to Neural Networks’ by Kevin Gurney from University of Sheffield from 1997 we can ready that “Neural networks are often used for statistical analysis and data modelling, in which their role is perceived as an alternative to standard nonlinear regression or cluster analysis techniques”[[4]](#footnote-4). Of course ANN cannot be used for everything due to their limitations (You can ready about that in Chapter 5). But there are certain activities which they can be very useful for example face and handwriting recognition.

Basically they widely used for:

* Classification,
* Prediction,
* Pattern recognition,

Many ANN algorithms has been widely used and developed for Data mining purposes. For example - backpropagation algorithm. “The backpropagation algorithm performs learning on a multilayer feed-forward neural network. It iteratively learns a set of weights for prediction of the class label of tuples.”[[5]](#footnote-5)

Backpropagation algorithm is widely used for patter recognition problems. Chen C.H. in her ‘Neural network in pattern Recognition and their applications’ from 1991 mentioning that already 27 year ago ANN had a great impact on pattern recognition and artificial intelligence. She kindly describe many journal articles about that issue.

Data mining processes we can divide between 3 groups: preparation, core data mining, interpretation of results. In all that processes we can use neural networks.

During the Literature Review I discover very interesting position. *Computational Finance ’99* edited by Yaser S., Abu-Mostafa and others. It is a set of papers from yearly conference called Computational Finance (old ‘Neural Network in the Capital Markets’). It is very technical papers, many of them contains raw algorithms and examples. But we can see that ANN in Data mining techniques is used in Capital Market Analysis and they are very popular as analysis methods and forecasting techniques or bankruptcy predictions.

# Learning in ANN

Neural Networks need to be trained before they will be able to make correct actions and decisions. Learning looks that we show the neural network situations and network learn on its own mistakes, like a baby. That situations are called training samples. All training samples are set up before activity and based on strategy.

According to Mr. Kreisel ‘The learning strategy is an algorithm that can be used to change and thereby train the neural network, so that the network produces a desired output for a given input.’[[6]](#footnote-6)

S. Haykin connects all learning algorithms with network architecture (structures). We will talk about that in next chapter.

### Supervised learning

Supervised learning goal is to use mappings to learn how to get desired outcome. For example we create function when adding an inputs we supervise our machine and we know what is she learning based on our inputs. We know the correct answer before and train until we happy with machine result are this same as ours. It is used very often with classification when output is a category or problems with regression when a output is a variable with real value like for example currency.

Popular algorithms for supervised learning:

* Linear regression
* Random forest
* Support vector machines

### Unsupervised – reinforcement learning

Unsupervised learning is much more complicated. It is basically just your raw data as input and no correct output supervised by a teacher. Because there is no teacher at all. It is very often used to see the data distribution or structure. It is used widely in data mining processes. Used mainly for clustering analysis and for discovering association rules.

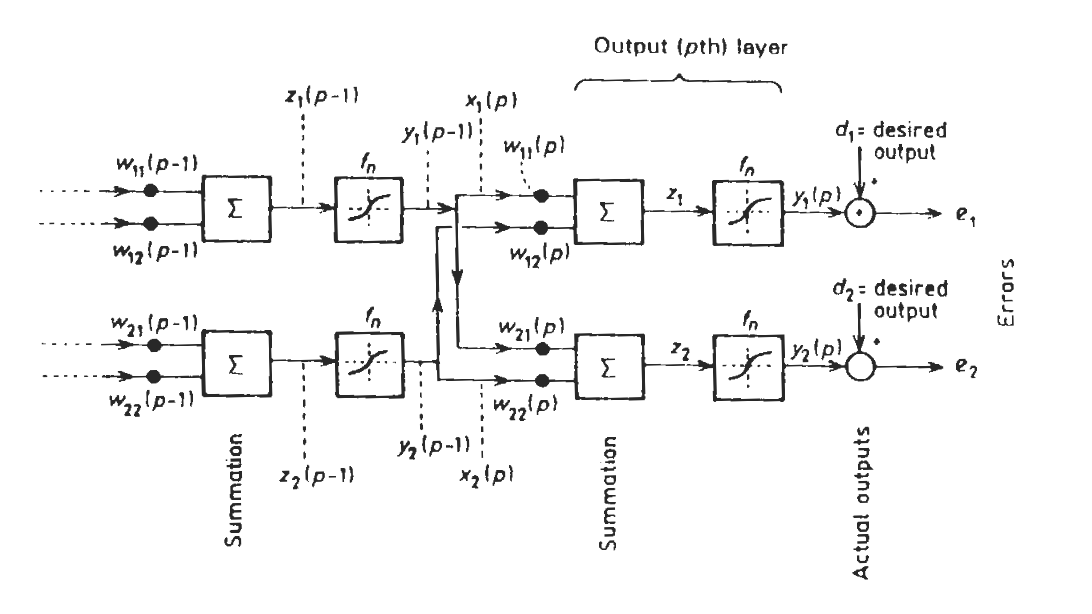
Popular algorithms:

* K- mean
* Apriori algorithm

### Semi – Supervised Learning

It is a mix of both techniques. We have a teacher with some correct answers for some inputs and for others we just put raw data. It is the most common technique. Used in many cases, like face detection, photo recognition, when many data is not labelled. Always unsupervised technique will be cheaper so it is the best of both. Used in data mining for recognition or prediction problems.

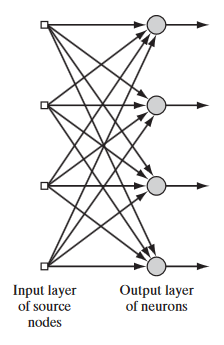
### Back Propagation algorithm

It is popular algorithm used for Multilayer Feedforward Networks (See Chapter 3.3). “The back propagation (BP) algorithm was proposed in 1986 by Rumelhart, Hinton and Williams for setting weights and hence for the training of multi-layer perceptrons. This opened the way for using multi-layer ANN.”[[7]](#footnote-7)  
Picture 3. Graupe, *A multi-layer perceptron* D 2007, *Principles of artificial neural networks*, Second Edition, University of Illinois, Chicago, USA, p. 60,

# Network Architecture

We can find 3 different network Architectures:

### Single Layer Feedforward Networks

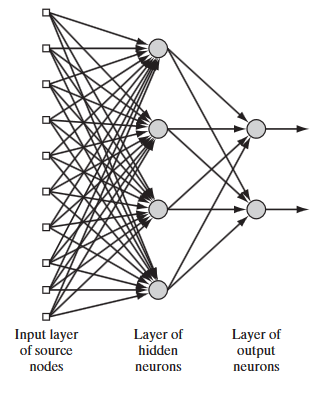


Picture 3. Haykin, S *Single Layer of neurons Feedforward Networks* 1999, Neural Networks and Learning Machines, Third Edition, McMaster University, Hamilton, Ontario, Canada, p.52

As you can see on Picture 3, Mr Haykin explains that this particular network is based on single layer of neurons. It is the simples form. Input layer are faced to output layer and there is no feedback, that’s why it’s called feedforward.

Other important aspect and limitation is lack of connection between neurons in this architecture. That limit the activities which we can use this particular scheme.

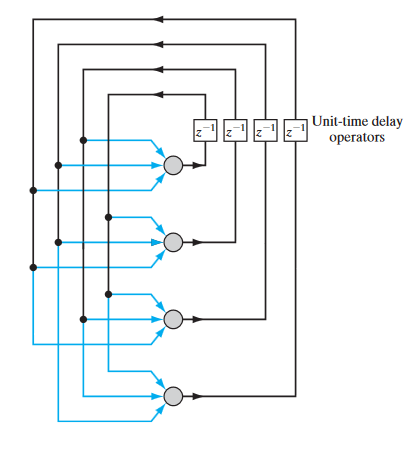
### Multilayer Feedforward Networks



Picture 4. Haykin, S *Multi-Layer of neurons Feedforward Networks* 1999, *Neural Networks and Learning Machines, Third Edition*, McMaster University, Hamilton, Ontario, Canada, p.53

This type of network has more layers of neurons. It is not visible for input output, but has function between. It is the Network when we can use our popular backpropagation algorithm.

### Recurrent Networks



Picture 5. Haykin, *Recurrent Network* 1999, *Neural Networks and Learning Machines, Third Edition*, McMaster University, Hamilton, Ontario, Canada, p.53

It is not feedforward type, so it can give feedback. As we can see on picture 5 it contains loops.

“The beauty of recurrent neural networks lies in their diversity of application. When we are dealing with RNNs they have a great ability to deal with various input and output types.”[[8]](#footnote-8)

Few examples:

* Classification
* Image Captioning
* Translation of language

Basically outputs and inputs can be with different type of files or sentences, pictures, lengths and size.

# ANN contra traditional PC

There are many benefits of neural networks. Very good summary of them we can find in S. Haykin ‘Neural Networks and Learning Machines’. The most important are that ‘neural network derives its computing power through, first, its massively parallel distributed structure and, second, its ability to learn and therefore generalize’[[9]](#footnote-9). Basically, its mean that ANN is able to react on new data which never was train before through generalisation of examples – conditioning.

Another benefit mentioning is adaptivity. Adaption to any changes and deal with problems. We can show the system how big level of adaptivity we want.

High parallelity of neural networks helps to work even if one neural network fails to do its work, Traditional PC is a one structured machine and if one feature is not working whole machine is not working properly.

ANN also handles noisy data very well and we can expect high accuracy and that is very big advantage over other data mining methods.

They are well suited for continuous-valued inputs and outputs, unlike most decision tree algorithms[[10]](#footnote-10).

Drawback is that neural networks have to work together, consistent and integrated system is a basic. The activity can be very simple but the ANN anyway needs to be trained. Learning sometimes can be long and demotivated. ANN are amazing tools but not for everything.

Another drawback is they poor ability to interpret data. Symbolic meaning of data, something what is not obvious for humans too.

Also other concerns are behind the lack of knowledge what is hidden in ‘black box’. That’s mean behind all that learning algorithms and training. It is not clearly visible and we not able to predict outcomes.

# Conclusion

Artificial Neural Networks can be widely used in many aspects and studies and with many disciplines. For example they start to be very popular in financial analysis to predict bankruptcy, forecasting, hedge funding and other trading, price policy in the company, in biology, in weather forecasting, medicine, and mostly in Computer Science for Security monitoring, gaming etc.

We can see unbelievable and infinite number of practical use of neural networks. Some are still discovered. We can and we should use it as it can develop amazing tools which can unsupervised make decisions after training. It is very exciting tool as it’s kind of copy of human brain.

# Bibliography

1. Gurney, K 1997, *An introduction to neural networks*, University of Sheffield, London and New York,
2. Kriesel, D 2005, *A brief Introduction to Neural Networks*, University of Bonn in Germany, dkriesel.com, [Accessed 21/02/2018],
3. Haykin, S 1999, *Neural Networks and Learning Machines*, Third Edition, McMaster University, Hamilton, Ontario, Canada,
4. Han et al. (), *Data mining Concept and Techniques, Third Edition,* University of Illinois at Urbana–Champaign,
5. Witten and Frank 2005, *Data Mining Practical Machine Learning Tools and Techniques,* Second Edition, Department of Computer Science, University of Waikato,
6. Graupe, D 2007, *Principles of artificial neural networks,* Second Edition,University of Illinois, Chicago, USA,
7. Chen, C.H 1991, Neural Networks in Pattern Recognition and Their Applications, Electrical and Computer Engineering Department, University of Massachusetts Dartmouth, USA
8. Brownlee, J. 2016. *Supervised and Unsupervised Machine Learning Algorithms.* [online] Available at: http:// machinelearningmastery.com/supervised-and-unsupervised-machine-learning-algorithms/ [20/02/2018].
9. Gupta, D. 2017. *Fundamentals of Deep Learning – Introduction to Recurrent Neural Networks.* [online] Available at: https://www.analyticsvidhya.com/blog/2017/12/introduction-to-recurrent-neural-networks/ [20/02/2018].
10. Abu-Mostafa, Yaser, S. 2000. *Computational Finance 1999*, Massachusetts Institute of Technology

1. Gurney, K 1997, *An introduction to neural networks*, University of Sheffield, London and New York, p.13. [↑](#footnote-ref-1)
2. Gurney, K 1997, *An introduction to neural networks*, University of Sheffield, London and New York, p.14. [↑](#footnote-ref-2)
3. Kriesel, D 2005 *A brief Introduction to Neural Networks,* University of Bonn in Germany, p.25, dkriesel.com [Accessed 21/02/2018] [↑](#footnote-ref-3)
4. Gurney, K 1997, *An introduction to neural networks’,* University of Sheffield, London and New York, p.19. [↑](#footnote-ref-4)
5. Han et al. (), *Data mining Concept and Techniques, Third Edition,* University of Illinois at Urbana–Champaign, Morgan Kaufmann, p.398 [↑](#footnote-ref-5)
6. Kriesel, D 2005 *A brief Introduction to Neural Networks,* University of Bonn in Germany, p.65, dkriesel.com, [Accessed 21/02/2018] [↑](#footnote-ref-6)
7. Graupe, D 2007, Principles of artificial neural networks, Second Edition, University of Illinois, Chicago, USA, p. 59, [↑](#footnote-ref-7)
8. Gupta, D. 2017. *Fundamentals of Deep Learning – Introduction to Recurrent Neural Networks.* [online] Available at: https://www.analyticsvidhya.com/blog/2017/12/introduction-to-recurrent-neural-networks/ [20/02/2018]. [↑](#footnote-ref-8)
9. Haykin, S 1999, *Neural Networks and Learning Machines, Third Edition*, McMaster University, Hamilton, Ontario, Canada, p. 33 [↑](#footnote-ref-9)
10. Han et al. (), *Data mining Concept and Techniques, Third Edition,* University of Illinois at Urbana–Champaign, Morgan Kaufmann, p.398 [↑](#footnote-ref-10)