ISSN: 1309 – 6575
Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi
Journal of Measurement and Evaluation in Education and Psychology
Research Article; 2025; 16(1); 1-12

Investigating the Performance of Artificial Neural Networks in Predicting Affective Responses

İzzettin AYDOĞAN * Osman TAT **

Abstract

In this study it is aimed to examine the performance of an artificial neural network trained using items reflecting a latent trait in predicting responses to an item reflecting the same trait. This latent trait is the awareness of being able to communicate with people from different cultures, which is included in the PISA 2018 assessment. Relevant scale items were used as research variables. In addition to determining the extent to which the predicted responses overlap with the actual responses by analyzing the artificial neural network models, it was examined how the predicted responses affect the assumed latent construct and the reliability of the responses. Thus, the performance of artificial neural networks in predicting responses to affective items was evaluated. The responses expected from individuals for the items examined overlap with the responses given by individuals at a relatively moderate. However, it is observed that although the prediction values partially weaken the model fit indices, they still manage to keep them strong. In addition, the predicted values improved the factor loadings and the variance explained for the latent trait. Similarly, it is noticed that the predicted values also positively affect the reliability.

Keywords: Artificial neural networks, machine learning, affective responses, prediction

Introduction

Current advances in machine learning and artificial intelligence are largely driven by artificial neural networks (ANNs) (Goel et al., 2023). The ability of ANNs to analyze complex data, especially those that cannot be simplified by traditional statistical methods, is gradually improving (Tu, 1996). ANNs consist of structures in which neurons are connected to each other by synapses with adjustable weights. Synapses connecting neurons, which are the building blocks of ANNs, function in communication. Information exchange between neurons takes place through synapses. Information flows from the synapse of one neuron to the dendrite of another neuron (Goel et al., 2023). The fact that the weights are adjustable allows the network to be trained by back-propagating the errors throughout the network. The aim of training is to adjust the weights to minimize the error between actual and predicted values (Lillicrap et al., 2020). ANNs have attracted much attention due to their ability to model non-linear relationships between variables (Tu, 1996). Although it is seen as a simple variant, ANNs are biologically similar to the working principles of the human brain (Hasson et al., 2020).

ANNs exhibit successful performances in many important fields such as health, climate, physics, chemistry, biology, engineering, industry, agriculture (Lau et al., 2019; Park et al., 2020). Considering the purpose and frequency of use of ANNs, it is possible to say that they are mostly used for diagnosis, prediction and forecasting. It is widely used in areas such as predicting some features through some predictors with regression logic, missing data assignment, recognition, and classification. Although its application area in educational research is limited (Tu, 1996), studies conducted in relation to education and psychology contents (Aybek & Okur, 2018; Aydoğan & Zırhlıoğlu, 2018; Aydoğdu, 2020; Al-Saleem et al., 2015; Chavez et al, 2023; Flitman, 1997; Guarín et al., 2015; Huang & Fang, 2012; Lau et al., 2019; Rodríguez-Hernández, 2021; Shahiri & Husain, 2015; Umar, 2019; Zacharis, 2016) especially focus on predicting students' academic performance. The basic logic of ANN models

* Assoc. Prof.,Van Yüzüncü Yıl University, Faculty of Education, Van-Türkiye, izettinaydogan@yyu.edu.tr, ORCID ID: 0000-0002-5908-1285
**Assoc. Prof.,Van Yüzüncü Yıl University, Faculty of Education, Van-Türkiye, osmantat@yyu.edu.tr, ORCID ID: 0000-0003-2950-9647

To cite this article:
Aydoğan,İ. ,Tat,O., (2024). Investigating the performance of artificial neural networks in predicting. *Journal of Measurement and Evaluation in Education and Psychology, 16*(1), 1-12. https://doi.org/10.21031/epod.1525454
Received: 31.07.2024
Accepted: 19.11.2024

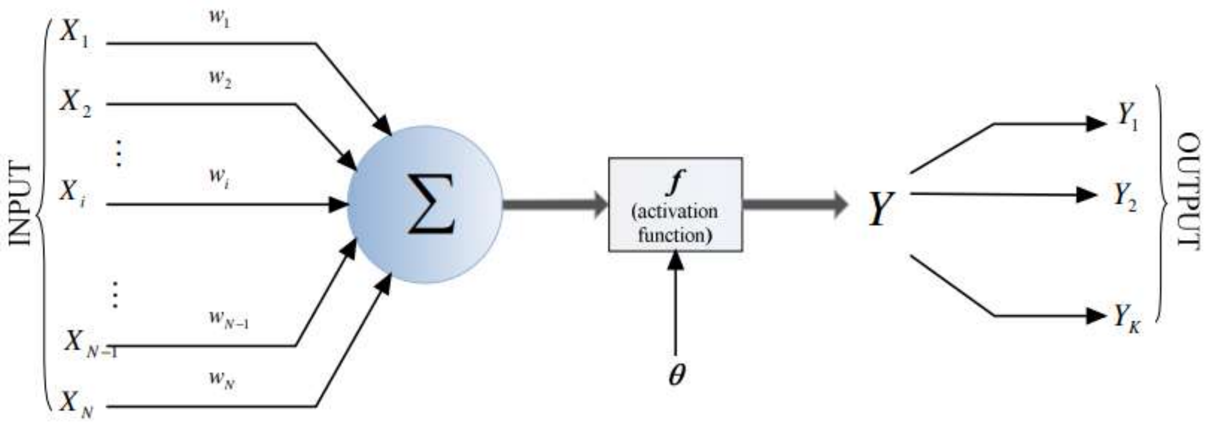
designed in these studies is to make predictions about students' cognitive performance through a number of covariate characteristics such as gender, parent's occupation, socio-economic status, etc. However, in this study, we aim to examine the performance of a network trained for a unidimensional scale, that is, using items that reflect the same latent trait, in predicting responses to another item with the same trait. In other words, using Programme for International Student Assessment (PISA) participants selected from the Lebanese sample, we examine how ANNs trained with responses to items on a scale predict responses to another item that is also part of the same scale. This provides an opportunity to compare the expected and observed values of the responses to a set of items with the same emotional integrity for the ANN trained based on machine learning. Therefore, in addition to determining the extent to which the responses predicted by ANNs overlap with the actual responses, we plan to monitor how the predicted items affect the latent construct and the reliability of the responses. In this context, we aim to evaluate the performance of the ANN method by performing similarity, validity and reliability analyses for actual and predicted responses. The focus here is to answer the question of to what extent we can accurately predict students' responses to another item that is part of the same emotion based on their emotional integrity, or to what extent the expected responses of individuals to a question posed within the same emotional integrity overlap with their responses. The reason for choosing the Lebanese sample is that Lebanon is a society where the emotional state reflected by the implicit feature in the scale we used is strongly experienced. We used the items of the student's intercultural communicative awareness scale administered in PISA 2018 as research data. The latent trait in the scale is the awareness of being able to communicate with people from different cultures. In 2018, Lebanon ranked first among the world countries in terms of the number of refugees per capita (McCarthy, 2019). The number of refugees per thousand inhabitants in Lebanon was 156 in 2018, which is more than twice the number in the second ranked country.

Artificial Neural Networks

ANNs are models that realize the features taught in the training phase through artificial neurons similar to the neuron structure in the human brain based on the principle of continuous improvement (Kose & Arslan, 2017; Vandamme et al., 2007). It has a complex and powerful structure that models non-linear relationships (Kardan et al., 2013). ANNs consist of three layers called input, hidden and output. The information transferred from the input data to the neurons in the input layer is processed through an aggregation function taking into account the weight values and transmitted to the activation function (See Figure 1).

Figure 1.

An Artifical Neuron (Grosan & Abraham, 2011, p.283).



The activation function is the component where calculations are made for the most accurate output. The information coming from the aggregation function is processed here to generate output values and transmitted to the output neurons (Rashid & Ahmad, 2016; Vandamme et al., 2007). In this process, the weight values are constantly adjusted to provide the best output. If the activation values reach the threshold value during the current iterations, the training phase is terminated and the network has learned. In this phase, new examples are shown to the network to test the learning. After the training phase is completed, the weight values remain constant. In this way, it is ensured that the learned network produces output using the current weight values (Öztemel, 2003).

Methods

Data and Participants

The research data were obtained from the PISA 2018 assessment (https://www.oecd.org/pisa/data/2018database/). PISA is a monitoring and assessment program implemented by the Organisation for Economic Co-operationand Development (OECD) for fifteen-year-old students in many countries around the world. Measures such as demographic information about students and their families, learning environments, information and communication technologies (ICT) and financial competence, as well as affective and cognitive measuresare provided. Thanks to the measurements applied and the results obtained accordingly, it provides the opportunity for countries to evaluate their own education systems and to review the educational outcomes of other countries. In this context, PISA provides important data and results to educators, researchers and administrators in terms of monitoring and evaluating educational processes (OECD, 2018).
The research group consists of Lebanese students who participated in the PISA 2018 assessment. 5614 Lebanese students participated in the PISA 2018 assessment. However, in order to clean the missing data in the data set and to meet the assumptions of the analysis techniques used in the research, some data were deleted and the research was conducted with the remaining 4631 student data.

Variables

The variables of the study consisted of seven items of the student's intercultural communicative awareness (*Awacom*) scale, which PISA officials stated as a part of the global competence domain. *Awacom* includes items that measure individuals' awareness of communicating with people from different cultures (See Table 1).

Tablo 1.
Items of Student's Intercultural Communicative Awareness Scale

Items label	PISA codes	Items
Item1	ST218Q01HA	I carefully observe their reactions.
Item2	ST218Q02HA	I frequently check that we are understanding each other correctly.
Item3	ST218Q03HA	I listen carefully to what they say.
Item4	ST218Q04HA	I choose my words carefully.
Item5	ST218Q05HA	I give concrete examples to explain my ideas.
Item6	ST218Q06HA	I explain things very carefully.
Item7	ST218Q07HA	If there is a problem with communication, I find ways around it (e.g. by using gestures, re-explaining, writing etc.).

The data were obtained through PISA student questionnaires. The scale items are in a four-point Likert response format: strongly disagree-disagree-agree-agree-strongly agree.

Data Preprocessing

The data used in the study were derived from PISA 2018 Lebanese sample data. The Lebanese sample consists of 5614 students; however, since confirmatory factor analysis (CFA), which is a member of structural equation models (SEM), and artificial neural networks (ANN) techniques used in the analysis processes are affected by missing (Ennett et al., 2001; Tabachnick & Fidell, 2019) and extreme (Tabachnick & Fidell, 2019) values, data with these characteristics were deleted by list-based data deletion method. In order to meet the assumptions of the stated techniques, 4631 data suitable for the realization of the research were reached after the deleted data and the research was conducted with this data set.

Data Analysis

Before proceeding with the analysis procedures, CFA analysis was conducted to examine whether the assumed latent construct was provided in terms of the research data. SEM can be categorized in to two categories: measurement and structural models. In measurement models, observed variables and latent variables are associated (Şen, 2020). With the measurement model created, it was revealed whether the seven items in *Awacom* reflect the assumed students' awareness of being able to communicate. Since multivariate normality was not achieved as a result of Mardia's multivariate skewness and kurtosis statistics (skewness and kurtosis <.05) (Tabachnick & Fidell, 2019), the robust standard error maximum likelihood (MLR) estimator was used as a parameter estimator for CFA (Rosseel et al., 2024). The fit indices (RMSEA=.065, CFI=.973, TLI=.96, SRMR=.024) obtained by analyzing the measurement model indicate that *Awcom* can be represented by seven items (Hox et al., 2018). The Cronbach alpha value calculated to determine the reliability of the responses to these seven items was .89, indicating that the reliability of the responses was high (George & Mallery, 2003).

For the aim of the study, responses to two randomly selected items from *Awacom* were predicted by neural networks trained on responses to all items reflecting the same latent trait. The items whose responses were predicted were labeled as Item3 and Item7 (SeeTable 1). The ANN models used in the predictions were created according to different conditions of splitting the data at different rates and the number of layers and the number of neurons in the layers. In machine learning methods, of which ANN is a member, the data are divided in to two parts as training and testing sets in order to avoid the problem of overlearning. The models trained with train data are controlled through other data sets (Brownlee, 2020). For this reason, nine different models were created depending on the split ratio, number of layers and neurons. The models with the most appropriate RMSE values were used as prediction models. Prediction procedures were performed for the relevant items in the test data set. In the analysis processes performed with ANN, the items consisting of four ordinal categories (1-2-3-4) were scaled between 0 and 1 (0-.333-.667-1) according to the model's assumption (Brownlee, 2020). Estimations were made according to the scaled values. After the analysis, these values were converted back to ordinal values by considering close ranges. Accuracy ratio, marginal homogeneity test (Agresti, 2013) and Kappa (Cohen, 1960) statistics were used to reveal the similarity between the estimated values and the actual values for Item3 and Item7. In this way, similarities between predicted and actual values were determined. In addition, sensitivity analyses (Beck, 2018; Lek et al., 1996) were conducted to determine the relationship between responses to actual Item3 and Item7 items and responses to other items. Then, how the subsets containing the predicted Item3 and Item7 items and the subsets containing the actual Item3 and Item7 items represent the assumed *Awacom* latent construct was also examined through CFA analyses. Thus, it was observed how the model fit indices, variance explained by the items and standardized factor loadings changed. In addition, Cronbach's alpha value was used to investigate how the reliability of the responses for the actual and predicted subsets changed. R [caret package (Kuhn, 2023), lavan package (Rosseel et al., 2024), neuralnet package (Fritsch, 2019), neuralnettools package (Beck, 2022), nnet package (Ripley & Venables, 2023)] Mplus and SPSS statistical programs were used for analysis.

Findings

The findings obtained from the prediction of two items labeled Item3 and Item7 by the networks trained with the items of the *Awacom* scale are presented under two separate headings. As a reminder, the rationale for prediction is based on the performance of ANNs that learn from the responses to items on the same scale, i.e. items that measure similar attributes, in predicting the responses to each item on the scale. These findings include the selection method of the networks, the performance of the networks, how the predicted items relate to the other items in the scale, the similarity of the actual and predicted values, the fit metrics in the verification of the assumed latent trait over the test subsets formed by the predicted and actual items, the variance values explained by the items, factor loadings and reliability values.

Predicting Item3 Responses

ANN models with nine different features were developed for predicting the item labeled Item3, where Item3 is used as output data and the other six items are used as input data. The networks have different ratios of test-train data and different numbers of layers and neurons. In evaluating the performance of the networks, the RMSE values produced by the trained network on all, train and test data were taken into account (See Table 2). In selecting the best network, it was preferred that the RMSE value was small and close for all data sets.

Table 2.

Features of ANN Models for Predicting Item3 Responses

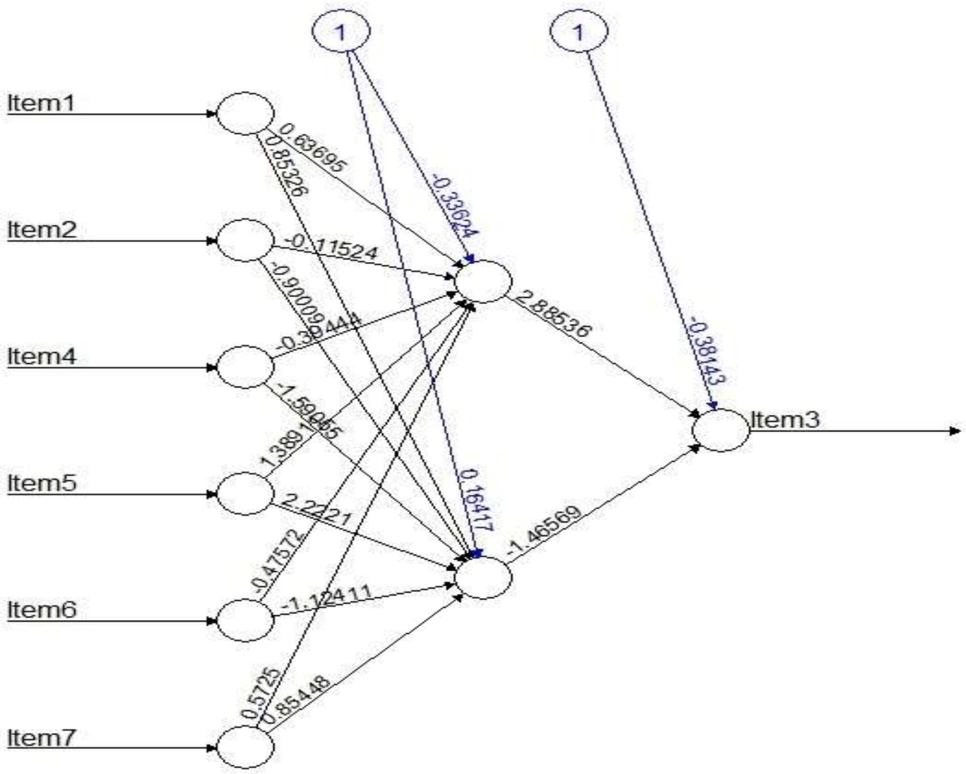
Models	Train/Test Spliting	Hiddens	RMSE		
			All	Train	Test
Model1*		2	.619	.618	.623
Model2	70/30	3	.617	.613	.625
Model3		3:2	.618	.612	.631
Model4		2	.619	.616	.629
Model5	75/25	3	.615	.609	.631
Model6		3:2	.614	.607	.636
Model7		2	.619	.617	.628
Model8	80/20	3	.618	.613	.638
Model9		3:2	.616	.612	.633

* Selected to best model

In this context, the most ideal model for predicting Item3 responses was found to be a single hidden layer network with two neurons in the hidden layer (See Figure 2). It is understood that the selected network performs well with 70-30% of the train and test data.

Figure 2.

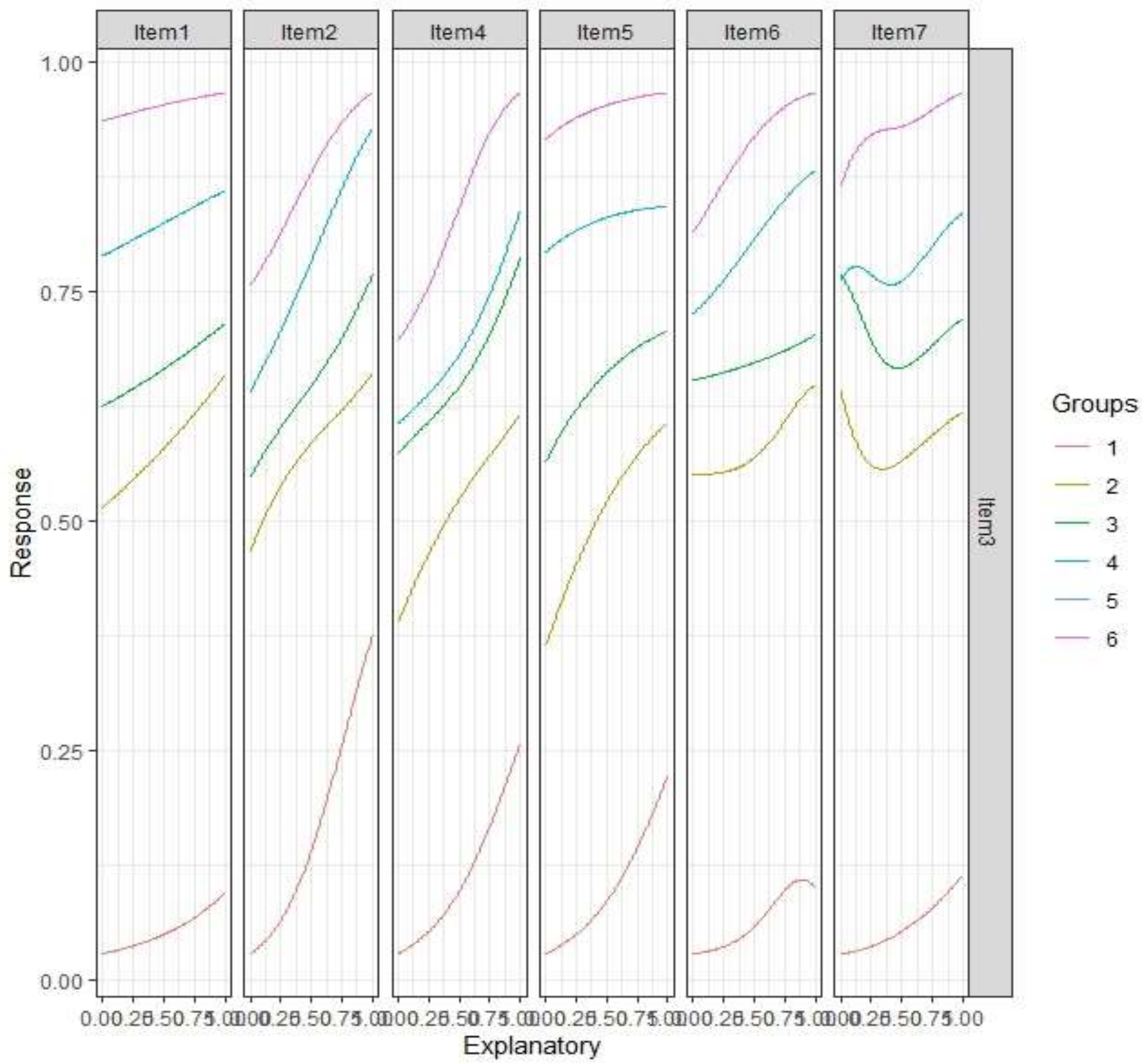
Network Structure of Model1



According to the output of the sensitivity analysis conducted to determine the relationship between the responses to Item3 predicted by Model 1 and the responses to other items, it is understood that the predicted responses to Item3 have a relatively linear relationship with the responses to other items (See Figure 3). It is noteworthy that the responses to the items other than Item7 are positively correlated with the responses to Item3. However, it is not possible to say that the same is the case for Item7. When the judgments expressed by the items are analyzed (See Table 1), the similarity and linearity of the relationship between the responses to the five items other than Item7 and the estimated Item3 values indicate that the predictions support the relevant latent construct (Beck, 2018; Lek et al., 1996).

Figure 3.

Results of Sensivity Analysis for Model1



Predicting Item7 Responses

Similarly, ANN models with nine different features were created where the output variable was Item7 and the input variables were the other six items in the *Awacom* scale. The differentiation in the networks is due to the differences in the ratio considered in the split of the data set and the number of layers and neurons. According to the RMSE values produced by the trained model for all, train and test datasets, the model with the smallest and closest RMSE values was selected as the best model (See Table 3).

Table 3.

Features of ANN Models for Predicting Item7 Responses

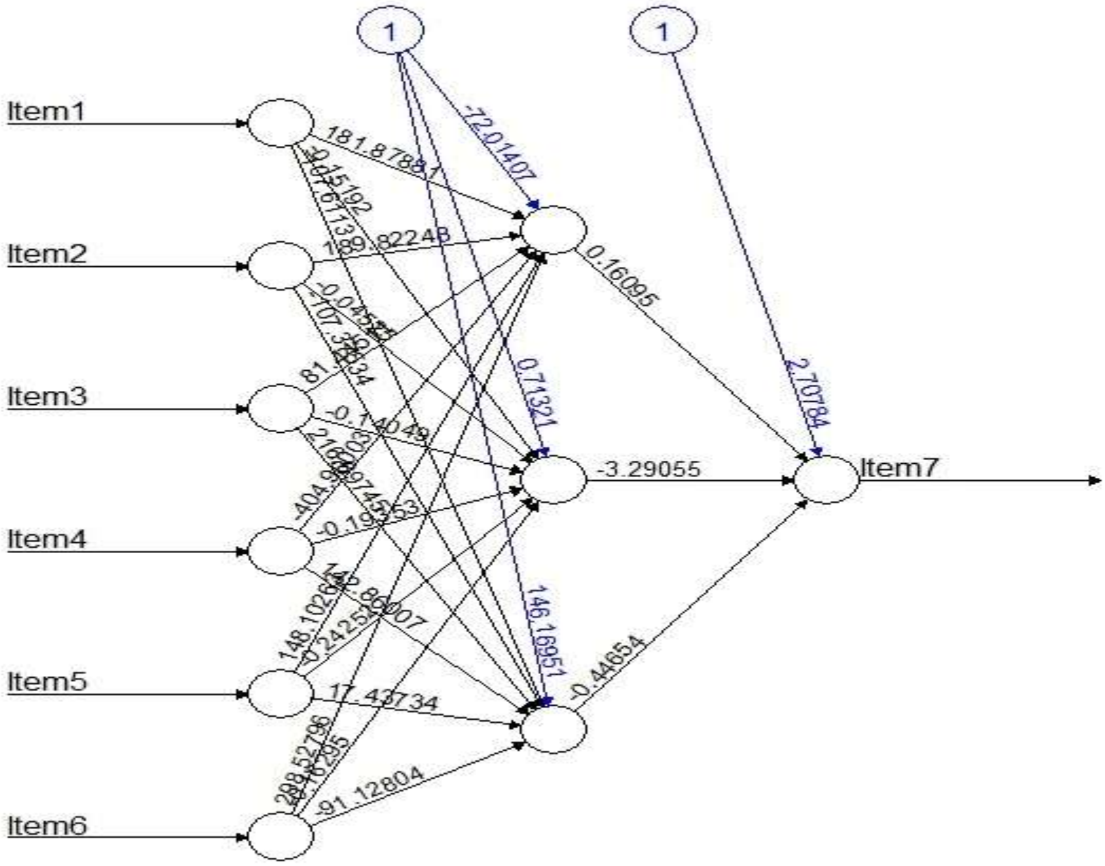
Models	Train/Test Spliting	Hiddens	RMSE		
			All	Train	Test
ModelA		2	.709	.699	.732
ModelB*	70/30	3	.708	.695	.734
ModelC		3:2	.708	.694	.740
ModelD		2	.710	.697	.750
ModelE	75/25	3	.710	.695	.752
ModelF		3:2	.706	.690	.753
ModelG		2	.709	.697	.759
ModelH	80/20	3	.708	.694	.762
ModelI		3:2	.707	.692	.765

* Selected to best model

Based on the RMSE values, it is understood that the lowest and closest values for all data sets are obtained at 70-30% separation of the data set. It is observed that the model in this group, which has ideal values, is a single interlayer network structure with three neurons (See Figure 4). Therefore, ModelB was preferred as the ideal model for predicting the responses to Item7.

Figure 4.

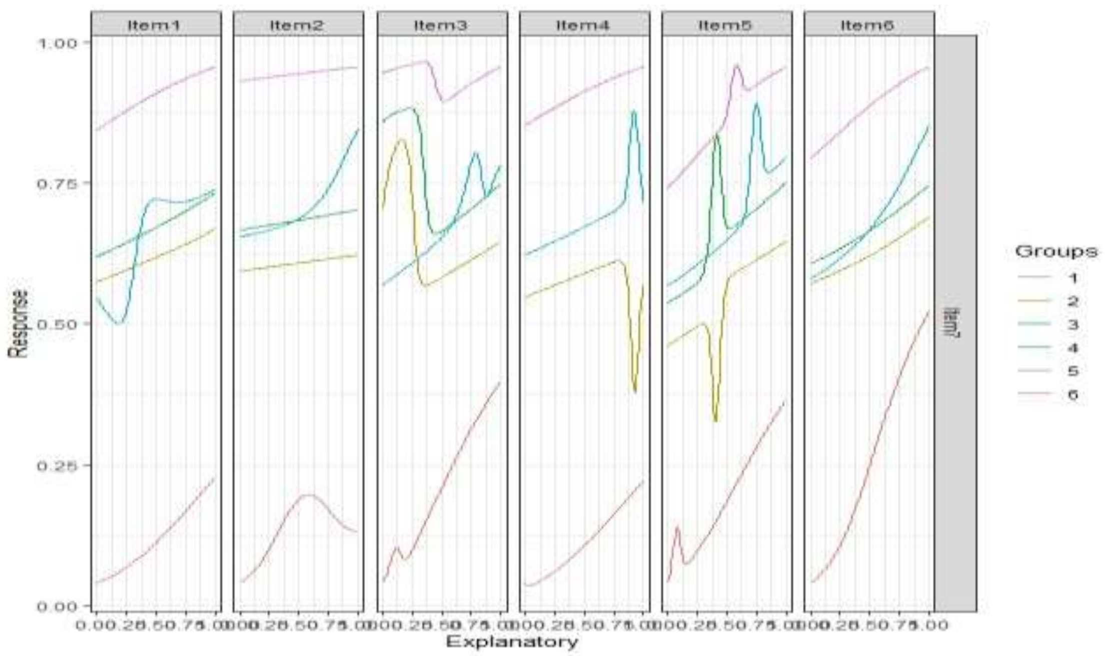
Network Structure of ModelB



It can be said that the responses to Item7 estimated for ModelB are generally not in a linear relationship with the responses to the other six items used to train the model (See Figure 5). This finding obtained by sensitivity analysis supports the relationship between Item3 estimated for Model1 and the other items. As observed in Figure 3, the responses to Item3 showed a similar and linear relationship with the responses to the other items except Item7. In this context, it is considered as an expected situation that the responses to Item7 have a non-linear relationship with the responses to other items.

Figure 5.

Results of Sensivity Analysis for ModelB



Performance of ANN Models for Actual and Predicted Values

Accuracy, marginal homogeneity test and Kappa statistics were used to determine how much the Item3 values predicted by Model1 and Item7 values predicted by ModelB corresponded to the actual values (See Table 4). It is understood that the responses to both estimated items are similar to the actual responses at an average rate of .60. There was no statistically significant difference between the mean responses to the two predicted items and the mean responses to the actual items (MH test, p >.05). According to Kappa values, there was a moderate similarity between predicted and actual Item3 values and a low similarity between predicted and actual Item7 values (Landis & Koch, 1977).

Table 4.

Similarity Values of Actual and Predicted Responses

Models	N	Match ratio	MH* test (p)	Kappa
Model1				
Actual Item3 Predicted Item3	1389	.63	>.05	.42
ModelB				
Actual Item7 Predicted Item7	1389	.58	>.05	.37

*Marginal homogeneity

CFA analyses were conducted to determine the extent to which the test subsets containing the estimated Item3 and Item7 and the test subsets containing the actual values provided the *Awacom* latent construct, and reliability analyses were conducted to determine how the reliability of the responses to the items estimated for these subsets affected the reliability (See Table 5). At the same time, the variance values and factor loadings explained by the predicted items for the latent construct were examined. The findings revealed that unlike similarity analyses, construct validity and reliability analyses increased the reliability of the responses to the predicted items and supported the latent construct. It shows that the model fit indices were relatively weakened by the two estimated items, but the model fit remained strong. On the other hand, it is understood that the variance values explained by the responses to the estimated items for the latent construct, the standardized factor loadings for the latent construct and the reliability values improved.

Table 5.

Comprassion Model Performance of Actual and Predicted Test Subsets

Subsets	RMSEA	CFI	TLI	R-square	Loading	Alpha
For Item3						
Actual subset	.081	.972	.959	.613	.783	.891
Predicted subset	.106	.965	.948	.983	.992	.903
For Item7						
Actual subset	.081	.972	.959	.489	.700	.891
Predicted subset	.111	.963	.944	.942	.970	.909

Conclusion and Discussion

ANNs, which evolved from the idea of simulating the human brain, provide significant advantages for the realization of many researches with different purposes and content due to its ability to model complex, non-linear relationships, unlike traditional statistical methods, its excellent fault tolerance, and its ability to be a fast and highly scalable machine learning (Zou et al., 2009). This study takes advantage of these important advantages of ANNs and examines how ANNs predict the responses of Lebanese students participating in PISA 2018 to other items that are part of the same emotion based on the integrity of emotion. Using some of the data, the network trained with the items of PISA's student's intercultural communicative awareness scale was able to predict two items of the same scale from another data set.

It is understood that the values predicted by ANN solutions for the responses to two randomly selected items match the actual values at a relatively moderate level. In this context, the values produced by the trained network are expected responses depending on the emotional integrity shaped by the responses to the items. Therefore, the responses expected from individuals for the items examined overlap with the responses given by individuals at a relatively moderate level. However, in the validity and reliability analyses conducted for the latent trait represented by the predicted items together with the other items, it is observed that although the predicted values partially weaken the model fit indices, they still manage to keep them strong. In addition, the estimated values improved the factor loadings and the variance explained for the latent trait. Similarly, when the latent trait aspect is considered, it is noticed that the estimated values also positively affect the reliability.

Especially in recent years, the use of advanced versions of ANNs such as convolutional neural networks (CNN), recurrent neural networks (RNN), emotional neural network (EANN) for predicting individual emotions based on machine learning using images, text, dialog, body movements, etc. in deep emotional fields such as affective state, affective computing, deep learning (Ashwin & Guddeti, 2020; Bakkialakshmi et al., 2022; Carstensen et al., 2016; Chan et al., 2020; Feng, 2022; Jadhav

&Sugandhi, 2018; Jamisola, 2016; Liu et al., 2023; Orozco-del-Castillo et al., 2021; Wang et al., 2022) has gained rapid momentum. Research using these techniques focuses on the relationship between the development of emotional responses by utilizing physiological responses. However, our research is based on a much simpler logic and purpose than these studies. We used ANN to predict the responses to an item in the same emotional state by utilizing emotional integrity. Noting that most of the researches (Aybek & Okur, 2018; Aydoğan & Zırhlıoğlu, 2018; Aydoğdu, 2020; Al-Saleem et al., 2015; Chavez et al., 2023; Flitman, 1997; Guarín et al., 2015; Huang & Fang, 2012; Lau et al., 2019; Rodríguez-Hernández, 2021; Shahiri & Husain, 2015; Umar, 2019; Zacharis, 2016) conducted with ANN in education and psychology are for cognitive prediction by utilizing covarities, we evaluated to what extent this performance of ANNs can be used to predict the responses to any item in scale applications frequently used in education and psychology.

Funding

The research was not financially supported by any institution or organisation.

Declarations

Conflict of Interest: The authors have no relevant financial or non-financial interests to disclose.

Ethical Approval: We declare that all ethical guidelines for the author have been followed. This study does not require any ethics committee approval as it includes open-access data.

The authors of this article declare (Declaration Form #: 2711240128) that Gen-AI tools have NOT been used in any capacity for content creation in this work.

References

Agresti, A. (2013). *Categorical data analysis.* Wiley.

Al-Saleem, M., Al-Kathiry, N., Al-Osimi, S., & Badr, G. (2015). Mining educational data to predict students' academic performance. In *Machine Learning and Data Mining in Pattern Recognition: 11th International Conference, MLDM 2015, Hamburg, Germany, July 20-21, 2015, Proceedings 11* (pp. 403-414). Springer International Publishing.

Ashwin, T. S., & Guddeti, R. M. R. (2020). Automatic detection of students' affective states in classroom environment using hybrid convolutional neural networks. *Education and Information Technologies*, *25*(2), 1387-1415. https://doi.org/10.1007/s10639-019-10004-6

Aybek, H. S. Y., & Okur, M. R. (2018). Predicting achievement with artificial neural networks: The case of Anadolu University open education system. *International Journal of Assessment Tools in Education*, *5*(3), 474-490. https://doi.org/10.21449/ijate.435507

Aydoğan, İ., & Zırhlıoğlu, G. (2018). Öğrenci başarılarının yapay sinir ağları ile kestirilmesi. *Van Yüzüncü Yıl Üniversitesi Eğitim Fakültesi Dergisi*, *15*(1), 577-610. http://dx.doi.org/10.23891/efdyyu.2018.80

Aydoğdu, Ş. (2020). Predicting student final performance using artificial neural networks in online learning environments. *Education and Information Technologies*, *25*(3), 1913-1927. https://doi.org/10.1007/s10639-019-10053-x

Bakkialakshmi, V. S., Sudalaimuthu, T., & Winkler, S. (2022). Effective Prediction System for Affective Computing on Emotional Psychology with Artificial Neural Network. *Easy Chair Preprint*.

Beck, M.W. (2018). NeuralNetTools: Visualization and Analysis Tools for Neural Networks. *Journal of Statistical Software*, *85*(11), 1 .https://doi.org/10.18637/jss.v085.i11

Beck, M.W. (2022). Visualization and analysis tools for neural networks, R package version 1.5.3. Retrieved from https://cran.r-project.org/web/packages/NeuralNetTools/index.html

Brownlee, J. (2020). *Data preparation for machine learning: data cleaning, feature selection, and data transforms in Python*. Machine Learning Mastery.

Carstensen, S. L., Madsen, J., & Larsen, J. (2016). Predicting Changes in Affective States using Neural Networks. *arXiv preprint arXiv:1612.00582*. https://doi.org/10.48550/arXiv.1612.00582

Chan, K. Y., Kwong, C. K., Wongthongtham, P., Jiang, H., Fung, C. K., Abu-Salih, B., ... & Jain, P. (2020). Affective design using machine learning: a survey and its prospect of conjoining big data. *International Journal of Computer Integrated Manufacturing*, *33*(7), 645-669. https://doi.org/10.1080/0951192X.2018.1526412

Chavez, H., Chavez-Arias, B., Contreras-Rosas, S., Alvarez-Rodríguez, J. M., & Raymundo, C. (2023, February). Artificial neural network model to predict student performance using nonpersonal information. In *Frontiers in Education* (Vol. 8, p. 1106679). Frontiers Media SA.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*(1), 37-46. https://doi.org/10.1177/001316446002000104

Ennett, C. M., Frize, M., & Walker, C. R. (2001). Influence of missing values on artificial neural network performance. In *MEDINFO 2001* (pp. 449-453). Ios Press.

Feng, H. (2022). A Novel Adaptive Affective Cognition Analysis Model for College Students Using a Deep Convolution Neural Network and Deep Features. *Computational Intelligence and Neuroscience*, *2022*(1), 2114114. https://doi.org/10.1155/2022/2114114

Flitman, A. M. (1997). Towards analysing student failures: neural networks compared with regression analysis and multiple discriminant analysis. *Computers & Operations Research*, *24*(4), 367-377. https://doi.org/10.1016/S0305-0548(96)00060-3

Fritsch, S., Guenther, F., Wright, M.N., Suling, M., Mueller, S.M. (2019). Training of neural Networks, R package version 1.44.2. Retrieved from https://cran.r-project.org/web/packages/neuralnet/index.html

George, D., & Mallery, P. (2003). *SPSS for Windows step by step: A simple guide and reference. 11.0 update* (4th ed.). Allyn & Bacon.

Goel, A., Goel, A. K., & Kumar, A. (2023). The role of artificial neural network and machine learning in utilizing spatial information. *Spatial Information Research*, *31*(3), 275-285. https://doi.org/10.1007/s41324-022-00494-x

Grosan, C., & Abraham, A. (2011). Artificial neural networks. *Intelligent Systems: A Modern Approach*, 281-323.

Guarín, C. E. L., Guzmán, E. L., & González, F. A. (2015). A model to predict low academic performance at a specific enrollment using data mining. *IEEE Revista Iberoamericana de tecnologias del Aprendizaje*, *10*(3), 119-125. https://doi.org/10.1109/RITA.2015.2452632

Hasson, U., Nastase, S. A., & Goldstein, A. (2020). Direct fit to nature: an evolutionary perspective on biological and artificial neural networks. *Neuron*, *105*(3), 416-434. https://doi.org/10.1016/j.neuron.2019.12.002

Hox, J., Moerbeek, M., & van de Schoot, R. (2018). *Multilevel analysis: Techniques and applications* (3rd ed.). Routledge.

Huang, S., & Fang, N. (2012, October). Work in progress: Early prediction of students' academic performance in an introductory engineering course through different mathematical modeling techniques. In *2012 Frontiers in Education Conference Proceedings* (pp. 1-2). IEEE.

Jadhav, N., & Sugandhi, R. (2018, November). Survey on human behavior recognition using affective computing. In *2018 IEEE Global Conference on Wireless Computing and Networking (GCWCN)* (pp. 98-103). IEEE.

Jamisola, R. S. (2016). Conceptualizing a Questionnaire-Based Machine Learning Tool that Determines State of Mind and Emotion. *Lovotics*, *4*(115), 2. http://dx.doi.org/10.4172/2090-9888.1000115

Kardan, A. A., Sadeghi, H., Ghidary, S. S., & Sani, M. R. F. (2013). Prediction of student course selection in online higher education institutes using neural network. *Computers & Education*, *65*, 1-11. https://doi.org/10.1016/j.compedu.2013.01.015

Kose, U., & Arslan, A. (2017). Optimization of self-learning in Computer Engineering courses: An intelligent software system supported by Artificial Neural Network and Vortex Optimization Algorithm. *Computer Applications in Engineering Education*, *25*(1), 142-156. https://doi.org/10.1002/cae.21787

Kuhn, M., Wing, J., Weston, S., Williams, A., Keefer, C., Engelhardt, A., … & Hunt, T. (2023). Classification and regression training, R package version 6.0-94. Retrieved from https://cran.r-project.org/web/packages/caret/index.html

Landis, J, R., & Koch, G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*, 159-174.https://doi.org/10.2307/2529310

Lau, E. T., Sun, L., & Yang, Q. (2019). Modelling, prediction and classification of student academic performance using artificial neural networks. *SN Applied Sciences*, *1*(9), 982. https://doi.org/10.1007/s42452-019-0884-7

Lek, S., Delacoste, M., Baran, P., Dimopoulos, I., Lauga, J., & Aulagnier, S. (1996). Application of neural networks to modeling nonlinear relationships in ecology. *Ecological Modelling*, *90*, 39–52. https://doi.org/10.1016/0304-3800(95)00142-5

Lillicrap, T. P., Santoro, A., Marris, L., Akerman, C. J., & Hinton, G. (2020). Backpropagation and the brain. *Nature Reviews Neuroscience*, *21*(6), 335-346. https://doi.org/10.1038/s41583-020-0277-3

Liu, J., Ang, M. C., Chaw, J. K., Kor, A. L., & Ng, K. W. (2023). Emotion assessment and application in human–computer interaction interface based on backpropagation neural network and artificial bee colony algorithm. *Expert Systems with Applications*, *232*, 120857. https://doi.org/10.1016/j.eswa.2023.120857

McCarthy, N. (June19, 2019). Lebanon has by far the most refugees per 1,000 population. Retrieved from https://www.statista.com/chart/8800/lebanon-has-by-far-the-most-refugees-per-capita/

OECD. (2018). PISA 2018 Technical report. OECD Publishing, Retrieved from https://www.oecd.org/pisa/data/pisa2018technicalreport/

Orozco-del-Castillo, M. G., Orozco-del-Castillo, E. C., Brito-Borges, E., Bermejo-Sabbagh, C., & Cuevas-Cuevas, N. (2021, November). An artificial neural network for depression screening and questionnaire refinement in undergraduate students. In *International Congress of Telematics and Computing* (pp. 1-13). Springer International Publishing.

Öztemel, E. (2003).*Yapay sinir ağları.* PapatyaYayıncılık.

Park, C. W., Seo, S. W., Kang, N., Ko, B., Choi, B. W., Park, C. M., ... & Yoon, H. J. (2020). Artificial intelligence in health care: Current applications and issues. *Journal of Korean Medical Science, 35*(42). https://doi.org/10.3346/jkms.2020.35.e379

Rashid, T. A., & Ahmad, H. A. (2016). Lecturer performance system using neural network with Particle Swarm Optimization. *Computer Applications in Engineering Education*, *24*(4), 629-638. https://doi.org/10.1002/cae.21737

Ripley, B., & Venables, W. (2023). Feed-forward neural networks and multinomial log-linear models, R package version 7.3-18. Retrieved from https://cran.r-project.org/web/packages/nnet/index.html

Rodríguez-Hernández, C. F., Musso, M., Kyndt, E., & Cascallar, E. (2021). Artificial neural networks in academic performance prediction: Systematic implementation and predictor evaluation. *Computers and Education: Artificial Intelligence, 2*, 100018. https://doi.org/10.1016/j.caeai.2021.100018

Rosseel, Y., Oberski, D., Byrnes, J., Vanbrabant, L., Savalei, V., Merkle, E., … & Jorgensen, T. (2024). Latent variable analysis, R package version 0.6-18. Retrieved from https://cran.r-project.org/web/packages/lavaan/lavaan.pdf

Shahiri, A. M., & Husain, W. (2015). A review on predicting student's performance using data mining techniques. *Procedia Computer Science*, *72*, 414-422. https://doi.org/10.1016/j.procs.2015.12.157

Şen, S. (2020). *Mplus ile yapısal eşitlik modellemesi uygulamaları.* Nobel Akademik Yayıncılık.

Tabachnick, B. G., & Fidell, L. (2019). *Using multivariate statistics* (7th ed.). Pearson.

Tu, J. V. (1996). Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. *Journal of Clinical Epidemiology*, *49*(11), 1225-1231. https://doi.org/10.1016/S0895-4356(96)00002-9

Umar, M. A. (2019). Student academic performance prediction using artificial neural networks: A case study. *International Journal of Computer Applications*, *178*(48), 24-29. http://dx.doi.org/10.5120/ijca2019919387

Vandamme, J. P., Meskens, N., & Superby, J. F. (2007). Predicting academic performance by data mining methods. *Education Economics*, *15*(4), 405. https://doi.org/10.1080/09645290701409939

Wang, Y., Song, W., Tao, W., Liotta, A., Yang, D., Li, X., ... & Zhang, W. (2022). A systematic review on affective computing: Emotion models, databases, and recent advances. *Information Fusion*, *83*, 19-52. https://doi.org/10.1016/j.inffus.2022.03.009

Zacharis, N. Z. (2016). Predicting student academic performance in blended learning using artificial neural networks. *International Journal of Artificial Intelligence and Applications*, *7*(5), 17-29. http://dx.doi.org/10.5121/ijaia.2016.7502

Zou, J., Han, Y., & So, S. S. (2009). Overview of artificial neural networks. *Artificial neural networks: methods and applications*, 14-22. https://doi.org/10.1007/978-1-60327-101-1_2