Ákos Nyitrai[1], Attila Gere[1], László Sipos[1]

Food science applications and international trends of artificial neural networks

1. Summary

Recently, research has been focusing increasingly on the system of artificial neural networks, and its results are used in many places by industrial practices. The success of these networks lies in their ability to recognize the complex relationships and patterns in data, as well as to predict unknown samples, thus enabling value and category predictions with high certainty. Artificial neural networks are very efficient tools for modeling non-linear trends within data. In many cases, they perform well where traditional statistical tools provide unsatisfactory results or unable to solve a given research problem. In our work, the operation principle and structure (topology) of artificial neural networks are summarized, as well as the classification and application possibilities of the networks. The latest food science applications are presented separately, based on the usage type (prediction, classification, optimization). Results show that artificial neural networks possess many beneficial properties, making them especially suitable for solving food science tasks.

2. Introduction

The idea of artificial neural networks was based on the neurotransmission mechanisms of the human brain. After studying the human brain and the nervous system, the idea was born to create an artificial network similar to the living organism. The brain consists of a large number of brain cells, i.e., neurons: according to the latest measurements, the average person has 10 to 15 billion cortical neurons [1]. During its operation, a neuron receives a number of impulses from other neurons through several dendrites. Depending on the impulse received, the neuron may transmit signals to other neurons through simple axons, linking dendrites to other neurons. Thus, a single impulse information may travel through several millions of neurons before reaching the appropriate layer of the brain, or the output layer of the network. The role of the threshold is that if this value is reached, the neuron will be activated, while below this value it will not happen.

The artificial neural network is also composed of brain-like elements. Following the example of biological systems, these processing elements have also been named neurons. After the biological sample, an output is generated by each of the neurons in the artificial network, based on a certain number of inputs. The output thus obtained is a relatively simple function of the inputs.

Originally, artificial neural networks (ANN) were non-linear approximation procedures designed to mimic human brain function, and they could be characterized by the most important brain characteristics (versatility, parallelism, adaptive response to external stimuli, efficient shape recognition ability even in the case of noisy data, etc.). The first research objective was to represent the human system accurately, however, today neural networks are used as mathematical, rather than biological models. Artificial neural networks resemble human brain function in two ways: they possess an ability to learn and are able to sort information. ANNs are basically a family of non-linear calculation methods [3].

The scientific foundations of the system were laid down by McCulloch and Pitts in 1943 [4], when the world's first rudimentary neural network was created.

[1] Szent István University, Faculty of Food Science, Department of Postharvest Science and Sensory Evaluation

They relied on three sources: basic knowledge regarding physiology and the operation of cerebral neurons, Russell and Whitehead's formal analysis of propositional logic, and Turing's calculation theory [5]. By combining these three different theories, they created the concept of an artificial neuron, which they called Threshold Logic Unit (TLU) at the time. This idea was further developed by Hebb [6], and finally the concept of the perceptron was created by Rosenblatt in 1959 [7]. It is a neuron that does not have a memory, and there is no difference between inputs in terms of weighting. A network consisting of perceptrons is already capable of solving simple classification tasks. Following this, the initial enthusiasm waned, mainly due to the technical difficulties of practical implementation. In 1982, a study was published by Hopfield [8], in which two concepts of key importance were presented, making it possible to overcome the limits formulated earlier: the non-linear relationship between the total input of the neuron and the output obtained, and the possibility for feedback between the outputs and the inputs. This important research result and the introduction of the backward algorithm in 1986 prompted researchers to develop neural networks to use problem-solving algorithms. This way, they can be an effective problem-solving alternative in the areas of finance, medical diagnosis, process control, engineering tasks, geology, weather forecasting, data processing, data mining, traceability, as well as the mapping of consumer preferences, among other things [9].

3. Characteristics of artificial neural networks and their operating principle

Neural networks have been used successfully because of several very important properties [10]:

1. they are very complex non-linear computing tools capable of modeling extremely complex functions,

2. learning ability (the data structure automatically learns from the representative data with the help of the training algorithm developed over time),

3. widespread applicability (not limited to numerical data, images and texts can be analyzed as easily as figures),

4. parallelism, ability to generalize, high speed and fault tolerance.

The neural network can be thought of as a modeling of the functional relationship between the input and the corresponding „output" variables:

$$y = f(x)$$

where x and y are the input and output vectors, and f indicates the relationship or activation function. Depending on the application, vector **y** displays the output value: in a reduced dimension space (exploratory analysis), the binary vector of class memberships (classification), or a real value dependent vector (regression).

A neural network is a system that performs calculations on the input data, and the output data is one or more calculated values. When a neural network is built and trained for a specific task, roughly correct output values are obtained, based on the inputs. There are numerous criteria for measuring deviations, because it always depends on the nature of the actual task which of these should be used for the analysis of the results.

The most commonly used criteria are as follows: minimization of the mean squared error (MSE), the binary cross entropy (BCE) or the multiclass cross-entropy (MCCE). In addition, the network can also be analyzed using the Kullback-Leibler divergence or the hinge loss function, but these are less popular approaches. The special feature of ANNs is that the functional relationship described by the equation is perfect, as opposed to traditional mathematical models. In fact, in the case of neural network modeling, the function-like relationship between the output and the input space is clearly defined. The neural network consists of interconnected units called nodes and neurons. Neurons are characterized by the fact that the n weighted inputs and the input selected for a constant (*bias*) follow a certain, typically non-linear function. The summing and the subsequent function are collectively known as the transfer function, which gives the weighted sum of the inputs, and is a subsequent, usually non-linear function. Each neuron within the network performs a part of the calculations: some of the data are taken as the input, and after a few simple calculations the output data is provided. The output data of a neuron is transmitted to another neuron, except for neurons that provide the ultimate result of the entire system.

Neurons are arranged in layers. The neurons of the input layer receive the inputs for the calculations, such as length or weight. These values are enter the first hidden layer that performs the calculations and transmits the outputs to the next layer. This next layer may be another hidden layer, if there is such one. The output results of the neurons of the last hidden layer are transmitted to the neuron(s) that calculate the final result of the network [12].

The peculiarity of artificial neural networks is based on the fact that, during their operation, they use a large number of parallelly linked simple arithmetic units called neurons, because of the biological similarity. In mathematical language, a neuron can be defined as a non-linear summing function characterized by certain parameters, whose variables are called the inputs of the neurons, and its value is the output. In this context, parametrization can occur in two different forms:

1. Parameters are linked to the input of the unit, so a "global input" is built for the neuron as a linear combination of the inputs, weighted by the parameters (called weights w_i); the output of the unit then appears as a non-linear function (f) of the global input:

$$y = f(w_0 + \sum_{i=1}^{n} w_i x_i$$

2. Parameters are assigned to the non-linearity of the neuron, i.e., they participate in the determination of the unit. This happens, for example, if „f" is a Gaussian radial basis function:

$$y = \exp[-\sum_{i=1}^{n} \frac{(x_i - x_i)^2}{2w_0^2}]$$

Several versions of this are known in science, and the most appropriate one is generally selected depending on the nature of the tasks to be solved. In other words, the transfer function is nothing but a summing function that gives the weighted sum of the inputs, and a subsequent, usually *non-linear* function. Some commonly known transfer functions are shown in **Figure 3**.

In the simplest cases, the inputs of the neurons occupy equal inputs and do not have a memory. For such a neuron, the x_i scalar inputs are summed with weights w_i (i=0,1,..., N), and the weighted sum will be the f(.) non-linear element **[13]**.

The network represents the construction of a non-linear function of two or more neurons. The relationships between the neurons are typically illustrated in a graph where units and units acting on the same input variable are arranged in layers. Weights that regulate the non-linear functions are usually represented as lines associated with the different units in the layers.

Training of a network means the process of optimization of the calculation parameters, where the goal is to obtain approximately correct outputs for the input data. This process is carried out on the one hand on the data used for training, and on the other hand based on the algorithms. Different calculation parameter combinations are selected by the training algorithm, and these combinations are evaluated by using them for each training. This way it is determined how good the answers provided by the network are. Each parameter combination is in fact a test. The training algorithm selects the new parameter combinations based on the results of the previous test. A neural network is actually a computational model that can be used to develop various computer systems. A neural network can be constructed from simple process elements, in which case each element acts as a neuron. There are different neural networks that can differ in structure, in the internal calculations carried out by the neurons, or in training algorithms **[13]**, **[14]**, **[15]**.

ANNs have in fact become popular in the past decade, as neural network types, as well as their testing, validation and visualization have been integrated into different statistical and data mining software packages. Neural networks are typically created by software developers in separate modules: MatLab (Neural Network Toolbox), Statistica (Neural Networks), Palisade (NeuralTools), SPSS (Modeler), Alyuda (Neurointelligence), NeuroDimension (Neurosolution), etc.

ANNs are more robust and outperform other calculation methods in six categories: pattern recognition, clustering, function modeling, forecasting, optimization and control **[16]**. In the field of food science, there are three main areas of application:

1. exploratory analysis

2. prediction

3. classification

4. The structure of artificial neural networks

Generally, the behavior of a network is determined by the following: structure (number of hidden layers and the nodes in them), weight of the relationships (relationship parameters) and the distortion (a parameter related to the neurons) transformation function that calculates the value of the output signal **[17]**. According to their structure and operation, networks can be classified into several groups. By analyzing current literature data, it has been shown that today there is an unfathomable number of neural network types. In a previous study, Maren listed 48 different types of artificial neural networks, while Pham distinguished more than 50 types **[18]**, **[19]**. There was a huge overlap between these types and, in terms of practical application, only a few networks are used by researchers in their publications. In the following, these networks are presented in detail:

1. Multi Layer Feed Forward Neural Net (*MLFNN*)

2. Radial Basis Function Neural Net (*RBFNN*)

3. Kohonen/Self-Organizing Maps (*SOM*)

4. Generalized Regression Neural Net (*GRNN*) and Probabilistic Neural Nets (*PNN*)

We can distinguish 3 types of layers: the input layer (it contains as many units as the number if independent variables (x), and it only feeds the data into the network), the output layer (it consists of as many neurons as the number of components of the dependent vector (y)), and the hidden layers. These hidden layers are the location of complex calculations and algorithmic solutions. The optimum number of hidden layers is unique in each case, therefore, the

analysis of this is usually part of the neural network measurement [20], [21], [22] (**Figure 4**).

The Multi Layer Feed Forward Neural Net (*MLFNN*) is characterized by a lack of circular feedback [24], information flows only in one direction: input layer → hidden layer (one or more) → output layer, i.e., the output vector of a layer is the input vector of the next layer. MLF is the basis of neural networks, a number of complex procedures are based on this network structure. The disadvantages of MLFNN include that sigmoid functions are slow to approximate and that the solving algorithm is not visible. Generally, this is called the black box of the network, as we do not know the details of the solution, only the final result is calculated by the network. The essence of neural networks is learning from a sample set, therefore, network performance greatly depends on the multitude of trained samples. If the network is built on the basis of a bad, limited or too noisy sample set, then output values will be calculated by the system based on faulty relationships. The success of training is an essential element, because the network tends to overlearn input data. In this case, the network does not learn the relationships between the elements of the sample set, but the individual samples, which leads to obvious distortion. The advantages of MLF networks are that compared to other network types (*Generalized Regression Neural Net, GRNN; Probabilistic Neural Net, PNN*) they are smaller in size, much more reliable and less sensitive to trained data outside the range. Another advantage of MLF networks is that it is suitable for the generalization of very small amounts of training data sets [25].

The structure of radial basis function neural nets (*RBFN*) is identical to that of the MLFN network (one input, one output and a hidden layer), the main difference being in the functions of the neurons. The radial basis function network is a system in which the hidden layer achieves a non-linear rendering with radial basis functions [13]. In fact, the output layer consists of linear units, and the hidden layer forms so-called radial basis functions, as shown by the name of the network. The radial function is a real-value function, the value of which depends only on the distance between its input vector and a so-called centroid (*barycenter*) [26]. In the case of RBFN modeling, there are three main parameters: output weights (w_{kj}), centroids (β_j) and shell factors (μ_j), and they are typically optimized in this order (**Figure 5**).

Kohonen-based structures, self-organizing maps, (*SOM*) are designed for non-linear rendering from a multidimensional space into a single plane, with which the distances between the samples (similarities and differences) can be determined. The basic elements of the Kohonen network, which are found in one layer, typically in the grid points of a plane, make up the linear summing function. Each input is connected to all of the nodes of the network [28], [29], [13] (**Figure 6**).

The neighborhood of a neuron is generally considered a square, rectangular or hexagonal, meaning that each neuron has 4, 6 or 8 closest neighbors. From an algorithmic point of view, Kohonen modeling performs competitive learning, i.e., only one neuron of the 2D layer is selected after each input is entered into the network („winner takes all" principle), and this neuron will be the position where the sample is mapped. The winning neuron is the one whose weight vector is most similar to the input sample [30] (**Figure 7**).

Generalized Regression Neural Nets (*GRNN*) and Probabilistic Neural Nets (*PNN*) operate on similar principles. *GRN* networks can be used for value predictions/function estimations, while *PNN* networks can be used for category predictions/classification. The main advantages of *GRNN/PNN* networks are that it is not necessary for the user to make decisions about network design, i.e., no setting of the topology is required (number of hidden layers and nodes), the network optimizes itself. In each case, these networks contain two hidden layers with one neuron per training case in the first hidden layer, and the size of the second layer is determined on the basis if the training data.

In the general regression neural net, the closer the known case is to the unknown one during the training, the greater its weight will be when calculating the unknown dependent value. Inputs are summed by two nodes of the summation layer, while the output node divides them for creating the prediction (**Figure 8**). For probability networks, the output values of the summation layer can be interpreted as estimates of the probability density function for each class. The output neuron selects the category with the highest probability density function, and chooses it as the predicted category (**Figure 9**).

Advantages of the *GRNN* include that it accelerates training, thanks to which the network will work faster than *MLFN* networks. Unlike traditional feed forward networks, *GRN* networks are always able to find the global minimum and so there is no risk of the system approaching a local minimum. *PRNN* networks not only classify, but also provide probabilities, based on which the case falls into the different dependent categories. Probabilistic networks can be capable of faster operation and more accurate predictions than simple feed forward methods. At the same time, *PRNN* networks classify new cases more slowly and require more storage space than *MLFN* networks.

5. Food science applications of artificial neural networks

The number of food science applications of artificial neural networks (*ANN*) has been increasing steadily over the past 15 years. In the Science Direct scientific database, a search for the words ANN and food provided 135,800 scientific papers, 91,405

of which were journal articles, 12,286 were book chapters, 6,782 were review articles and 406 were encyclopedias. If we look at the nature of the trend in recent years, from 2001 to 2017 an overall increase can be observed. The slope of the lines connecting the years shows the dynamics of the change. From this point of view, the section belonging to the year 2017 exhibits the greatest increase in publication (section with the highest slope) (**Figure 10**).

The spread of artificial neural networks in the field of food science is due to their abilities presented earlier, since they can solve a number of tasks, the most important of which are: modeling of complex relationships, classification, categorization, regression between non-linear variables. Their precondition is a large teaching data set, where there is a correlation between the measured and estimated data. The scope is very wide, therefore, any classification would be subjective. In the following, we sought to address the widest possible range of applications, based on the following classification: prediction, classification, optimization.

5.1 Prediction applications in food science

Castro et al. **[33]** used neural networks to predict the quality of the fruit of the macauba oil palm. This plant plays a prominent role in tropical countries as it is an excellent source of biomass/bioenergy. For the measurements, the biometric data of 543 fruits were used. Among physico-chemical variables, data that can be measured easily were used as input data (weight of the fresh fruit, axis length, axis width, width of the endocarp at two different locations). As output values, properties were selected that can replace those that require multiple measurements (dry weight of skin, pulp dry weight, endocarp dry weight, pulp oil content, seed dry weight, seed oil content). Data were analyzed using multivariate linear regression and neural networks. Data obtained using destructive and non-destructive tests were separated, and they were analyzed in separate systems. Samples were grouped by clustering according to their place of origin. During the measurements, one group was the teaching set, while the other was the validation data set. During validation, significantly better results were given by the feed forward networks built (R²=0.96 for destructive and R²=0.97 for non-destructive measurements) than multivariate linear regression (R²= 0.82; R²=0.81).

Singh et al. **[34]** used back propagation neural networks to predict the organoleptic qualification of UHT milk samples. The effects of physico-chemical and biochemical processes that take place during storage on the organoleptic properties of milk were investigated. In addition to the oxidative, proteolytic and lipolytic indices of milk, hydroxymethylfurfural (HMF) content and color data measured by reflectometry were used as input values. The prediction gave quantified values of total sensory quality and taste value. In their research, organoleptic evaluation was carried out by 5 trained judges. Besides the neural network, prediction was also performed by regression modeling. The accuracy of the regression models was R²=0.869 in the case of the taste value and R²= 0.917 in the case of total sensory quality. The efficiency of the artificial neural networks was higher than that of regression models (R²=0.95 and R²=0.97). As for the topology of the networks, the greatest accuracy for taste value testing was given by the network with the 5-15-1 structure. When analyzing total sensory quality, the network with two hidden layers proved to be the best, with a topology of 5-3-3-1. The prediction provided by the neural networks proved to be extremely accurate.

Krishnamurthy and his fellow researchers **[35]** used neural networks to predict the consumer preference of beef products. The input data set contained the results of two types of tests: the judgment of a panel of 10 experts (quantitative descriptive analysis: 28 properties were evaluated, of which 4 pertained to appearance, 7 to aroma, 11 to taste, 3 to aftertaste, 2 to mouth sensation and 2 to smoothness), and the result of a 100-person consumer preference test. During the consumer preference test of 10 samples, overall liking, taste, aftertaste and mouth sensation of the different products were evaluated. Following the experiment, regression fits were carried out (multivariate linear regression (*MLR*), principal component regression (*PCR*), partial least square regression (*PLSR*)) which gave accurate predictions, but did not prove to be robust enough to be able to handle the data provided by the panel of trained judges. The multi layer feed forward network with the best prediction ability had a 9-5-2 topology (9 neurons in the input layer, 5 in the hidden layer and 2 in the output layer), with which an accuracy of 98% was achieved. Clustering accuracy was 95% and 80%.

León-Roque et al. **[36]** determined the fermentation index of cocoa beans by involving neural networks. The novelty of the experiment was that they used a simple color-based image analysis, i.e., the fermentation index was provided by the system on the basis of the color of the cocoa bean. The study was carried out with 120 samples, coming from three different growing areas (40 cocoa beans from each). Digital pictures were taken of the whole cocoa beans and of the samples cut in two. After this, the degree of fermentation was determined by chemical instrumental analytical methods. During the statistical evaluation, feed forward networks were used, with the inputs being the R, G, B color values of the surface and the inside of the sample, as well as the absorption spectra of the extracts, and the predicted characteristics were the fermentation index values. The efficiency of the network with the best prediction ability for the sample set proved to be good (81%).

Vásquez et al. **[37]** investigated Swiss-type cheese samples. Their goal was to observe and evaluate

structural changes during maturation. In the study, 40 different cheese samples were tested during their maturation. During the process, hyperspectral images of the cheese samples were taken, and the measurement results of these were the input data of the feed forward neural network. The structural hardness of the samples was determined using texture profile analysis (TPA), and these formed the output values of the network. To map the relationships between the hardness and the spectral profiles, PLS regression and neural networks were tested. Based on the results, the feed forward neural network provided better results (R^2= 0.96) than PLS regression (R^2= 0.94), i.e., it could predict the hardness of Swiss-type cheese samples more accurately, based on their hyperspectral images.

Bahramparvar and her fellow researchers [38] used artificial neural networks to predict the consumer acceptance of ice creams. Based on the results of a panel of trained judges, consumer preference was predicted for 9 different ice cream test samples. When building the neural network, experimental organoleptic properties given by the judges (appearance, taste, texture, freezing character, firmness, viscosity, silkiness, liquefaction) were used as input data. The output was the global consumer acceptance values of the 10 judges. The neural network used was a simple feed forward multilayer neural network. The network was trained on 30% of the data, validated on 10%, and tested on 60%. Several network topologies were investigated: the number of hidden layer neurons varied between 2 and 20. More accurate predictions were obtained in the case of a network with 10 neurons in the hidden layer, with a mean absolute error (*MAE*) of 0.27, and a very high correlation coefficient (R^2=0.96). The tests also showed that the most important factors for consumer acceptance of ice creams were taste and texture.

5.2 Classification applications in food science

Wang et al. [39] created a tracking system. The food tracking system developed by them not only tracks processes, but also evaluates the quality of food at each stage of the supply chain. Quality assessment was carried out at the various stages using fuzzy logic and the final quality was classified using an artificial neural network (high, medium or low quality product). The sample set consisted of 20 pork samples. The input values of the fuzzy logic were determined at each stage by different influencing factors, for example, a disease of the animal or feed quality, etc., played an important role during farming, while during the retail distribution stage of the product decisive factors are the quality of the packaging or the degree of microbiological contamination. Based on these, meat quality was determined at each stage. The final quality values of the 5 stages formed the input data of the neural network. As a control of the network, meat products were also examined by an expert panel and classified into the same 3 categories. Best results were obtained with a 5-5-3 topology back propagation network (90% accuracy), as the result of the network and the classification of the expert panel differed in the case of 2 products out of 20.

Zarifneshat et al. [40] used neural networks to examine the quality of Golden Delicious apples. The goal of the research was to investigate the effect of various factors on the extent of bruising. The apples were tested using several different rheological measurements, including penetrometer distance and acoustic hardness measurements. Input variables of the neural network were the energy used for bruising, compressive force, surface curvature radius, temperature and acoustic hardness values. The predicted output value to be estimated was selected to be the volume of bruising. 120 were tested, 96 values were used to teach the network and 24 results were used to test it. A back propagation network type was the method of choice, and a declining learning-rate factor algorithm was also used. Regression values exceeded 90% in both cases, but the best results were obtained by the neural network calculated with the declining learning-rate factor algorithm (5-30-1 topology).

Da Silva Sauthier et al. [41] investigated mango samples by HPLC, the study focused mainly on the concentrations of the 12 bioactive phenolic components found in mango, and the application of neural networks appeared as exploration and classification sub-tasks. The test set of 42 samples was made up of 3 different varieties (Tommy, Rosa, Espada). During the measurement, data sets, such as total phenolic content, antioxidant activity or total anthocyanin content were separated. A Kohonen network was used to map sample distribution, and input data were the HPLC analytical results of the different fruits. Samples were divided into clearly separated groups by the network based on their functional composition. The study also showed that mango samples are rich in polyphenol-containing ingredients, such as ellagic acid, gallic acid, rutin or catechol, therefore, it can be an excellent raw material for functional foods or the pharmaceutical industry.

Yu et al. [42] studied the quality of green tea by rapid measurements, and the artificial neural network was built from the data of an electronic nose. Measurements were carried out on 5 green teas of different quality and value. Of the samples tested, the price of the cheapest product was 240 yuan per kilogram, while the most expensive was 3,600 yuan/kg, this is why it was important to develop a green tea classification system. During the analysis, volatile components of the samples were measured by the instrument of an electronic nose, and these data were then used as the input data of the neural network. In their research, data processing was carried out using cluster analysis, as well as back propagation (*BP*) and probabilistic (*PN*) neural networks. The

interesting part of their work is that the three principal components obtained as the result of principal component analysis were the input data of the neural network. Their results demonstrated that both the back propagation and the probabilistic network successfully completed the classification task: 100% and 98.7% in the case of the teaching set, and 88% and 85.3% success rates for the test set.

Anjos and her fellow researchers [43] used neural networks to determine the botanical origin of honeys. The honeys examined could be divided into two groups: single-flower honeys (honeys typically coming from a single plant such as acacia, orange, almond, lavender, etc.) and mixed flower honeys. The classification system was created on the basis of physical and chemical parameters: moisture content, electric conductivity, water activity, ash content, pH, free acidity, color coordinates (L*, a*, b*), as well as total phenol content (input parameters). In their work, a total of 49 samples were tested from 14 different regions. Based on the results of a back propagation neural network, single-flower honeys could be distinguished from mixed flower honeys during classification with an accuracy of 95%. The result of the research was that the botanical origin of honeys can be predicted on the basis of the color and electric conductivity of the samples.

Fadilah et al. [44] used neural networks combined with an intelligent visual system to determine the ripeness of the fruit of the oil palm (Fresh Fruit Bunch, FFB). The measurement was based on two pillars, on the one hand, samples were analyzed by a network trained with the pictures of the FFBs, and on the other hand, MLF neural networks were created using data simplified by principal component analysis. Based on the results, samples could be classified into 4 categories: unripe, underripe, ripe and overripe. 120 images were included in the testing set, 28 in the validation set, and 60 in the independent test set. After the development of the network it turned out that the best result of the network taught with the full data set was 91.67%. On the other hand, the classification accuracy of the network trained on the data set reduced by principal component analysis was 93.33%, i.e., 1.67% better.

Huang [45] used neural networks combined with a visual system to classify unhealthy and insect-infected betel nut samples. Chewing and consumption of betel nuts is the number one cause of oral cancer in Taiwan, and its commercial turnover exceeds USD 3.4 million per annually. Input data for the neural network built for the investigation consisted of 6 geometric parameters (main axis length, secondary axis length, number of axes, area, rim and image density) and the color characteristics (image grayness and R, G, B data). For quality classification, a back propagation network was used. The 144 test samples could fall into one of three categories: excellent, good and bad. Learning and testing of the network was performed on

images taken with a CCD camera and then digitized. The average success rate of the neural network built (10-18-3 topology) was 90.9% (excellent: 91.7%; good: 89.1%; bad: 92.3% successful classification). In summary, it can be stated that infected nuts can be eliminated clearly using the measurement system built, reducing the risk of the disease.

Aroca-Santos et al. [46] used artificial neural networks for the identification and quantitative determination of olive oils. Using visible spectroscopy data, the composition of 4 varieties of extra virgin olive oil (80 samples) was determined. The input values of the feed forward network were the data obtained by spectroscopy. There were 4 neurons in the output layer, corresponding to the 4 varieties of oil. The identification value was 100%, and the error rate during quantitative determination was only 4.98%. In their research, it was found that ANN combined with spectroscopy is a faster and more economical method of identification than the commonly used GC-MS method.

The use of artificial neural networks (*ANN*) in the practice of sensory qualification was demonstrated by Sipos et al. (2012). In their work, they emphasize that only a few research results related to sensory tests have been published in the international literature, which may be because researchers use conventional statistical methods for their analyses, and novel methods are not known among them. They suggest in their work to apply artificial neural networks to monitor the performance of sensory judges (panel) [47].

5.3 Optimization applications in food science

The goal of Kono et al. [48] was to determine the optimum storage conditions of frozen cooked rice. The information required for the research was storage time (1, 5, 10, 30, 90 days) and storage temperature (-5, -15, -30, -40 °C), which were the input variables of the network. To measure the ice crystals, the first step was a washing with a fluorescent liquid (0.01 V/V% Rhodamine B solution). This was followed by freezing, illumination with a special lamp and then digital imaging. Images were processed to obtain the equivalent diameter of the crystals, which were the output data of the neural networks. In addition, a sensory test was also carried out, in which 690 consumer judges participated. The subject of the research was to determine what ice crystal size is associated with the best organoleptic properties. There was a strong correlation between overall flavor and the size of the ice crystals (R^2= 0.96, if the frozen rice is thawed at room temperature, and R^2= 0.93, if the sample is thawed in a microwave oven). In their work, they came to the conclusion that a crystal size below 13 μm is the most ideal for tastiness. The back propagation neural network had a 2-3-1 neuron structure and it determined the optimum storage conditions for ice crystal sizes: storage at -25 °C and

natural thawing or storage at -15 °C and microwave thawing were the parameters that provided the best organoleptic properties to the finished product.

6. Conclusions

In summary, it can be stated that artificial neural networks possess many beneficial properties, making them suitable for solving different tasks. By recognizing the complex relationships and patterns inherent in data, they can solve problems in the case of which traditional methods would not yield satisfactory results. High efficiency is achieved by artificial neural networks primarily in the modeling of non-linear trends, therefore, they are used for the modeling of various complex relationships, classification, categorization or for the regression of non-linear variables. Nowadays there are many sophisticated algorithms available for the training of neural networks, which can mean an alternative to conventional methods (linear regression, discriminant analysis). However, when using artificial neural networks, several factors need to be taken into consideration (objective of the task (prediction, classification, optimization), reliability of the data set, data set size, expected accuracy, etc.).

In our work, the foundations, characteristics, simplified operating principles and general structure of artificial neural networks were summarized. We also covered their widespread applicability and their use in food science experiments. Through various classification, prediction and optimization tests we gave an account of the effective use of neural networks, which were outstandingly successful in the solution mechanism. Taking into account current trends, it is assumed that the application of neural networks still has untapped potential, therefore, the number of scientific research projects using different ANNs is expected to increase in the coming years. The spread of robotics and artificial intelligence systems points to the strengthening of the role played by neural networks in the future.

7. Acknowledgement

8. References

[1] David, Vojtěch (2016): Izotropická frakcionace: metoda umožňující spolehlivý odhad procesní kapacity mozku.

[2] Basheer, Imad A., and M. Hajmeer (2000): Artificial neural networks: fundamentals, computing, design, and application. Journal of microbiological methods 43.1 3-31.

[3] Zupan, J., Gasteiger, J. (1991): Neural networks: A new method for solving chemical problems or just a passing phase? (a review). Anal. Chim. Acta 248, 1–30.

[4] McCulloch, W. S. and Pitts, W. H. (1943): A logical calculus of the ideas immanent in nervous activity. Bulletin of Mathematical Biophysics, 5, 115–137.

[5] Russell, Stuart, Peter Norvig, and Artificial Intelligence (1995): A modern approach. Artificial Intelligence. Prentice-Hall, Egnlewood Cliffs 25 27.

[6] Hebb, D.O. (1949): The organization of behavior: A neuropsychological theory. New York, John Wiley and Sons. 335.

[7] Rosenblatt, Frank (1958): The perceptron: A probabilistic model for information storage and organization in the brain. Psychological review 65.6 386.

[8] Hopfield, J. (1982): Neural networks and physical systems with emergent collective computational abilities, Proc. Natl. Acad. Sci. U.S.A., 79 (8) 2554–2558.

[9] Widrow, B. and Hoff, M.E. Jr. (1960): Adaptive Switching Circuits. IRE WESCON Convention Record. 96–104.

[10] 11 Balogun, W. A., Salami, M-J. E., Aibinu, A. M., Mustafah, Y. M., Isiaka. S. B.S. (2014): Mini Review: Artificial Neural Network Application on Fruit and Vegetables Quality Assessment. International Journal of Scientific & Engineering Research, Volume 5, Issue 6.

[11] Marini, Federico (2009): Artificial neural networks in foodstuff analyses: Trends and perspectives A review. Analytica Chimica Acta 635.2 121-131.

[12] Marini, F., Magrì, A. L., Bucci, R. (2007): Multilayer feed-forward artificial neural networks forclass modeling. Chemometrics and Intelligent Laboratory Systems, 88 (1), 118–124.

[13] Altrichter M., Horváth G., Pataki B., Strausz Gy., Takács G., Valyon. J. (2006): Neurális hálózatok, Budapest, Panem Kiadó. 1-413.

[14] Bishop, C.M. (1996): Neural Networks for Pattern Recognition, Clarendon Press, Oxford.

[15] Zupan, J., Gasteiger, J. (1999): Neural Networks in Chemistry and Drug Design, 2nd ed., Wiley VCH Ed., Weinheim, 1999.

[16] Dreyfus, G. (ed.) (2005): Neural Networks – Methodology and Applications, Springer, Heidelberg.

[17] 19 Braga, A. P., de Carvalho, A. C. P. L. F., and Ludermir, T. B. (2000): Redes Neurais Artificiais: Teoria e Aplicações. Ed. LTC, Rio de Janeiro.

[18] Maren, Alianna J. (1991): Neural networks for enhanced human-computer interactions. IEEE control systems 11.5 34-36.

[19] Pham, D. T., and E. Oztemel (1994): Control chart pattern recognition using learning vector quantization networks. The International Journal of Production Research 32.3 721-729.

[20] Hornik K. (1991): Approximation capabilities of multilayer feedforward networks, Neural Networks 4, 251-257.

[21] Hornik, K., Stinchcombe, M., White, H. (1989): Multilayer Feedforward Networks are Universal Approximators. Neural Networks. 2

[22] Hornik, M. Stinchcombe, H. (1990): WhiteUniversal approximation of an unknown mapping and its derivatives using multilayer feedforward networks. Neural Networks, 3, pp. 551-560.

[23] Anonymus (2017): http://python3.codes/neural-network-python-part-1-sigmoid-function-gradient-descent-backpropagation/ (Hozzáférés/Aquired: 02.02.2017)

[24] Zell, Andreas (1994): Simulation Neuronaler Netze [Simulation of Neural Networks] (in German) (1st ed.). Addison-Wesley. p. 73.

[25] Borosy, A. P. (2001): Mesterséges ideghálózatok. In: Horvay szerk. Sokváltozós adatelemzés (Kemometria). Budapest, Nemzeti tankönyvkiadó. 312–329.

[26] Buhmann, M.D. Ablowitz, M.J. (2003): Radial Basis Functions: Theory and Implementations, Cambridge University Press, Cambridge, 2003.

[27] Kumaraguru, P., V. R. Prakash, and M. Abraham. Performance Analysis of Radial Basis Function Neural Network for Internet Traffic Classification.

[28] Kohonen, T. (2001): Self Organizing Maps, 3rd ed., Springer, New York

[29] Marini, F., Zupan, J., Magri A.L. (2005): Class-modeling using Kohonen artificial neural networks. Anal. Chim. Acta, 544, 306–314.

[30] Melssen, W.J., Wehrens, R., Buydens, L.M.C. (2006): Supervised Kohonen networks for classification problems. Chemom. Intell. Lab. Syst., 83, 99–113.
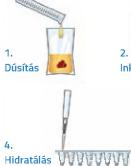
[31] Song, Jingge, et al. (2017): A globally enhanced general regression neural network for on-line multiple emissions prediction of utility boiler. Knowledge-Based Systems 118 4-14.

[32] Kusy, Maciej, and Piotr A. Kowalski (2018): Weighted probabilistic neural network. Information Sciences 430 65-76.

[33] Castro, Carla Aparecida de O., et al. (2017): High-performance prediction of macauba fruit biomass for agricultural and industrial purposes using Artificial Neural Networks. Industrial Crops and Products 108 806-813.

[34] Singh, R. R. B., et al. (2009): Prediction of sensory quality of UHT milk–a comparison of kinetic and neural network approaches. Journal of Food Engineering 92.2 146-151.

[35] Krishnamurthy, Raju, et al. (2007): Prediction of consumer liking from trained sensory panel information: Evaluation of neural networks. Food Quality and Preference 18.2 275-285.

[36] León-Roque, Noemí, et al. (2016): Prediction of fermentation index of cocoa beans (Theobroma cacao L.) based on color measurement and artificial neural networks. Talanta 161 31-39.

[37] Vásquez, Nadya, et al. (2018): Comparison between artificial neural network and partial least squares regression models for hardness modeling during the ripening process of Swiss-type cheese using spectral profiles. Journal of Food Engineering 219 8-15

[38] Bahramparvar, Maryam, Fakhreddin Salehi, and Seyed Razavi (2014): Predicting total acceptance of ice cream using artificial neural network. Journal of Food Processing and Preservation 38.3 1080-1088.

[39] Wang, Jing, Huili Yue, and Zenan Zhou (2017): An improved traceability system for food quality assurance and evaluation based on fuzzy classification and neural network. Food Control 79 363-370.

[40] Zarifneshat, S., Rohani, A., Ghassemzadeh, H. R., Sadeghi, M., Ahmadi, E., & Zarifneshat, M. (2012): Predictions of apple bruise volume using artificial neural network. Computers and Electronics in Agriculture, 82, pp. 75-86

[41] da Silva Sauthier, Maria Celeste, et al. (2018): Screening of Mangifera indica L. functional content using PCA and neural networks (ANN). Food Chemistry.

[42] Yu, Huichun, et al. (2008): Quality grade identification of green tea using E-nose by CA and ANN. LWT-Food Science and Technology 41.7 1268-1273.

[43] Anjos, Ofélia, et al. (2015): Neural networks applied to discriminate botanical origin of honeys. Food chemistry 175 128-136.

[44] Fadilah, Norasyikin, et al. (2012): Intelligent color vision system for ripeness classification of oil palm fresh fruit bunch. Sensors 12.10 14179-14195.

[45] Huang, Kuo-Yi. (2012): Detection and classification of areca nuts with machine vision. Computers & Mathematics with Applications 64.5 739-746.

[46] Aroca-Santos, Regina, et al. (2016): Neural networks applied to characterize blends containing refined and extra virgin olive oils. Talanta 161 304-308.

[47] Sipos L., Gere A., Kókai, Z., Szabó, D. (2012): Mesterséges ideghálózatok (ANN) alkalmazása az érzékszervi minősítés gyakorlatában. Élelmiszervizsgálati Közlemények. 58, 1-2, 32-45.

[48] Kono, Shinji, et al. (2015): Optimum storage temperature of frozen cooked rice predicted by ice crystal measurement, sensory evaluation and artificial neural network. International Journal of Refrigeration 56 165-172.

SCIENCE