https://doi.org/10.17221/45/2021-VETMED

Progressive trends on the application of artificial neural networks in animal sciences – A review

*Edyta Agnieszka Bauer**

Department of Animal Reproduction, Anatomy and Genomics, University of Agriculture in Krakow, Krakow, Poland

**Corresponding author: e.bauer@urk.edu.pl*

Citation: Bauer EA (2022): Progressive trends on the application of artificial neural networks in animal sciences – A review. Vet Med-Czech 67, 219–230.

Abstract: In recent years, artificial neural networks have become the subject of intensive research in a number of scientific areas. The high performance and operational speed of neural models open up a wide spectrum of applications in various areas of life sciences. Objectives pursued by many scientists, who use neural modelling in their research, focus – among others – on intensifying real-time calculations. This study shows the possibility of using Multilayer-Perceptron (MLP) and Radial Basis Function (RBF) models of artificial neural networks for the future development of new methods for animal science. The process should be explained explicitly to make the MLP and RBF models more readily accepted by more researchers. This study describes and recommends certain models as well as uses forecasting methods, which are represented by the chosen neural network topologies, in particular MLP and RBF models for more successful operations in the field of animals sciences.

Keywords: livestock animal science; machine learning; machine models

Introduction

A neural network-based modelling method is capable of mapping extremely complex functional relationships. One of the most valuable qualities of neural networks is their ability to process information using a parallel method that is completely different from serial operations of a traditional computer, and the capacity to learn based on examples and to replace traditional software. Neural networks show very high accuracy in predicting various occurrences and processes when it is difficult to clearly define a cause-and-effect relationship or there are no rules for establishing a logical connection between causes and effects. In such cases, artificial neural networks (ANNs) have the capacity to learn and are able to predict the probability of an occurrence in the future based on examples from the past (Hecht-Nielson 1989; Tadeusiewicz 1993; Bishop 1995; Rutkowski 2004; Han et al. 2012; Biecek and Burzykowski 2021).

LITERATURE REVIEW OF ANNs

ANNs were created in the 1940s, when McCullogh and Pitts (McCullogh and Pitts 1943; Jain et al. 1996) described the structure of the basic element of ANNs, i.e., a neuron, based on the anatomy of a nerve cell. The evolution of artificial neural systems is said to have begun in 1943. Nevertheless, it is assumed that the very domain of artificial neural networks did not appear until the release of the historic publication of Anderson and Rosenfeld

Financed by the National Science Center in Poland (Project No. 2017/25/N/NZ5/00793).

https://doi.org/10.17221/45/2021-VETMED

in 1988 (Anderson and Rosenfeld 1988). ANNs are highly simplified brain models that consist of a huge number of information processing elements. These elements are called neurons which are combined into a network by means of connections via modified parameters (*weights*) during the training process (Hopfield 1982; Rutkowski 2004; Asht and Dass 2012). The weights correspond to the impact of each input variable on the result. Networks have a layered structure that can be divided into an input layer, hidden layers and an output layer in terms of accessibility during the learning process (Hopfield 1982; Hecht-Nielson 1989; Tadeusiewicz 1993; Gorska et al. 1997; Han et al. 2012).

The description of the function of an artificial neuron, as presented by McCullogh and Pitts (1943), became the impulse that prompted more extensive research into neural networks and provided, among others, a theoretical foundation. The first prototype of a *Perceptron* network was created by Frank Rosenblatt and Charles Wightman in 1957 and termed a "*neural-like*" network (Rosenblat 1958; Gorska et al. 1997; Asht and Dass 2012). It was developed as a partially electromechanical and partially electronic system used for recognising alphanumeric characters (Rosenblat 1958; Hecht-Nielson 1989). As a single-layer network, *Perceptron* was good at recognising simple letters with, however, rather poor results at recognising complex characters (Minsky and Pappert 1969). The *Perceptron* prototype became the embodiment of the imagination of scientists and engineers, and provided a background for these types of systems which are developed today.

In 1969, Minsky and Papert (1969) showed that single-layer neural networks are subject to a number of limitations, e.g., that these types of networks are only suitable for classifying linearly separable patterns. A new method, termed an *adaptive linear combiner*, was developed and proved to be very useful (Hecht-Nielson 1989). A pattern can be a fingerprint image, words written in italics by hand, a human face or a voice signal. Minsky and Papert's (1969) publication suspended all research into ANNs until the 1980s, when a series of scientific papers authored by Anderson and Hopfield, were published and the technology of largescale integrated circuits were developed that work on non-linear and multi-layer neural networks resumed (Amit 1989; Korbicz et al. 1994; Grzesiak et al. 2006; Han et al. 2012). These networks were found to be superior to classic computers with respect to certain types of problems. Initially, the main application area of neural networks included tasks related to the recognition of signs and signals, e.g., speech, written characters, sound signals or images (Hecht-Nielson 1989; Ozesmi et al. 2005).

BIOLOGICAL BASIS OF ARTIFICIAL NEURAL NETWORKS

The prototype of an ANN is very similar to a human brain. A biological neuron is the basic building block of an animal's nervous system (Jain et al. 1996; Li and Ding 2013; Gurney 2014). Nerve cells, neurons and the connections between them, called dendrites, are the main elements of this complex structure (Hebb 1949; Amit 1989; Churchland and Sejnowski 1992; Jain et al. 1996; Rutkowski 2004). Neurons consist of a cell body containing a nucleus and numerous extensions (numerous and densely branched dendrites and a thicker axon, bifurcated at the end). Dendrites and axon terminals of individual neurons can be extensively branched. A single neuron can receive as many as over 100 000 different inputs (Churchland and Sejnowski 1992; Jain et al. 1996; Gurney 2014). Both dendrites and axons play an important role in that they integrate functions of many nerve cells. Dendrites receive impulses and send single signals to a cell body and axons transmit information from the cell body to the subsequent neurons, e.g., send the output signal. Axons are connected to dendrites of other neurons through synapses. This system of connections results in an aggregation of neurons called a network (Amit 1989; Churchland and Sejnowski 1992; Jain et al. 1996; Gorska et al. 1997; Gurney 2014).

The principle of the function of artificial neural networks

Neural networks, which consist of groups of artificial interconnected neurons, are used to perform complex calculations (Churchland and Sejnowski 1992; Tadeusiewicz 1993; Rutkowski 2004; Ozesmi et al. 2005; Han et al. 2012). The structure of a single artificial neuron is shown in Figure 1.

Neurons, as the basic elements of all neural networks, most often contain a continuous *non-linear* component (Hecht-Nielson 1989; Tadeusiewicz

https://doi.org/10.17221/45/2021-VETMED

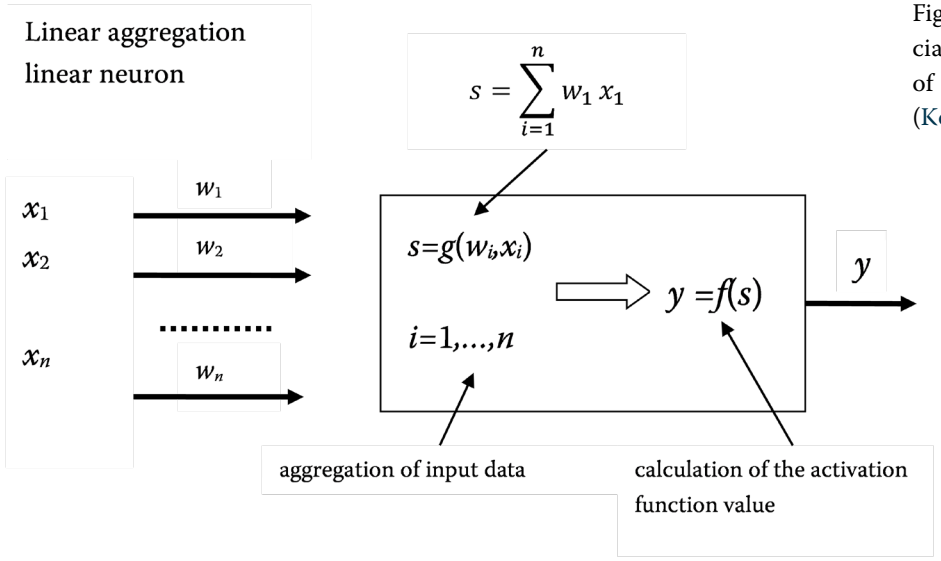

Figure 1. Scheme of an artificial neuron as a basic element of the artificial neural network (Korbicz et al. 1994)

1993; Li and Ding 2013). Non-linear activation functions can be used for the approximation of relationships between variables, where such relationships can be highly complex (Hecht-Nielson 1989; Ozesmi et al. 2005). Each neuron usually performs two tasks: aggregates the input data and weights, and calculates the output value using an activation function (called a transition function). The aggregation of the input data is termed neuron stimulation. There are two basic types of neurons in terms of how data are aggregated: linear and radial neurons (Whitley 1995; Jain et al. 1996; Ozesmi et al. 2005). The difference lies in the way the neuron stimulation signal is calculated.

In the case of linear neurons, the stimulation signal (s) is calculated as a scalar product of the inputs (x_i) and weights (w_i), i.e., $s = \sum w_i x_i$, while in the case of radial neurons, it is calculated as the distance between the input values (x_i) and weights (w_i), i.e., $s = \sum (x_i - w_i)^2$. On the other hand, in terms of the activation function, linear neurons can be divided into strictly linear neurons [with a linear activation function: $f(s) = s$] or sigmoid neurons [with a sigmoid activation function, e.g., $f(s) = 1/\left(1 + e^{-bs}\right)$]. For radial neurons, a Gaussian function – $f(s) = e^{-\beta s^2}$.

The activation function has a great impact on the functioning of neurons as well as of the entire neural network. A number of mathematical functions have been studied in terms of their application for ANNs (Whitley 1995); however, only a few of them have been found suitable for general use:

1. Linear activation function:

$$y = f(s) = \alpha s + \beta \quad (1)$$

where: α, β – straight line parameters, α specifies the slope of the straight line with respect to the x-axis, and β specifies the point of intersection of the straight line with the y-axis.

2. Sigmoid activation function:

$$f(s) = \frac{1}{1 + e^{-\int \beta s}} \quad (2)$$

where: β – specifies the slope of the function.
For large β values, the curve is steep and resembles that of a threshold function, and for small β values, it resembles the curve of a linear function. A sigmoid function takes values from 0 to 1. It is frequently used for building ANNs because neurons with a sigmoid activation function bear the greatest resemblance to the structure of a natural nerve cell (Jain et al. 1996; Gurney 2014).

3. Tangent activation function:

$$f(s) = tgh\left(\frac{\beta s}{2}\right) = \frac{1 - e^{-\beta s}}{1 + e^{-\beta s}} \quad (3)$$

where: β – the same meaning as in the sigmoid function.
The graph of a tangensoidal function resembles that of a sigmoid function, except that the values of the former are in the range of –1 to 1.

4. Gaussian functions:

$$f(s) = e^{-\beta s^2} \quad (4)$$

where: β – the width of the function.
The greater the value of β, the greater the focus of the function around the maximum (the curve is more *pointed*) (see Figure 2). This

https://doi.org/10.17221/45/2021-VETMED

function is used as an activation function for radial neurons (Withley 1995).

The threshold activation function is also referred to as unit jump function. It is described by the formula:

$$f(s) = \begin{cases} 1 \, for \, s \geq s_p \\ 0 \, for \, s < s_p \end{cases} \quad (5)$$

where: S_p – the fixed set threshold value.

The formula is used for a classification or recognition analysis and enables selecting one of the above-mentioned functions (Luger and Stubblefield 2004), where the input pattern is defined as a member of a predefined class (Tadeusiewicz 1993; Withley 1995).

A neuron model works similarly to a biological neuron, which, in neurophysiology, is referred to as the *all or nothing* principle (Tadeusiewicz 1993; Korbicz et al. 1994; Gurney 2014). The threshold activation function is rarely used because these types of networks are prone to learning problems. To avoid problems related to network training, the most commonly used activation functions are those that provide signals that change in a continuous manner, e.g., the functions described above.

Neural network models have a layered structure, e.g., neural networks are organised into three basic layers: input, hidden and output (Hecht-Nielson 1989; Tadeusiewicz 1993; Jain et al. 1996; Bala and Kumar 2017). The output values for one neuron layer are the input values for the next neuron layer. The exception is the first layer (*input layer*), which consists of neurons with input values based on independent variables, and the last layer (*output layer*), which consists of neurons that determine the final result of the calculations over the entire network (Jain et al. 1996; Ozesmi et al. 2005; Bengio 2009). Between the input and the output layer, there is a hidden layer. It can be assumed that the more extensive the hidden layer, the deeper the relationships between the input data can be found by the neural network (Hopfield 1982; Tadeusiewicz 1993; Rutkowski 2004). This is due to the fact that hidden layer neurons generate some intermediate data which provide the basis for the process of finding the final solution to a problem to be solved by a neural network. However, it should be noted that a good neural network should be capable of generalisation, i.e., capable of generalising the learned relationships between the data based on a validation set, so that the network can be used for practical purposes (Hopfield 1982; Tadeusiewicz 1993; Huang et al. 2006; Bengio 2009; Li and Ding 2013). To achieve such a neural network requires not only the modification of the weights during training, but also the proper selection of the network structure, e.g., the number of hidden layers and the number of neurons in the individual layers (Hopfield 1982; Tadeusiewicz 1993; Bengio 2009; Han et al. 2012). Choosing an optimal network structure means working out a compro-

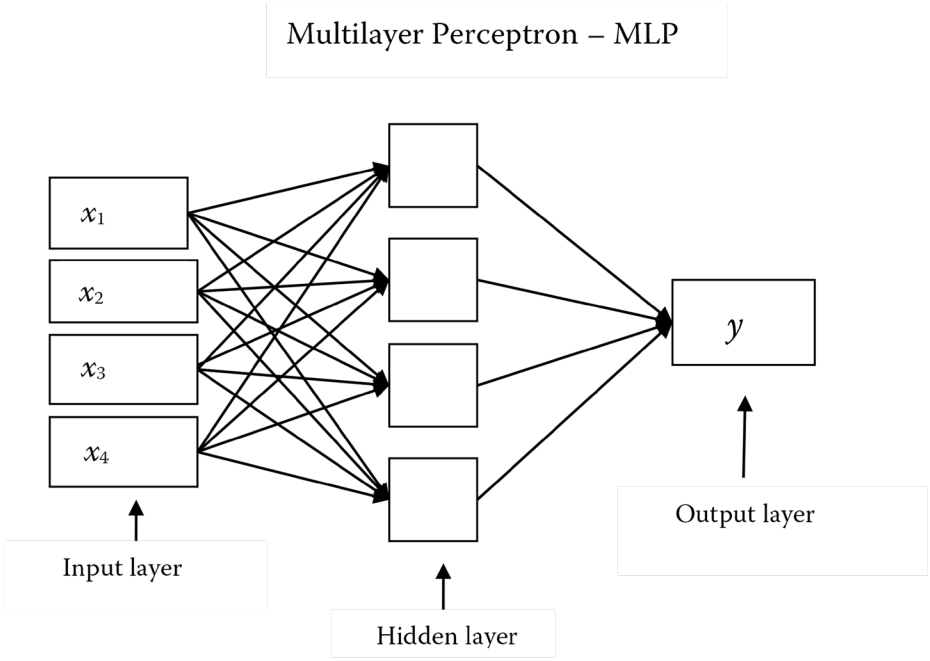

Figure 2. The structure of an example of an artificial neural network – Multilayer Perceptron (Tadeusiewicz 1993)

https://doi.org/10.17221/45/2021-VETMED

mise between an overly extensive and an excessively simplified network without over-fitting a network or setting a local rather than a global minimum. Another, important issue is the size of the teaching training set. Connections within a neural network, which allow communication between all the neurons of one layer with all the neurons of the subsequent layers, are referred to as *one-to-one* connections (Hopfield 1982; Tadeusiewicz 1993). In practice, neural networks with a *one-way* structure (*feed forward* networks) are commonly used, which means a signal travels in only one direction: from input neurons, through hidden neurons, to output neurons (Hopfield 1982; Tadeusiewicz 1993; Jain et al. 1996; Bengio 2009; Asht and Dass 2012; Bala and Kumar 2017). However, if feedback in the network is possible (i.e., later-stage neurons can communicate with earlier-stage neurons), the network can be used for performing complex recursive calculations (Bengio 2009).

Of key importance to the proper functioning of ANNs is to provide them with proper training, i.e., to teach a network based on examples. This process involves modifying parameters of all the individual neurons (i.e., *weight vector*) in such a way as to create a model which describes the relationship between the input and the output variables as accurately as possible (Jain et al. 1996; Ozesmi et al. 2005; Bengio 2009; Gurney 2014; Bala and Kumar 2017). In other words, the aim of the network training is to determine such weight values for which the response of a network is identical (or very similar) to the actual output value in the network training set. The network training process creates no algorithms for the input data processing as is the case with traditional computer software. A network is provided with examples of tasks and then weight factors for connections within the network are modified in accordance with the pre-defined learning strategy: supervised learning (*with a teacher*) or unsupervised learning (*without a teacher*) (Tadeusiewicz 1993; Asht and Dass 2012; Bala and Kumar 2017). The network training process does not use a single pre-defined strategy. It is rather a trial-and-error procedure and the algorithm oriented towards solving a problem emerges in the course of the network training process. A network is provided with various input patterns and uses them to find a solution to a problem by modifying the weights assigned to connections between the neurons so as to minimise the errors in the determination of the output values (Tadeusiewicz 1993; Jain et al. 1996; Bengio 2009; Asht and Dass 2012). Neural networks are trained based on a set of input data, i.e., a training set, which describe the occurrence to be modelled. It is usually divided into three subsets: training, testing and validation. The training subset is used during the network training stage (in the *assimilation period*) to select the weights (Ozesmi et al. 2005; Han et al. 2012). The testing subset allows one to check whether the network is being trained correctly without over-fitting or over-learning. The validation subset enables one to check whether a network has a generalisation capacity (Huang et al. 2006; Bengio 2009; Bala and Kumar 2017). The modification of the weight values continues until the approximation criteria (minimisation of the approximation error) or the error in the validation subset stops increasing (Asht and Dass 2012). The ultimate goal of the network training is to minimise the generalisation error. During the network training process, the error function is most often defined as the sum of squared deviations between the pre-defined and the output value of the network (*SOS*). The network structure and weights are assumed to have been properly chosen if a global minimum is achieved for the objective function (Hecht-Nielson 1989; Tadeusiewicz 1993; Huang et al. 2006; Li and Ding 2013).

There are two types of neural network training methods:

1. Supervised learning (*with a teacher*); and
2. Unsupervised learning *(without a teacher*).

The method of supervised learning (*with a teacher*) makes it possible to verify the response of a neural network because a correct output value is known for each set of input data. In this method, the weight values are selected so as to ensure the smallest possible error between the outputs from a network and the desired response.

There are many algorithms of *supervised learning*, and the most popular one is the so-called delta rule (Whitley 1995; Tadeusiewicz 1993; Bala and Kumar 2017).

Unsupervised learning is a *self-learning* method which is used when the expected (*desired*) response to a pattern included in the examples is unknown. This method requires no previous knowledge about the expected results and defines no preliminary weight values (Tadeusiewicz 1993; Jain et al. 1996; Bela and Kumar 2017). However, it is important for the input data to be classifiable,

https://doi.org/10.17221/45/2021-VETMED

which means that the data are similar in terms of certain qualities. In the *unsupervised learning* method, a sensible solution to a problem is determined by a network on its own, based on a sample of input data, which are used to collect necessary information, e.g., by detecting repeated signals in the input data and combining them into classes. The initial weight values are generated by the network. The most popular method of unsupervised learning is the so-called Hebb's rule (Hebb 1949; Tadeusiewicz 1993; Korbicz et al. 1994; Osowski 1996; Bala and Kumar 2017).

MULTILAYER PERCEPTRONS – MLPs

Multilayer networks are a class of feed-forward neural networks. Single-direction multilayer neural networks have been developed on the basis of a learning algorithm called the *error back propagation* (Jain et al. 1996; Huang et al. 2006; Asht and Dass 2012). These types of neural network are most commonly used in real-life projects and they enable a relatively easy and effective network training process. The *error back propagation* method is a generalisation of delta rules, oriented towards MLP neural networks (Tadeusiewicz 1993; Alsmadi et al. 2009; Han et al. 2012). The MLP training process is simply a numerical procedure for optimising the objective function. It is a gradient method that uses the fact that the gradient of a function indicates the direction of the greatest increase (or decrease) of the objective function.

This feature allows one to minimise the objective function by modifying the variables and weight factors for the neural network in the direction of the greatest increase of the function, i.e., in accordance with the delta rule (Ozesmi et al. 2005). The objective function of a neural network can take various forms, and it is most commonly defined as the square of the error or the square error. It is the difference between the outputs generated by a network for a given input vector and the actual output values (Jain et al. 1996). The activation function is defined as a linear function in the input layer, when the output value of a neuron is calculated. Additionally, a linear or non-linear (e.g., logistic) function can be used in the input, hidden and output layers (Hecht-Nielson 1989; Ozesmi et al. 2005). A typical structure of an MLP network is shown in Figure 3. The training of an MLP network is a multi-stage process composed of the following stages:

1. Random selection of the initial weights and determination of an input vector (*network training vector*);
2. Calculation of the output values using current weights, by sending information from the input layer through the hidden layers to the output layer of the network;
3. Comparison of the output values from the network with those desired for the training vector;
4. Modification of the weights assigned to the output neurons based on the accepted algorithm;
5. Calculation of the error for the training pattern;
6. If the error is smaller than the pre-defined error or the pre-defined number of iterations for the training set has been achieved, the network training should be terminated; otherwise, the process should be resume starting from point 2.

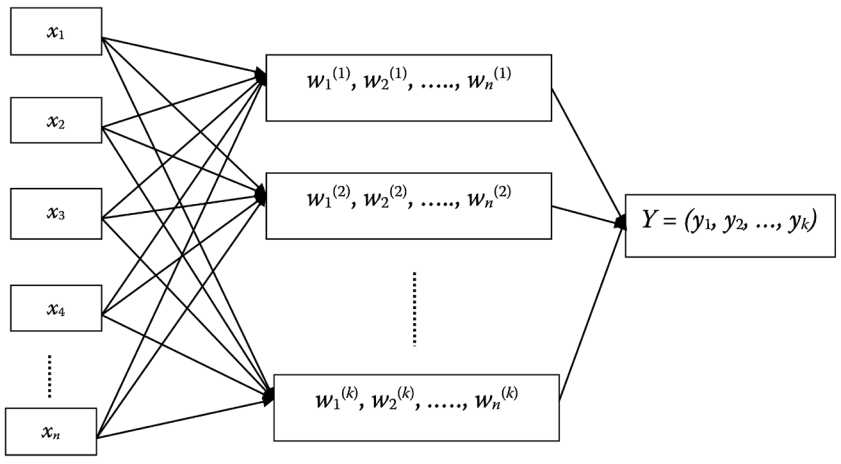

Figure 3. Structure of a single layer MLP neural network – example (Tadeusiewicz 1993)

https://doi.org/10.17221/45/2021-VETMED

RADIAL BASIS FUNCTION – RBF

Radial networks are a natural supplement to sigmoid networks. This is due to the above-described with MLP models differences between linear neurons and radial neurons which have been previously described [space division by a hyperplane (*linear neuron*) or hyperspheres (*radial neuron*)]. The input and output layers of an RBF network work similarly to those of MLP networks (Tadeusiewicz 1993; Jain et al. 1996; see Figure 4). The hidden layer of an RBF network consists of radial neurons, and each of them has its own centre. The aggregation of input data involves calculating the distance between the input data and the neuron centres. As a result of radial neurons, RBF networks enable one to model all types of functions using a single hidden layer, which eliminates the need for defining the number of hidden layers when designing a network. Given the local nature of the approximation provided by the basic functions, choosing the number of hidden neurons in RBF networks is a simpler process when compared to MLP networks. For the hidden layers of RBF networks, Gaussian functions are frequently used as the activation function (Jain et al. 1996; Asht and Dass 2012).

The training of RBF networks consists of the following three stages:

1. Determination of the radial centres (*hidden neurons*);
2. Selection of parameters determining the shape of the basic functions for the hidden neurons (*including the width of the radial function*);
3. Assignment of weights to the neurons in the output layer.

Radial centres (i.e., points around which data groups are symmetrically distributed) can be selected using different methods, e.g., at random or using a taxonomic method called the *k-means* method (Bishop 1995; Osowski 1996; Rutkowska et al. 1997; Asht and Dass 2012). To determine the shape (*width*) of a basis function, the *k-nearest neighbours algorithm* is used and, to teach the output layer, the pseudo-reversal method is used (Osowski 1996; Rutkowska et al. 1997).

MLP AND RBF DIFFERENCES

The two leading and main types among neural networks, the Multilayer Perceptron and Radial Basis Function, are the most popular used networks in animal sciences. The major differences between the these two types of neural networks include the procedure of network training and the structure of the hidden layer. In MLP networks, the hidden layer consists of linear neurons, while, in RBF networks, it consists of radial neurons (Hecht-Nielson 1989; Tadeusiewicz 1993; Jain et al. 1996). Linear neurons represent a hyperplane in multidimensional space and divide the space into two classes, while radial neurons represent a hypersphere that circularly divides the space around the designated centres (Hecht-Nielson 1989; Tadeusiewicz 1993). In addition, there is only one hidden layer in RBF networks, while MLP networks can feature several hidden layers. RBF networks have a greater capacity for precise data interpolation, while MLP networks are used for more accurate data extrapolation (Hebb 1949; Hecht-Nielson 1989; Tadeusiewicz 1993; Osowski 1996; Asht and Dass 2012).

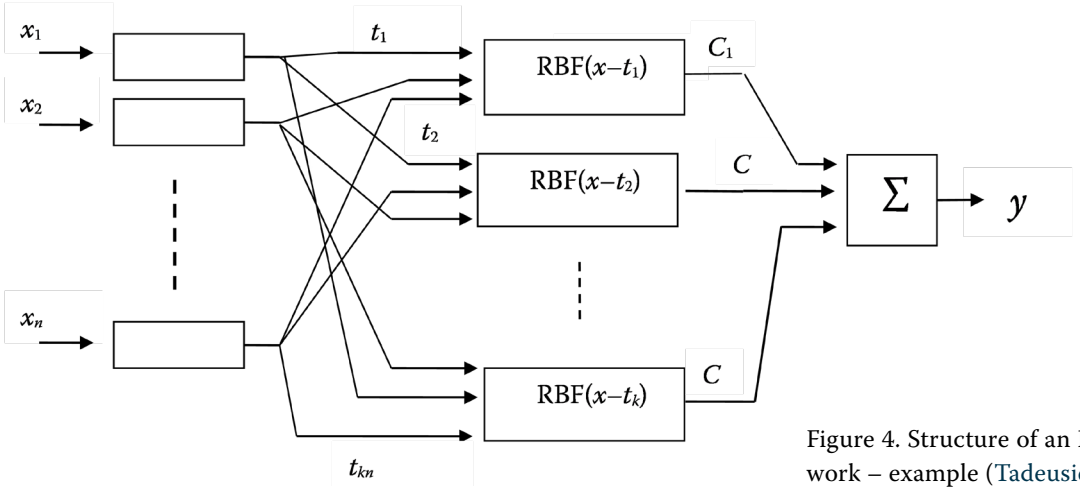

Figure 4. Structure of an RBF neural network – example (Tadeusiewicz 1993)

https://doi.org/10.17221/45/2021-VETMED

Artificial neural network in animal science

Neural networks are also used for solving optimisation and decision-making tasks and for the rapid search of large databases. They are successfully used in forecasting, e.g., sales, prices, stock market results, weather or demand for electric power, in various optimisation processes, e.g., in commercial operations or waste disposal, as well as in the analysis and interpretation of biological and medical research results. ANN models, in particular MLP- and RBF-based models, are still rarely used in animal science. There are some scientific studies related to cattle breeding, e.g., studies into the detection of abnormalities in heifers or the effect of milking of cows on the content of somatic cells in milk, the prediction of lactation efficiency of dairy cows, the detection of cows with possible insemination problems or in the search for causes of the decreasing numbers of Polish cows in herds (Grzesiak et al. 2006; Grzesiak et al. 2010; Zborowski and Grzesiak 2011).

Different types of ANNs, e.g., MLP- and RBF-based networks, have been applied for the classification of dystocia in heifers characterised by high and similar vulnerability to difficult and easy calving (Zborowski and Grzesiak 2011). The results showed the best research results achieved with the use of MLP models in detecting dystocia. In dairy science, a study on the influence of the milking procedure was conducted using an MLP network by Jedrus et al. (2008). It seems to be reasonable to use modern construction methods and of neural model operations to analyse the impact of non-nutritional factors on the chemical composition of the produced milk. The mapping of the relationships between the urea content and subsequent lactation using neural models has been investigated in a number of studies (Gulinski et al. 2008; Fatehi et al. 2012). Bauer and Zychlinska-Buczek (2017) reported that an RBF model seemed to be an appropriate tool for the continuous detailed observations of the urea content in milk obtained from highly productive cows with the use of a wider base of dependent variables.

Few studies have considered predicting the relationship between the rumen fill scores, feed and animal characteristics to improve the predictive capacity of ANN models. Artificial neural networks have been successfully used in the simulation of milk production in goats (Fernandez et al. 2006), *in vitro* methane and carbon dioxide production in the rumen (Dong and Zhao 2014) and for modelling the solid and liquid passage rate in ruminants (Moyo et al. 2017; Moyo and Nsahlai 2018). Craninx et al. (2008) used the machine learning of artificial neural networks with different architectures and training algorithms to determine the prediction accuracy of the chosen model for the branched chain milk fatty acid content in the rumen. A different artificial modelling method was used in a study of the model to predict the rumen weight and fill score in ruminants, as presented by Aruwayo et al. (2020). Another aspect of using ANN models to predict different concentrations of elements in the rumen was presented by Li et al. (2019). The volatile fatty acid content in the rumen was studied by Dong and Zhao (2014) using multi-linear regression models and a three-layer neural network with a backpropagation process.

Mastitis is one of the costliest diseases affecting dairy farms throughout much of the world (National Mastitis Council 1998). Since 1999, some studies have shown that it is possible to develop a neural network model (MLP) to classify the bacterial aetiology of bovine mastitis in herds and cows. Yang et al. (2000) used ANNs to determine the most influential variables associated with the presence or absence of clinical mastitis. The objective of the study conducted by Heald et al. (2000) was to confront the classification capabilities of neural network models with a more conventional statistical analysis, e.g., linear discriminate analysis.

The dairy yield production has also been studied using MLP and RBF models of artificial neural networks. Different value and mix records of milk production were modelled by Salehi et al. (1998). These scientists compared two dedicated neural predictors, and even though the misclassifications were marginal, the network required more time to be developed. An ANN with a feed-forward, backpropagation training algorithm was used by Adamczyk et al. (2005) to predict and determine an objective measurement of the slaughter value of young bulls. Adamczyk et al. (2005) summarise their research results using an ANN, which, as a new tool, can be considered an interesting alternative to traditional models in investigating animal production traits. Different authors have developed a model (MLP) with study contain others ruminants – to classify the behaviour of the sheep in the herd, an MLP-based feed-forward backpropagation neural network was initialised and trained by Nadimi et al. (2012). An ANN (MLP) was suc-

https://doi.org/10.17221/45/2021-VETMED

cessfully compared with a great number of other models to predict the lactation in one domestic Turkish sheep breed (Karadas et al. 2017).

Davoodi and Khanteymoori (2010) employed ANN backpropagation learning models to analyse the performance of learning algorithms and to predict horse racing results. Their study showed an acceptable prediction level; however, some algorithms examined in the study required a longer training period. Some scientists have studied the use of neural networks for diagnosing equine lameness. In such cases, ANN models, e.g., MLP-based models, provide a perfect learning solution and can be successfully used for a non-human diagnosis of equine lameness (Schobesberger and Peham 2002). Network training based on the *error back propagation* algorithm, an ANN model, was successfully developed by Eksteen and Breetzke (2011). A typical endemic disease of African horses was analysed using basic ANN components. In a study relating to horses, a feed-forward network with two hidden layers has been used to predict the incidence of hoof cracking (Suchorski-Tremblay et al. 2001). Moreover, lameness in dairy herds is posing a serious problem to dairy farmers. To detect lameness, Mertens et al. (2012) developed ANN models with classification algorithms. Besides, ANN models like MLP and RBF have been used to detect the lameness of cows based on the body weight (Gupta et al. 2014). In dairy cattle farming, artificial neural networks can be successfully used to predict milk production per the 305-d lactation (Grzesiak et al. 2006; Sharma et al. 2007). The substantial heating problem in dairy farming was previously studied using ANN models (Brown-Brandl et al. 2005). Cattle heat stress can, in extreme cases, besides a decrease in the milk production, cause serious health problems. Brown-Brandl et al. (2005) proposed a regression model to make an accurate prediction that could easily be inserted into a spreadsheet. The study also demonstrated that it is not very difficult to improve the training of ANNs to achieve successful performance. It has been concluded that ANN models provide a useful method for further research in multidisciplinary applications like livestock animal science.

Research opportunities

This paper describes two of the most useful neural networks models. Both models have been extensively investigated using traditional mathematical methods. Each model has its specific strengths and weaknesses, and therefore they can be differently applied to enhance the system-wide performance when solving complex problems. Possible application areas could include monitoring animals and their movements to consider their overall welfare and identify possible diseases, and many other scenarios where computer methods are involved.

Recently, meta data analyses, used in combination with cloud computing for real-time, large-scale data analysis, stored in the cloud, are of considerable importance to agriculture, and, at the same time, artificial models are compared with other existing popular techniques. Studies indicate that new artificial network models are characterised by high accuracy. Studies indicate the high accuracy of new artificial model techniques (Schmidhuber 2015; Kamilaris and Prenafeta-Boldu 2018). Some studies have indicated that ANN models allowed one to refine the accuracy of predicting models and showed better predicting performance, e.g., in sheep (Ince and Sofu 2013), goats (Raja et al. 2012) and dairy cattle (Goyal 2012). Computational ANN-based intelligence helps to address different problems and minimise hazards. Further research is required with the use of MLP- and RBF-based artificial neural network models to achieve better performance for use in animal science in the near future. However, the success rate of each artificial neural network model is highly dependent on the quality of the data set used (Minski and Papert 1969; Tadeusiewicz 1993; Kamilaris and Prenafeta-Boldu 2018). Which of the two artificial models is chosen depends on the problem that needs to be solved. Accordingly, based on general observations, it is concluded that the performance of MLP- and RBF-based artificial neural network models can be improved in terms of the accuracy, training set and time. However, there are problems in which neural networks should not be used because they do not show the expected results or provide the right solution (Tadeusiewicz 1993).

The purpose of this paper was to present the role and importance of ANNs in predicting and describing problems studied in the field of animal science. Overall, the ANN approach described in this study may be an appropriate and powerful decision supporting and monitoring tool, and could be implemented in a computerised information system to improve the animal's health.

https://doi.org/10.17221/45/2021-VETMED

Conflict of interest

The author declares no conflict of interest.

REFERENCES

Adamczyk K, Molenda K, Szarek J, Skrzynski GC. Prediction of bulls' slaughter value from growth data using artificial neural network. J Cen Euro Agric. 2005;6(2): 133-42.

Alsmadi S, Omar B, Noah A. Back propagation algorithm: The best algorithm among the multi-layer perceptron algorithm. J Comp Sci Net Security. 2009;9(4):378-83.

Amit D. Modeling brain functions: The world of attractor neural networks. Cambridge: Cambridge University Press; 1989. 504 p.

Anderson JA, Rosenfeld E. Neurocomputing – Foundations of research. Cambridge (Mass), London (GB): MIT Press; 1988. 729 p.

Aruwayo A, Ahmed Ks, Muhammad IR. Ruminant production and nutrition chemical composition of some selected non-conventional feed resources katsina. 8th Asan-Nias Joint Annual Meeting; Sep 8-12; 2020; p. 1096-8.

Asht S, Dass R. Pattern recognition techniques: A review. Int J Comput Sci Telecomm. 2012;3(8): 25-9.

Bala R, Kumar D. Classification using ANN: A review. Int J Comput Intell Res. 2017;13(7):1811-20.

Bauer EA, Zychlinska-Buczek J. Oszacowanie wplywu okreslonych skladnikow mleka na koncentracje mocznika przy zastosowaniu sztucznych sieci neuronowych [Estimating the effect of some milk components on urea concentration using artificial neural networks]. W Zoot. 2017;3:56-63. Polish.

Bengio Y. Learning deep architecture for AI. Found Trends Mach Learn. 2009;2(1):1-127.

Biecek P, Burzykowski T. Explanatory model analysis: Explore, explain and examine predictive models. 1st ed. New York, USA: Chapman and Hall/CRC; 2021. 309 p.

Bishop A. Neural networks for pattern recognition. New York: Oxford University Press; 1995. 482 p.

Brown-Brandl T, Jones DD, Woldt W. Evaluating modeling techniques for cattle heat stress prodiction. Bio Eng. 2005;91(4):513-24.

Churchland PS, Sejnowski TJ. The computational brain. Cambridge (Mass): MIT Press; 1992. 544 p.

Craninx M, Vlaeminck B, Fievez V. Artificial neural networks to model the rumen fermentation pattern in dairy cattle. Commun Agric Appl Biol Sci. 2008;60(2):226-38.

Davoodi E, Khanteymoori AR. Horse racing prediction using artificial neural networks. Recent advances in neural networks, fuzzy systems & evolutionary computing. Proceedings of the 11th WSEAS International Conference on Neural Networks (NN›10); Proceedings of the 11th WSEAS International Conference on Evolutionary Computing (EC›10); Proceedings of the 11th WSEAS International Conference on Fuzzy Systems (FS›10); "G. Enescu" University, Iasi, Romania; 2010 Jun 13-15; [Wisconsin, USA]: WSEAS Press; 2010. p. 155-60.

Dong R, Zhao G. The use of artificial neural network for modeling in vitro rumen methane production using the CNCPS carbohydrate fractions as dietary variables. Livest Sci. 2014 Apr;162:159-67.

Eksteen S, Breetzke GD. Predicting the abundance of African horses sickness vectors in South Africa using GIS and artificial neural networks. South Afr J Sci. 2011 Jul; 107(7-8):1-8.

Fatehi F, Zali A, Honarvar M, Dehghan-Banadaky M, Young AJ, Ghiasvand M, Eftekhari M. Review of the relationship between milk urea nitrogen and days in milk, parity, and monthly temperature mean in Iranian Holstein cows. J Dairy Sci. 2012 Sep;95(9):5156-63.

Fernandez H, Hughes S, Vigne JD, Helmer D, Hodgins G, Miquel C, Hanni C, Luikart G, Taberlet P. Divergent mtDNA lineages of goats in an Early Neolithic site, far from the initial domestication areas. Proc Natl Acad Sci U S A. 2006 Oct 17;103(42):15375-9.

Goyal S. Artificial neural networks for dairy industry: A review. J Adv Comput Sci Tech. 2012;1(3):101-15.

Gorska T, Grabowska A, Zagrodzka J. Mozg a zachowanie [Brain and behavior]. Warszawa: Wydaw. Naukowe PWN; 1997. 514 p. Polish.

Grzesiak W, Blaszczyk P, Lacroix R. Methods of predicting milk yield in dairy cows – Predictive capabilities of Wood's lactation curve and artificial neural networks (ANNs). Comput Electron Agric. 2006;54(2):69-83.

Grzesiak W, Zaborski D, Sablik P, Zukiewska A, Dybus A, Szatkowska I. Detection of cows with insemination problems using selected classification models. Comput Electron Agric. 2010;74(2):265-73.

Gulinski P, Mlynek K, Salamonczyk E. Zmiany zawartosci mocznika w mleku w zaleznosci od wybranych czynnikow srodowiskowych [Milk urea concetration changes depending on selected environmental factors]. Med Weter. 2008;64(4):465-8. Polish.

Gupta RK, Lathwal SS, Mohanty TK, Ruhli AP, Singh Y. Detection of lameness of cows based on body weight using artificial neural network. 2014 International Conference on Computing for Sustainable Global Development (INDIACom); IEEE; 2014. p. 337-41.

Gurney K. An introduction to neural networks. Hoboken: CRC Press; 2014. 148 p.

https://doi.org/10.17221/45/2021-VETMED

Han J, Kamber M, Pei J. Data mining: Concepts and techniques. Waltham, MA: Morgan Kaufmann/Elsevier; 2012. 703 p.

Heald CW, Kim T, Sischo WM, Cooper JB, Wolfgang DR. A computerized mastitis decision aid using farm-based records: An artificial neural network approach. J Dairy Sci. 2000 Apr;83(4):711-20.

Hebb DO. The organization of behaviour: A neuro-psychological theory. New York: Wiley; 1949. 335 p.

Hecht-Nielsen R. Theory of the backpropagation neural network. International Joint Conference on Neural Networks. IEEE; 1989. p. 593-605.

Hopfield JJ. Neural networks and physical systems with emergent collective computational abilities. Proc Natl Acad Sci U S A. 1982 Apr;79(8):2554-8.

Huang GB, Zhu QY, Siew CK. Extreme learning machine: Theory and applications. Neurocomputing. 2006;70(1-3): 489-501.

Ince D, Sofu A. Estimation of lactation milk yield of Awassi sheep with artificial neural network modeling. Small Rumin Res. 2013;113(1):15-9.

Jain AK, Mao J, Mohiuddin KM. Artificial neural networks: A tutorial. Computer (Long Beach Calif). 1996;29(3):31-44.

Jedrus A, Nizewski P, Lipinski M, Boniecki P. Neuronowa analiza wplywu sposobu doju i wybranych cech zootechnicznych krow na liczbe komorek somatycznych w mleku [Neuron analysis of the influence of the way of milking and selected characteristsics of cows on somatic cell count in milk]. Tech Rol Ogrod Leśna. 2008;4:22-4. Polish.

Kamilaris A, Prenafeta-Boldu FX. Deep learning in agriculture: A survey. Comput Electron Agric. 2018;147:70-90.

Karadas K, Tariq M, Tariq MM, Eyduran E. Measuring predictive preformance of data mining and artificial neural network algorithms for predicting lactation milk field in indigenouns Akkaraman sheep. Pak J Zool. 2017;49(1):1-7.

Korbicz J, Obuchowicz A, Ucinski D. Sztuczne sieci neuronowe: Podstawy i zastosowania. Warszawa, Poland: Akademicka Oficyna Wydawnicza PLJ; 1994. 251 p.

Li MM, Sengupta S, Hanigan MD. Using artificial neural networks to predict pH, ammonia, and volatile fatty acid concentrations in the rumen. J Dairy Sci. 2019 Oct;102 (10):8850-61.

Li H, Ding S. Research and development of granular neural networks. Appl Math Inf Sci. 2013;7(3):1251-61.

Luger G, Stubblefield W. Artificial intelligence: Structures and strategies for complex problem solving. 5th ed. Amsterdam: The Benjamin/Cummings Publishing Company, Inc; 2004. 928 p.

Mertens K, Vangeyete J, Van Weyenberg S, De Campeneere S, Van Nuffel A. Automatic lameness detection in dairy cows using artificial neural networks. Proceedings of the Power and Machinery. International Conference of Agricultural Engineering – CIGR-AgEng; Valencia, Spain; Jul 8-12; 2012.

McCullogh WS, Pitts WH. A logical calculus of the ideas immanent in nervous activity. Bull Math Biophys. 1943; 5(4):115-33.

Minsky M, Papert S. Perceptrons: An introduction to computational geometry. London, England: MIT Press; 1969. 258 p.

Moyo M, Gueguim Kana EB, Nsahlai VI. Modeling of digesta passage rates in grazing and browsing domestic and wild ruminant herbivores. S Afr J Anim Sci. 2017;47(3):362-77.

Moyo M, Nsahlai IV. Rate of passage of digesta in ruminants; Are goats different? In: Kukovics S, editor. Goat science. London: IntechOpen; 2018. p. 39-79.

Nadimi ES, Jorgensen RN, Blanes-Vidal V, Christensen S. Monitoring and classifying animal behavior using ZigBee-based mobile ad hoc wireless sensor networks and artificial neural networks. Comput Electron Agric. 2012; 82:44-54.

National Mastitis Council. Current concepts of bovine mastitis. 4th ed. Madison, WI, USA: National Mastitis Council, Inc.; 1998. 64 p.

Osowski S. Sieci neuronowe w ujeciu algorytmicznym [Neural networks in an algorithmic approach]. Warszawa: Wydawnictwa Naukowo-Techniczne; 1996. 349 p. Polish.

Ozesmi SL, Ozesmi U, Tan CO. Methodological issues in building, training, and testing artificial neural networks in ecological application. Ecol Modell. 2005;195 (1):83-93.

Raja TV, Ruhil AP, Gandhi RS. Comparison of connectionist and multiple regression approaches for prediction of body weight of goats. Neural Comput Appl. 2012;21(1):119-24.

Rosenblatt F. The perceptron. A theory of statistical separability in cognitive system. Project Para Report No. VG-1196-G-1. Buffalo, N.Y.: Cornell Aeronautical Laboratory; 1958. 268 p.

Rutkowska D, Pilinski M, Rutkowski L. Sieci neuronowe, algorytmy genetyczne i systemy rozmyte [Neural networks, genetic algorithms and fuzzy systems]. Warszawa, Lodz: Wydaw. Naukowe PWN; 1997. 410 p. Polish.

Rutkowski L. Adaptive probabilistic neural networks for pattern classification in time-varying environment. IEEE Trans Neural Netw. 2004 Jul;15(4):811-27.

Salehi F, Lacroix R, Wade KM. Improving dairy yield predictions through combined record classifiers and specialized artificial neural networks. Comput Electron Agric. 1998; 20(3):199-213.

Schobesberger H, Peham C. Computerized detection of supporting forelimb lameness in the horse using an artificial neural network. Vet J. 2002 Jan;163(1):77-84.

https://doi.org/10.17221/45/2021-VETMED

Schmidhuber J. Deep learning in neural networks: An overview. Neural Netw. 2015 Jan;61:85-117.

Sharma AK, Sharma RK, Kasana HS. Prediction of first lactation 305-day milk yield in Karan Fries dairy cattle using ANN modeling. Appl Soft Comput. 2007;7(3):1112-20.

Suchorski-Tremblay AM, Kok R, Thomason JJ. Modelling horse hoof cracking with artificial neural networks. Can Biosyst Eng. 2001;43(7):7.15-7.22.

Tadeusiewicz R. Sieci neuronowe [Neural networks]. Warszawa: Akademicka Oficyna Wydawnicza; 1993. 195 p. Polish.

Whitley D. Genetic algorithms and neural networks. Genetic Algor Eng Com Sci. 1995;16(8):13-52.

Yang XZ, Lacroix R, Wade KM. Investigation into the production and conformation traits associated with clinical mastitis using artificial neural networks. Can J Anim Sci. 2000;80(3):415-26.

Zborowski D, Grzesiak W. Detection of heifers with dystocia using artificial neural networks with regard to ERalpha-BGLI, ERalpha-SNABI and CYP19-PVUII genotypes. Acta Sci Pol Zootechnica. 2011;10(2):105-16.

Received: March 25, 2021
Accepted: December 29, 2021
Published online: March 2, 2022