Full Length Research Paper

A CFBPN artificial neural network model for educational qualitative data analyses: Example of students' attitudes based on Kellerts' typologies

Nurettin Yorek* and Ilker Ugulu

Faculty of Education, Dokuz Eylul University, Izmir, Turkey.

In this study, artificial neural networks are suggested as a model that can be 'trained' to yield qualitative results out of a huge amount of categorical data. It can be said that this is a new approach applied in educational qualitative data analysis. In this direction, a cascade-forward back-propagation neural network (CFBPN) model was developed to analyze categorical data for determine students' attitudes. The data were collected using a conceptual understanding test which includes open-ended questions. The results of this study indicate that using CFBPN model in analyzing data from educational research examining attitudes, behaviors, or beliefs may help us obtain more detailed information about the data analyzed and hence about the characteristics of the participants involved.

Key words: artificial neural networks; qualitative data analysis; education; biology

INTRODUCTION

Kellerts' attitudinal typologies

There have been many studies conducted on attitudes towards nature and living things. Studies by Stephen Kellert have special importance because they proposed a particular systematic for attitudes towards nature. By conducting comprehensive studies Kellert (1996) determined nine basic attitude types and showed that attitudes vary with respect to age (Table 1). Whereas children's relationship with animals is more of emotional at ages 6-9, they focus on learning about living things at ages 10-13. In later developmental stages ethical/ ecological issues emerge related to living things and their habitat.

Most of the studies conducted based on Kellert's typologies try to determine in which of the 9 categories participants fall. However, in some instances Kellert (1991) made some additions to the 9 basic attitudes. For instance, in a study related to Japanese society Kellert included naturalistic and theistic attitudes (Kellert, 1991). In some other cases, Kellert preferred to use attitudes with similar properties as grouped (such as naturalistic-ecologistic) (Kellert, 1993b).

Data analyses in educational researches

The majority of the data collected within the scope of

*Corresponding author. E-mail: Nurettin.yorek@deu.edu.tr.

Table 1. Kellerts' typologies of basic attitudes toward wildlife (according to Thompson and Mintzes, 2002; p.647).

| Term	Definition
Naturalistic	Primary focus an interest and affection for wildlife and the outdoors.
Ecologistic	Primary concern for the environment as a system, and for interrelationships between wildlife species and natural habitats.
Humanistic	Primary interest in and strong affection for individual animals such as pets or large wild animals with strong anthropomorphic associations.
Moralistic	Primary concern for the right and wrong treatment of animals, with strong ethical opposition to presumed overexploitation or cruelty toward animals.
Scientistic	Primary interest in the physical attributes and biological functioning of animals.
Aesthetic	Primary interest in the physical attractiveness and symbolic characteristics of animals.
Utilitarian	Primary interest in the practical value of animals.
Dominionistic	Primary interest in the mastery and control of animals, typically in sporting situations.
Negativistic	Primary orientation an active avoidance of animals due to dislike or fear.

researches conducted on social cases such as education rely on quantitative data. In this respect, most commonly used statistical analysis procedures are descriptive statistics, t-test, ANOVA/MANOVA, correlation, regression, and psychometric statistics (Hsu, 2005). The main reason for this is that quantitative data take less time to collect and analyze by using package software. However, results obtained from qualitative analyses provide more in-depth data on subjects and thus considered to be more 'valuable' for researchers. In this respect, it is considered that new methods to analyze a large amount of qualitative data in a short span of time with minimal loss are required.

   The main reason that encourages us to conduct this study is our persuasion that artificial neural networks have such a huge potential. Customized artificial neural network architectures and training algorithms specific to individual studies are considered to be used in the analyses of qualitative data. For example, in this study, artificial neural networks are suggested as a model that can be 'trained' to yield qualitative results out of a huge amount of categorical data.

Artificial neural networks

Artificial neural networks (ANNs) are mathematical models inspired by biological neural networks contained in human brain. Having similar characteristics to those of biological neural networks (i.e. consistency, flexibility, parallel function, and tolerance to errors, etc.), these systems attempt to learn tasks and determine how they will react to new tasks by means of creating their own experiences through the data obtained by using the predetermined samples (Sagiroglu et al., 2003).

   Neural networks can be used to model complex relationship without using simplifying assumptions, which are commonly used in linear approaches. The other
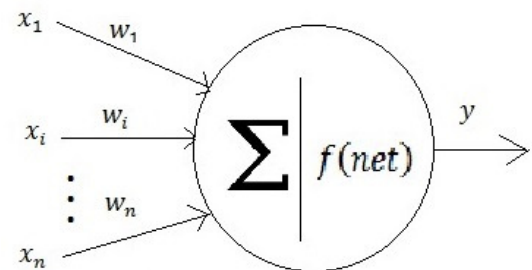
Figure 1. Representation of a neuron.

advantages of the ANNs are the ability to represent both linear and nonlinear relationships, the ability to learn these relationships directly from the data used, need not take into account a detailed information of structures and interactions in the systems, and they are regarded as ultimate black-box models. At least in some cases if not always, i.e. for prediction using the trained network, the ANN systems are alternative to experimentation and save a lot of time which may have been consumed since experimentation is so difficult and in some cases are impossible. Artificial neurons based on biological model were first defined by McCulloch and Pitts. McCulloch-Pitts (MCP) neuron model is given in Figure 1.

   In all neural network models, x_i input values are multiplied by w_i connection weights and then summed up. Summation unit is compatible with the body of biological neuron. It sums up weighted inputs and then gives the *net* output, such that:

$$net = w_1x_1 + w_2x_2 + ... + w_nx_n + x_0b = \sum_{i=1}^{n} w_ix_i + x_0b \qquad (1)$$
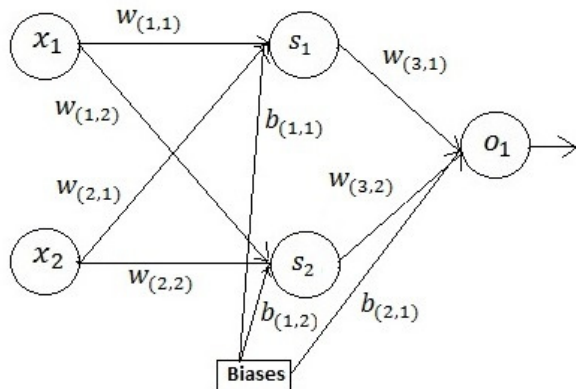
$$y = f(net)$$

Figure 2. Feed forward ANN (FFANN).

Input values (x_i) are multiplied by weights (w_i) assigned to connections and applied to additive function (Σ) together with a bias value ($x_0 * b$). Bias value is applied to neuron externally. Output is obtained by application of f activation function to additive function output.

Each input has an effect on output in proportion to weights assigned to connections, and threshold value is independent from system inputs. In case the values of all inputs are equal to 0, the output function is represented as $f(x_0 * b)$.

In case of biological neurons, neuron yields to output when input exceeds the activation value. In order to apply this feature in ANNs, an activation function, usually nonlinear, which produces cell output by means of processing net input obtained from additive function, is used. Various activation functions are applied based on the model used. Most commonly used functions are step functions (unipolar and bipolar), linear functions (standard linear, and symmetric piecewise linear), and sigmoid functions (logarithmic sigmoid and tangent sigmoid).

Structures of ANN

ANNs can be analyzed in two separate groups as single-layer and multilayer ANNs, which are determined based on the number of layers in their structures (i.e. network architectures). In a single-layer neural network, neurons represent output layer. Neurons receiving input values are not considered as input layer due to the fact that no calculations are made in this layer. Data received from input layer is calculated in output layer and network output is obtained.

On the other hand, multilayer networks are different from other networks, such that multilayer networks have one or more hidden layers and also weighting is applied

in input layer. In multilayer networks, at least one hidden layer (s_1, s_2) is represented between output layer (o_1) and input layer (x_1, x_2) (Fausett, 1994) (Figure 2).

Learning in ANNs

In a neural network, learning can be defined as reaching optimum weight values between neurons, which provides approximation between the output values calculated by output values versus a given input vector set and the expected output values.

A training algorithm is used for learning process and compositions of weights are determined by these algorithms. The objective of the learning process is to obtain an output value with a maximum approximation to the expected output by means of reducing errors using learning algorithms. For this purpose, weights in the system are iterated in each network with an aim to reduce errors. If artificial neural networks have achieved their goal with the input-output pairs, weight values are saved. The process during which weights are constantly iterated until the expected result is achieved is defined as "learning" (Lawrence et al., 1997).

Delta Rule, also applies to this study, is one of the most commonly used learning rules (Sagiroglu et al., 2003). Reducing the discrepancy between the expected output value and the predicted output value of neuron, this learning rule is based on the concept that strengthens and constantly changes input connections. This rule is based on the principle of reducing mean square error by means of changing weight values of the connections and it attempts to reduce errors by means of back propagation from output layer towards input layer. Therefore, Delta Rule is also called back propagation or least mean square learning rule.

Feed forward back-propagation ANNs (FFANN)

In feed forward ANNs, one layer contains some neurons which are connected to those of the following layer. Each connection is weighed. A neuron is described with its own activation level, which is responsible for the propagation of the information from the input layer to the output layer. However, to obtain reliable weights, the neural network must learn about the known input- and output-samples. During the learning process, an error between theoretical and experimental outputs is computed. Thus, the weight-values are modified through an error back propagation process which is executed on several sampling data, until achieving as small error as possible. After this last step, the neural network can be considered as trained and able to be used in calculating other responses to new entries that have never been

presented to the network. It is important to emphasize that the learning speed of the neural network depends not only on the architecture but also on the algorithm used (Gallant, 1993; Guney, 1997).

Back-propagation algorithm

Back propagation of errors learning model, first introduced by Rumelhart (1986), is one of the most commonly used models amongst other artificial neural network learning models (Rumelhart, 1986). In back propagation algorithm, learning mechanism is based on iterative gradient descent method which minimizes errors between the expected outputs and the predicted outputs of the network. In learning rule, error calculated in network output is used in the calculation of new values of the weights. Supposing that y_i represents output value of the ith neuron in the output of artificial neural network after k times of iteration of the training, d_i represents the expected value, and e_i represents the error signal of the i neuron, then the calculation of error value is defined by the following Equation:

$$e_i = \left(d_i - y_i\left(k\right)\right) \tag{2}$$

When an input data is applied to a network, various processes are performed on this data until it reaches output layer. Output obtained as a result of these processes is compared to the expected output and approximation function is defined by the following Equation:

$$E = \frac{1}{2}\sum_i \left(e_i(k)\right)^2 = \frac{1}{2}\sum_i \left(d_i - y_i(k)\right)^2 \tag{3}$$

The difference between the calculated values and the expected values is calculated as an error signal for each output node. Based on these error signals, connection weights are rearranged for each neuron. This arrangement allows for convergence of the network to a condition where all data can be coded, and the gradient of weight values is determined by the method of the steepest falling gradient (Rumelhart, 1986), which can be represented by the following Equation:

$$\Delta w_{ij} = -\eta \frac{\partial E\left(\vec{w}\right)}{\partial w_{ij}} \tag{4}$$

In the above Equation, η is coefficient of learning. Each iteration process in back propagation algorithm consists of two stages as forward propagation and back propagation. During forward propagation, output values of ANN versus input signals applied to ANN at that time are determined. During back propagation, the previously assigned weights are rearranged on the basis of resulting output errors. Each change of weight in ANN is performed based on the following equation:

$$\Delta w_j = \eta \delta_j y_i \tag{5}$$

For neurons in output layer, δ_j is defined as

$$\delta_j = e_j\left(k\right) f_j' \tag{6}$$

Whereas, for neurons in hidden layers, it is defined as

$$\delta_j = f_j' \sum_m \delta_m w_{mj} \tag{7}$$

f_j is the activation function of j neuron. By these definitions, the flow of error signals from output towards input is considered to be similar to the flow of signals forward during forward propagation. Iteration process continues until the error value is reduced to a certain level and therefore training process of the network is completed. Weight values of the connections between layers are obtained from the network upon completion of its training and these values are stored to be used during test process (Yao 1999).

METHOD

Qualitative data obtained from student answers to open-ended questions were used to train and test the ANN model. 80% of this data was used for training of the network and the remaining 20% was used for testing of the network (Hagan et al., 1996). Detailed information and algorithms of ANNs are explained in above section 1.3.

Subjects

The participants included 214 students (127 female and 87 male) who were selected via cluster sampling method (Bogdan and Biklen, 2006) from eight high schools in Izmir, a large city in western Turkey. Schools accepted students from different parts of the city and students varied in terms of socioeconomic status.

Data collection

In this study a conceptual understanding test was used. The test included open-ended questions and was developed by researchers. In addition, to clarify vague concepts and to obtain in-depth information about the topics interviews were conducted with students and teachers. The final version of the test used in this study is presented in Table 2.

Table 2. Conceptual understanding test of the living things.

1. It is estimated that there are millions of species living on Earth. If you were asked to classify all the living things (types, species) into main groups, without leaving anyone out, *at least* how many groups can you from?
2. When all the living things were considered, what do you think is the biological position of *human*? Explain.
3. When all the living things were considered, in your opinion is(are) there any living thing(s), existence of which is(are) unimportant (to have little or no use)? If yes, which ones? If no, why? Explain your reason.
4. When you rank the following names of the living things from the most significant to the least, according to your criteria of importance, which one ranks first? How did you determine the level of importance? Explain. Rat, nettle, mushroom, honeybee, daisy.

Table 3. Student rankings of the living things in terms of significance.

Students	Rank1	Rank2	Rank3	Rank4	Rank5	Attitudes
X17	Rat	Honeybee	Daisy	Mushroom	Nettle	Humanistic
X20	Mushroom	Rat	Nettle	Honeybee	Daisy	Utilitarian
X21	Honeybee	Rat	Nettle	Mushroom	Daisy	Ecologistic
X27	Honeybee	Rat	Nettle	Daisy	Mushroom	Utilitarian
X28	Mushroom	Rat	Daisy	Nettle	Honeybee	Utilitarian
X29	Honeybee	Rat	Mushroom	Nettle	Daisy	Ecologistic
X147	Daisy	Mushroom	Honeybee	Nettle	Rat	Utilitarian
X148	Honeybee	Mushroom	Daisy	Nettle	Rat	Utilitarian
X149	Daisy	Honeybee	Mushroom	Nettle	Rat	Ecologistic
X150	Daisy	Honeybee	Mushroom	Nettle	Rat	Scientistic

Data analyses

The fourth question in the conceptual understanding test was used to train and test our artificial neural network. Students answer this question at two stages. First, they were asked to list the names of 5 living things according to importance. Then, they were asked to explain the criteria they used to write the name of the first living thing as the first on the list.

Students' answers to the 4th question evaluated along with their answers to other questions and their attitudes were tried to determine based on Kellert's typologies. For example, let's assume that two students list the names of the living things as the same and both wrote the honeybee first. Let us further assume that one student's reason to write honeybee first was that "honeybees are living things that have important functions in nature." The reason put forth by the other student was that "honeybees make honey for us." In this case, according to Kellert's typologies first student can be characterized as 'ecologistic' and the second as 'utilitarian.' All students' answers were analyzed in this way and the data were tabulated. Some data are presented in Table 3.

Creating CFBPN model

In this study, we developed a cascade-forward back-propagation network (CFBPN) model. The CFBPN model is structurally very similar to the FFANN model. Every neuron at input and hidden layers are connected to each other. In addition, all input layer neurons and output layer neurons have direct connections with each other. While the hidden layer takes data only from the input layer, the output layer takes data from both the input layer and the

hidden layer. According to Filik and Kurban (2007) the fact that input layer (independent variables) and the output layer (dependent variables) is connected provides CFBPN model some advantages over the FFANN model in some cases.

To test the proposed ANN model collected data were divided into two groups as training data and test data. Training data were used to develop the ANN model. Test data were not used in training, they were used to verify and test the ANN model.

The model started training with the randomly chosen weight matrix. Then, results from the output layer compared to expected results and a back propagation error value is defined. This error value is back propagated in the network and weights were rearranged. This process continued until there is minimum error value or there is no change in weights. In addition, number of neurons hidden layer should have was determined to be able to obtain appropriate results. There is not a definite method of determining the number of neurons hidden layer should have. It has been determined through trial and error depending on researcher experience. The number of neurons in input and output layers is determined according to number of dependent and independent variables. Since listing names of five living things is the independent variable, there are 25 neurons in the input layer to represent this listing. Since there are four attitudinal typology types determined, there are four neurons in the output layer to represent this. Table 4 shows how the data in Table 3 were coded to train our network.

As the set of species contained in the research problem is given to the students in the following sequence as "rat, nettle, mushroom, honey bee, and daisy", the same sequence has also been used in coding. For each species, a vector containing four '0' and one '1' has been used. Therefore, the code used by a student to list these five species is a vector consisting of 25 bits. For example, when the

Table 4. Input vector data coding for ANN model.

Living things	Rat	Nettle	Mushroom	Honey bee	Daisy
Codes	10000	01000	00100	00010	00001

Table 5. Output vector data coding for ANN model.

Attitudes	Codes
Ecologistic (E)	1000
Utilitarian (U)	0100
Humanistic (H)	0010
Scientistic (S)	0001
E/U	1100
E/H	1010
E/S	1001
U/H	0110
U/S	0101
H/S	0011
E/U/H	1110
E/U/S	1101
E/H/S	1011
U/H/S	0111
E/U/H/S	1111

following sequence as 'nettle-honey bee-daisy-mushroom-rat' is coded, (01000) in the first order for nettle, (00010) in the second order for honey bee, (00001) in the third order for daisy, (00100) in the fourth order for mushroom, and (10000) in the fifth order for rat are entered, respectively. In other words, the complete sequence is coded as the following: 0100000010000010010010000. Written in vector notation, it is the input vector: (0, 1, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 1, 0, 0, 1, 0, 0, 1, 0, 0, 0, 0).

A similar method has been used in coding attitudes (Table 5.). (**1**000) for ecologist, (0**1**00) for utilitarian, (00**1**0) for humanistic, and (000**1**) for scientific have been used, respectively. However, a coding method inspired by the studies on rough set conducted by Narli *et al.* (2010) has been developed for the students who list the species in the same order, however having different attitude characteristics. Therefore, codes have been used in combinations in case of two or more attitudes. For example, the code (**11**00) has been used in case of ecologist-utilitarian, and the code (1011) has been used in case of ecologist-humanistic-scientific.

In the next stage, the transfer function for each stage must be determined by trial and error approach. In this regards, different types of transfer functions including logarithmic sigmoid, hyper-bolic tangent sigmoid, linear, and radial basis transfer functions were used to find the proper transfer function for the proposed neural network (Hagan et al., 1996).

Next, we determined transfer functions necessary for every stage. To determine the most appropriate transfer functions for our model, logarithmic sigmoid (logsig), hyperbolic tangent sigmoid (tansig), linear, and radial basis transfer functions were tried. As a result, logsig (Eq. (8)) and tansig (Eq. (9)) functions were decided as appropriate for the hidden layer and output layer respectively

(Figure 3.).

$$f(net) = \frac{1}{1 + e^{-net}} \tag{8}$$

$$f(net) = \frac{1 - e^{-2net}}{1 + e^{-2net}} \tag{9}$$

The CFBPN model used in our study has 25 neurons in input layer, 10 neurons in hidden layer and four neurons in output layer. The model is shown schematically in Figure 4.

As explained earlier, training stage is one of the important steps. In general, back propagation algorithm is used for training. We have used Levenberg-Marquardt algorithm in the training phase of the present study. Convergence rate of back propagation algorithm is low and its risk to find a local minimum is quite high. Levenberg-Marquardt Algorithm (LM) is generally used for problems vulnerable to such risks. Whereas back propagation algorithm (BP) attempts to reduce errors by use of first order derivatives, LM interpolates between Newton method and BP method, and then it attempts to reduce errors by use of second order derivatives (Hagan and Menhaj, 1994; Rumelhart and McClelland, 1986). LM method is a damped least-squares method based on the concept of maximum neighborhood (Levenberg, 1944).

In the next step, we determined the number of hidden layers. It was reported that networks having one hidden layer are appropriate for nonlinear approaches (Cybenko, 1989). Our model thus includes only one hidden layer.

RESULTS AND DISCUSSION

First, 184 cases were used for training of the network and 30 cases were used for testing the network. According to the results obtained via trial and error approach, optimum number of neurons that hidden layer should have was determined as 10. Therefore, to have a minimum error our network model should have a network architecture as 25x10x4 (Figure 5). The activation (transfer) functions used in hidden and output layers are logsig and tansig respectively.

In addition, the synaptic parameters including weights and biases are given in Tables 6, 7 and 8 which enable any one to reproduce every used data points in the present study.

The network was trained until optimum parameters are obtained (more than 1000 epoch). Then, it was tested using test data. The training and test phases are displayed in Figures 6 and 7 respectively. As a result, there is great consistency between the results our model produced and the expected results.
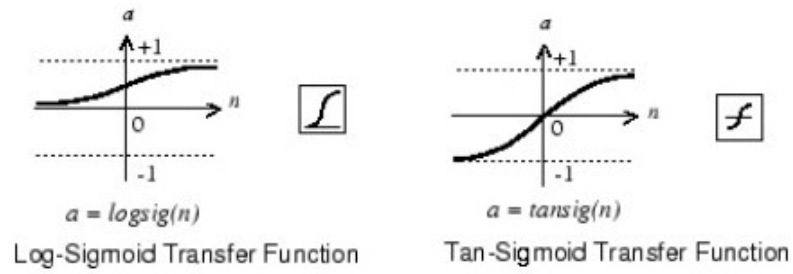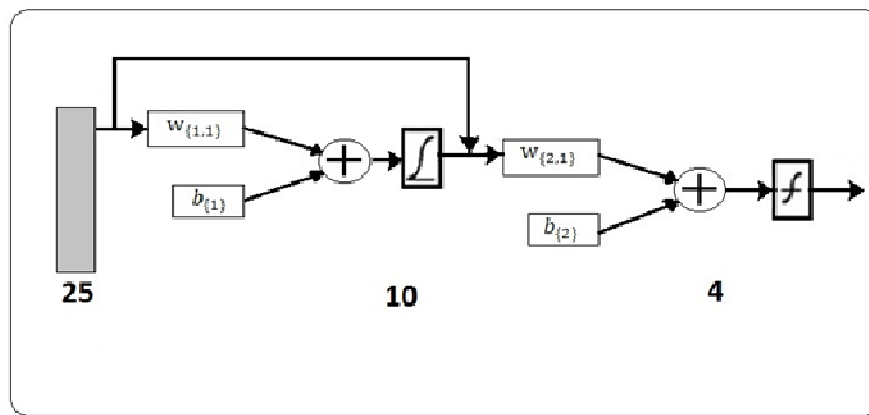
Figure 3. Transfer functions.

Figure 4. The schematic representation of CFBPN model.
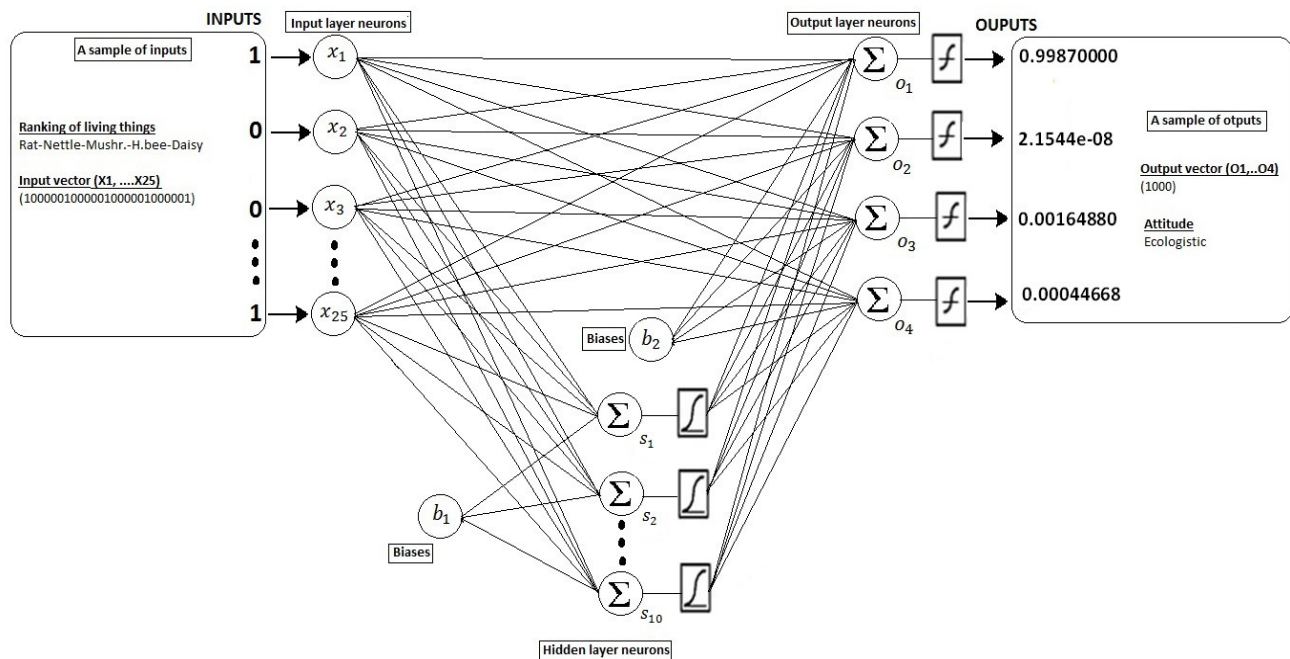

Figure 5. Trainable CFBPN artificial neural network architecture.

Table 6. Connection weights matrix from the input layer to hidden layer (W_{ij}) and biases

		S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}
						Hidden layer neurons					
Inputs	Rank1 $(X_1,..X_5)$	0.78864	1.4503	-1.8079	1.452	0.5519	-0.35577	-1.2072	0.037516	3.1423	2.1863
		-0.76022	0.13729	0.63891	1.0559	2.3707	-3.2429	-0.25737	2.3275	-0.54864	0.38329
		-0.95737	-0.17876	1.6528	0.58333	-1.0237	0.97174	2.6421	-1.1106	0.58025	-1.954
		0.36849	-2.1475	-2.6011	-3.455	-1.0759	1.9243	1.7404	-2.016	1.5536	-2.6801
		1.3728	-1.1545	2.0002	1.624	-0.16809	0.13441	-1.5541	0.47475	-0.35286	0.98813
	Rank2 $(X_6,..X_{10})$	-0.57162	2.3206	-1.1575	-0.52872	-3.724	0.37781	2.0017	-1.8524	-1.2134	-1.6003
		1.8729	-0.56485	0.72145	0.9362	0.88649	2.4606	0.92906	-2.8851	-2.1532	-0.10247
		0.44667	-3.3594	2.4802	-3.0291	3.0585	-1.0951	-4.4233	-2.0221	1.8581	-1.6579
		1.6994	1.1755	-1.7374	0.64896	2.2722	0.6353	-0.14444	2.2274	-1.5478	2.447
		-3.756	-2.1113	0.47761	1.9744	1.1089	-1.835	2.713	2.1359	-1.267	-2.4429
	Rank3 $(X_{11},..X_{15})$	-2.9365	0.65099	1.1628	-1.2544	1.0925	0.23389	-0.59234	0.33941	-0.68414	1.2433
		-0.50569	-3.4599	-0. 1137	0.96792	-3.6886	1.6913	-0.23712	1.1109	-2.4188	1.0576
		-2.65	1.5159	-4.3883	1.8249	2.8767	0.9495	-2.1025	-0.0026772	-0.90547	0.15393
		1.3603	1.6502	1.0117	0.13681	-0.41837	-1.5426	2.3729	1.0265	1.5974	-2.7632
		3.7646	-0.36254	-1.8205	-0.65485	-0.37411	-0.87901	-0.26321	-2.9487	-0.19173	-1.9
	Rank4 $(X_{16},..X_{20})$	2.0695	-1.9722	1.0685	0.69426	1.2338	0.22989	-0.37213	2.4683	-1.5162	-2.4725
		-0.71868	-0.98477	-3.4045	-0.88019	0.61636	-0.39837	2.6059	1.4761	0.29735	-1.5167
		1.0083	-0.60972	0.78559	-1.562	-2.64	-2.8013	1.0831	-0.95471	-1.6036	0.73547
		-3.0347	-3.6995	2.5029	0.0010816	1.2943	-0.20637	-2.1829	-0.044698	1.9217	-1.9484
		1.7884	4.4183	-0.24413	0.70473	-0.86411	1.2632	0.083992	-1.1271	1.5847	1.0705
	Rank5 $(X_{21},..X_{25})$	0.39689	-3.5193	0.55003	0.8385	0.47616	0.19385	1.1589	-0. 90451	1.0328	-0.99713
		-0.065859	2.2805	1.0507	-1.5964	1.3658	-0.54835	-2.1511	-1.5024	2.362	-2.5135
		1.5188	1.0353	-0.45812	1.4767	0.3819	3.616	2.2732	1.5689	-0.97686	0.35713
		0.081028	1.054	-0.31353	1.1559	-1.8619	-2.2108	-0.11367	-2.041	-3.0828	0.45892
		-3.2666	-3.2021	-0.30518	-3.5333	2.0617	0.27449	-1.1567	-0.011975	-1.9181	0.53516
	Biases (b_1)	-3.1279	-1.8378	2.0019	-1.0257	-0.68062	-0.4867	-1.1728	2.0629	2.689	3.7156

Conclusion

The present study used a CFBPN model approach to an educational research involving qualitative data. It can be said that this is a new approach applied in educational qualitative data analysis. Consequently, the present study has attempted to show the applicability of ANNs in detailed analyses of educational data. The ANNs are utilized mostly

Table 7. Weights connection matrix from the input layer to output layer (Wik) and biases

		Output layer neurons			
		O_1	O_2	O_3	O_4
Inputs	Rank1 $(X_1,..X_5)$	4.5734	-5.3634	0.38563	-4.4941
		-0.66081	-0.91009	-1.3472	0.72317
		0.68001	5.9236	-0.36622	-0.51813
		3.4679	1.9517	-3.2882	-0.45556
		2.4582	4.3011	2.1037	-1.5538
	Rank2 $(X_6,..X_{10})$	-1.1156	-1.9907	-3.1977	2.0466
		-0.17552	0.82335	-0.57041	-0.86524
		2.1132	-2.9039	-0.6195	-1.8202
		0.74707	-0.87652	0.59601	1.9969
		1.2532	1.179	0.56482	-0.41237
	Rank3 $(X_{11},..X_{15})$	0.041205	3.6995	-0.19798	0.51898
		-2.8629	2.4737	3.0117	-2.9654
		-1.0726	-0.27359	-2.5616	-0.19829
		2.221	-3.7873	-1.089	0.92936
		1.9347	-2.4868	-1.3744	-1.9162
	Rank4 $(X_{16},..X_{20})$	1.7668	0.025732	-0.49382	-0.032302
		-0.14938	-0.58613	-0.75787	-0.96644
		3.9681	-1.0538	1.9719	0.75513
		1.5356	2.0224	-0.47041	-4.9275
		-0.40513	-0.27916	-3.1219	2.6014
	Rank5 $(X_{21},..X_{25})$	0.88078	3.8167	1.0816	-0.80005
		4.6141	-1.735	-2.4077	1.1337
		0.3151	-1.7403	-2.0866	-0.15404
		-1.7809	1.327	1.0452	-1.0701
		1.4228	-1.3297	-0.70889	-1.1441
	Biases (b_1)	-4.6917	-0.47138	2.4025	-0.041698

Table 8. Weights connection matrix from the hidden layer to output layer (Wjk)

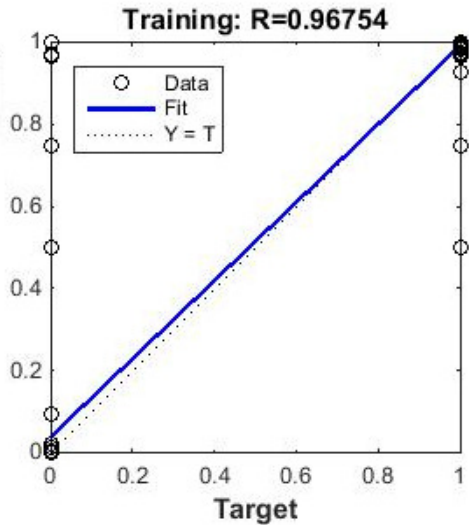|		Output layer neurons			
		O_1	O_2	O_3	O_4
Hidden layer neurons	S_1	-5.0116	8.3042	-3.015	-7.1749
	S_2	8.4891	-0.74355	8.4183	-11.4488
	S_3	-8.5945	-7.6702	-10.2047	-0.002188
	S_4	1.8644	-6.3863	-3.8443	8.91
	S_5	-0.74532	7.5856	-4.2982	9.861
	S_6	4.0926	-7.7658	0.92315	4.1332
	S_7	9.1429	-1.5493	-10.8214	0.060434
	S_8	6.047	6.6244	-3.89	-4.818
	S_9	8.5302	3.1134	7.4547	-3.849
	S_{10}	8.815	0.28353	-11.7286	-4.7165

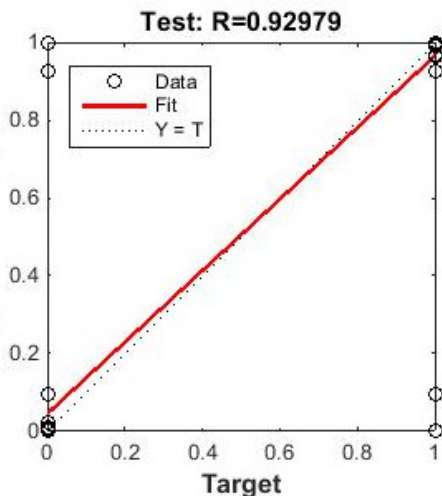**Figure 6.** Predicted attitudes vs. expected attitudes for training data set.

Figure 7. Predicted attitudes vs. expected attitudes for test data set.

in areas such as artificial intelligence (AI), machine learning, pattern recognition, decision support systems, expert systems, data analysis, and data mining.

Kellert (1993a, b) has represented some typologies in binary groups (e.g., utilitarian–dominionistic). However, these groupings occur between typologies that display very similar characteristics. Narli et al. (2010) argued that there may be students who have intermediate attitudes among the typologies identified by Kellert. In this study, it

has been showed that CFBPN model can be trained to uncover intermediate attitudes. Moreover, this can be done in much shorter time in this way.

Consequently, the results of this study indicate that using CFBPN model in analyzing data from educational research examining attitudes, behaviors, or beliefs may help us obtain more detailed information about the data analyzed and hence about the characteristics of the participants involved. In addition, precise knowledge about our students' attitudes and beliefs may prove helpful in curricular and instructional studies.

Conflict of Interests

The authors have not declared any conflicts of interest.

REFERENCES

Bogdan R, Biklen SK (2006). Qualitative Research for Education: An Introduction to Theories and Methods (5th Ed.). Pearson, New York

Cybenko GV (1989). Approximation by superpositions of a sigmoidal function. Math. Control Signal 2:303–314.

Fausett L (1994). Fundamentals of Neural Networks. Prentice-Hall, New Jersey.

Filik UB, Kurban M (2007). A new approach for the short-term load forecasting with autoregressive and artificial neural network models, Int. J. Comput. Int. Res. 3:66–71.

Gallant S (1993). Neural Network Learning and Expert systems, MIT Press, Cambridge.

Guney K (1997). An introduction to neural networks. UCL Press, London.

Hagan M, Menhaj M (1994). Training feed forward networks with the marquardt algorithm. In Proceedings of IEEE Transactions on Neural Networks. 6:989.

Hagan MT, Demuth HB, Beale M (1996). Neural network design. PWS Pub., Boston

Hsu T (2005). Research methods and data analysis procedures used by educational researchers. Int. J. Res. Met. Educ. 28:109-133.

Kellert SR (1991). Japanese perceptions of wildlife. Conserv. Biol. 5:297-308.

Kellert SR (1993a). The biological basis for human values of nature. In: Kellert SR, Wilson EO (eds) The biophilia hypothesis. pp. 42–69. Island Press, Washington DC

Kellert SR (1993b) Attitudes, knowledge, and behavior toward wildlife among the industrial superpowers: United States, Japan, and Germany. J. Soc. Issues 49:53-69.

Kellert SR (1996). The value of life: biological diversity and human society. Island Press, Washington DC

Lawrence S et al. (1997) Lessons in neural network training: Training may be harder than expected, In Proceedings of National Conference on Artificial Intelligence, AAAI-97, pp.540–545, Menlo Park, California

Levenberg K (1944). A method for the solution of certain nonlinear problems in least squares, Quart. Appl. Math. 2:164-168.

Narli S, Yorek N, Sahin M Usak M (2010). Can we make definite categorization of student attitudes? A rough set approach to investigate students' implicit attitudinal typologies toward living things, J. Sci. Educ. Tech. 19:456-469.

Rumelhart D (1986). Learning representations by backpropagation errors. Nature 323:533–536.

Rumelhart D, McClelland J (1986). Parallel Distributed Processing. MIT Press, Cambridge

Sagiroglu S et al. (2003). Muhendislikte Yapay Zeka Uygulamalari-I: Yapay Sinir Aglari (in Turkish), Ufuk Yayincilik.

Thompson TL, Mintzes JJ (2002). Cognitive structure and the affective domain: On knowing and feeling in biology. Int. J. Sci. Educ. 24:645-660.

Yao X (1999). Evolving artificial neural networks. In Proceeedings of the IEEE, 87:1423–1447.

Yorek N, Narli S (2009). Modeling of cognitive structure of uncertain scientific concepts using fuzzy-rough sets and intuitionistic fuzzy sets: Example of the life concept, Int. J. Uncertain. fuzz. 17:747-769.