DOI: 10.22103/jlst.2017.10043.1188

Prediction of breeding values for the milk production trait in Iranian Holstein cows applying artificial neural networks

S. Pour Hamidi[1], M. R. Mohammadabadi[1*], M. Asadi Foozi[1] and H. Nezamabadi-pour[2]

[1]Department of Animal Science, Faculty of Agriculture, Shahid Bahonar University of Kerman, Kerman, Iran.
[2]Department of Electrical Engineering, Shahid Bahonar University of Kerman, Kerman, Iran.
* Corresponding author, E-mail address: mmohammadabadi@yahoo.ca

Abstract The artificial neural networks, the learning algorithms and mathematical models mimicking the information processing ability of human brain can be used non-linear and complex data. The aim of this study was to predict the breeding values for milk production trait in Iranian Holstein cows applying artificial neural networks. Data on 35167 Iranian Holstein cows recorded between 1998 to 2009 were obtained from the Animal Breeding Center of Iran. Breeding values for the milk production trait were determined using the ASReml univariate animal model with 70% of all data used as training data, 15% as testing data and 15% as validating data, to prevent over-fitting of the artificial neural network. A feed-forward backpropagation multilayer perceptron algorithm with three-layer MLP; including 1 input layer, 1 hidden layer and 1 output layer and four-layer MLP; including 1 input layer, 2 hidden layer and 1 output layer was used. The most influential parameters for input characters in artificial neural network were sire, herd, calving year, twice-daily milking (Milk 2x), calving season and age based month. Breeding values for milk production was used as variable output. For network with 4 layers, the best selected structure for the first lactation trait contained input layers with 6 neurons, first hidden layer with 16 neurons and with 68 epoch, second hidden layer with 6 neurons and with 154 epoch and output layer with 1 neuron. The capability of artificial neural network model was higher and closer to the estimated breeding values; therefore it is possible to apply artificial neural networks, instead of commonly-used procedures for predicting the breeding values for milk production.
Keywords: artificial neural networks, breeding value, milk production, Iranian Holstein cows

Received: 12 Apr. 2017, *accepted:* 03 Sep. 2017, *published online:* 25 Dec. 2017

Introduction

According to the latest official statistics published by the Iranian Ministry of Agriculture, there are 18830 industrial dairy farms in Iran with 2048563 dairy cows. Between 2004 to 2008, milk production in Iran followed a growing trend. Despite the increased milk production in Iran, per capita consumption is lower than the international standard. Per capita consumption of milk is 95 kg, while in the world is 169 kg and in the Europe is 350 kg. According to the available data, it is logical that the breeding objectives be used to increase milk production in Iran (OSIT, 2008). In dairy cow industry, precise and proper prediction of milk production and performance of offspring is an important pre-requisite in selecting genetically superior sires (Lacroix et al., 1995; Salehi et al., 1998).

In a breeding program, genetic progress can be maximized through accurate identification of superior animals that will be selected as parents of the next generation and therfore breeding goals can be achieved (Lacroix et al., 1995; Salehi et al., 1998). A key component of this process is fast and reliable prediction of breeding values for the selection candidates. But, prediction of breeding values is often a computationally challenging and time consuming task, and therfore it is undertaken only periodically in most countries (Lacroix et al., 1995; Salehi et al., 1998). Rapid, lowcost alternatives that can provide approximate predictions of breeding values with acceptable accuracy could allow more timely selection and culling decisions by breeding companies or dairy producers. Rapid identification of superior males can lead to earlier collection and distribution of semen and more rapid genetic progress (Lacroix et al., 1995; Salehi et al., 1998). Moreover, investigations have shown that the conventional regression procedures cannot evaluate the multicolinearity between independent factors; hence it may result in biased outcomes (Raja et al., 2012; Ruhil et al., 2013). When the correlation between variables is high, multicollinearity takes place; therefore, it is difficult to obtain reliable estimates of the individual regression coefficients (Eyduran and Yavuzsonmez, 2010). In situations that correlation of some variables is very high; they are basically measuring the same phenomenon and give similar information, hence these variables can inversely affect the regression outcome. Difficulties caused by multicolinearity in the regression analysis have been documented (Eyduran and Yavuzsonmez, 2010; Ramzan and Khan, 2010). Artificial neural networks have been proposed in alleviating this limitation of the traditional regression methods, and can be used to handle non-linear and complex data, even when the data are imprecise and noisy (Raja et al., 2012). These networks contain a set of processing components, also known as neurons or nodes whose functionality is loosely based on biological neurons (Raja et al., 2012). These units are formed in layers that process the input information and pass it to the next layers. The capability of the network in processing is cumulated in the interunit connection strengths (or weights) that are acquired via a process of conformity to a collection of training pattern (Haykin, 1999; Raja et al., 2012). Moreover, the artificial neural network method entirely varies from the traditional statistical approaches, which need a specified algorithm to be transformed by a computer program (Grzesiak et al., 2003). Artificial neural networks have rarely been apllied to data in animal science, particularly in animal breeding (Ghazanfari et al., 2011; Ruhil et al., 2013; Shahinfar et al., 2012). Shahinfar et al. (2012) showed that artificial neural networks and neuro-fuzzy systems reliably predicted the breeding values for milk and fat yield in dairy cows. Silva et al. (2014) showed that artificial neural network was superior to linear models in estimating the breeding values, as well as applicability in predicting the genetic values. Although many studies have been performed on milk production traits and the genes affecting these traits in Iranian Holstein and native cows (Javanmard et al., 2008; Mohammadabadi et al., 2010; Pasandideh et al., 2015). We are not aware of any studies applying the artificial neural networks to predicting the breeding values milk production traits; therefore, the aim of this study was to predict the breeding value for milk production in Iranian Holstein cows by using the artificial neural networks.

Table 1. Data structure on milk yield trait in Iranian Holsteins

Statistic	Value
First lactation milk yield (kg)	6763
Standard deviation (kg)	1377
Number of records	6511
Number of sires	1098
Number of dams	27133

Materials and methods

Data on 35167 Iranian Holstein cows recorded between 1998 to 2009 were obtained from the Animal Breeding Center of Iran (Table 1).

After omiting the out-of-range and illogical data, the data were normalized and tandardized using the Excel, LINUX and NeuroSolution (http://www.neurosolutions.com) software. Breeding values for the milk production trait were estimated using the ASReml software via univariate animal model (1):

$$y=Xb+Zu+e \tag{1}$$

where, y; is the observation vector, X; matrix of design for fixed effects, b; fixed effects vector, Z; matrix of design for genetic effects, u; genetic effects vector and e; error random effects vector.

From all data, 70% data were used as training set, 15% as testing set and 15% as the validating set, to prevent over-fitting of artificial neural network. A feed-forward backpropagation multilayer perceptron (MLP herein) algorithm with three-layer and four-layer was used in MATLAB v7.0 software (The MathWorks, Natick, MA, USA). Three-layer MLP included 1 input layer, 1 hidden layer and 1 output layer, but four-layer MLP contained 1 input layer, 2 hidden layer and 1 output layer.

The most influential parameters for input characters in artificial neural network were the sire, herd, calving year, twice-daily milking (Milk 2x), calving season and age at calving in month. The breeding values for milk production trait was used as the variable output (Figure 1).

Each node in the input layer corresponds to one explanatory variable. Nodes in the hidden layer contain hyperbolic tangent activation functions (Hagan et al., 1996) as;

$$h=(e^{y_i}-e^{-y_i})/(e^{y_i}+e^{-y_i}) \tag{2}$$

and take a weighted sum of all input variables (formula 3).

$$y_i=\sum_j \omega_{ji}\chi_i \tag{3}$$

where, χ_i is an input variable and ω_{ji} is corresponding weight in layer j. Similarly, the output node(s) takes a weighted sum of all nodes in the second hidden layer and uses the same activation function to calculate the

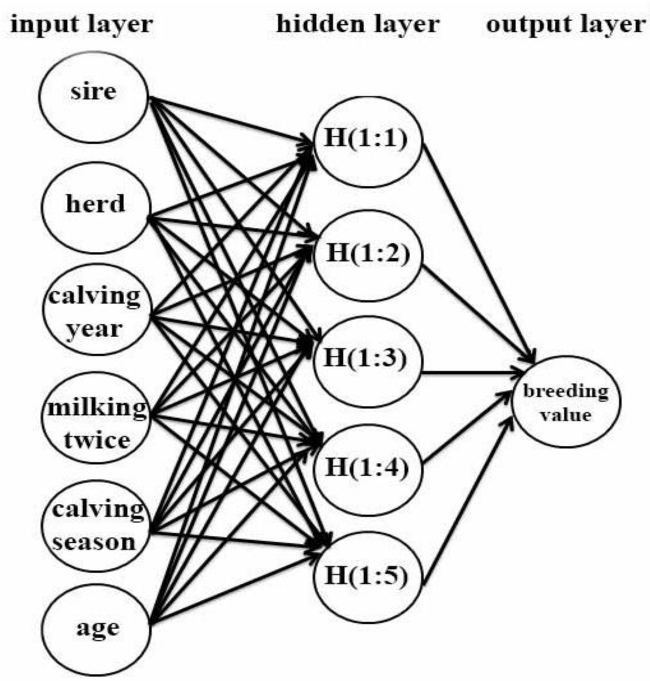

Figure 1. Variables and layers used in the training data sets for artificial neural networks in Iranian Holstein cows; 16 neurons in the hidden layers were the best but only, only 5 neurons are shown here.

output value. Learning (updating weights) in the back-propagation algorithm starts by summing up the errors over all network output unit(s). For each output unit k, the error term is:

$$E_k = o_k(1-o_k)(t_k - o_k) \tag{4}$$

where, t_k and o_k are target and output for kth output of dth training example, respectively. Then, for each hidden layer, the error term will be:

$$E_h = o_h(1-o_h)\sum_{k\in output}\omega_{kh}E_k \tag{5}$$

where, o_h is the output of the hidden layer and ω_{kh} is the weight of kth output neuron. Each weight in the network is updated by using the formulas 6 and 7.

$$\omega_{ji} = \omega_{ji} + \Delta\omega_{ji} \tag{6}$$

$$\Delta\omega_{ji} = \eta E_{jxji} \tag{7}$$

where, η is called the learning rate (e.g., 0.05), E_j is the error term for the jth node, and x_{ji} is the input value for jth node in ith layer to which the weight is applied (Mitchell, 1997). The tangent hyperbolic function also ranges from −1 to 1 and is differentiable, which has two advantages. First, it is necessary when using in backpropagation algorithm and second it gives a prediction range between −1 and 1 which is well suited for this study, because in our case, breeding values can take both positive and negative values (Haykin, 1999).

Results

After examination of the network with different neurons for the first lactation and evaluation of correlation coefficients for the testing and training data (Figure 2), the network with the 1 hidden layer and 16 neurons in the hidden layer was choosen as the best.

As shown in Figure 2, the highest correlation coefficient was related to the network including 16 and 17 neurons in the first hidden layer, but for reduction of the system complexity, 16 neurons were selected as the best structure for the network with 3 layers. The network with 3 layers yielded the largest correlation coefficient (R=0.82; Figure 3) and the lowest root mean square error (RMSE=0.07).

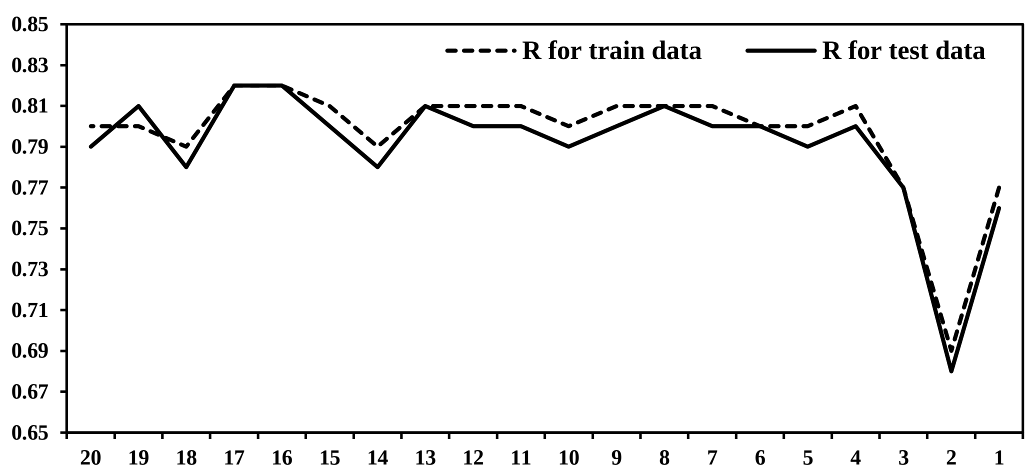

Figure 2. Correlation coefficients of the testing and training data sets related to the number of neurons in the first hidden layer for predicting the breeding value for the milk production trait in the first lactation

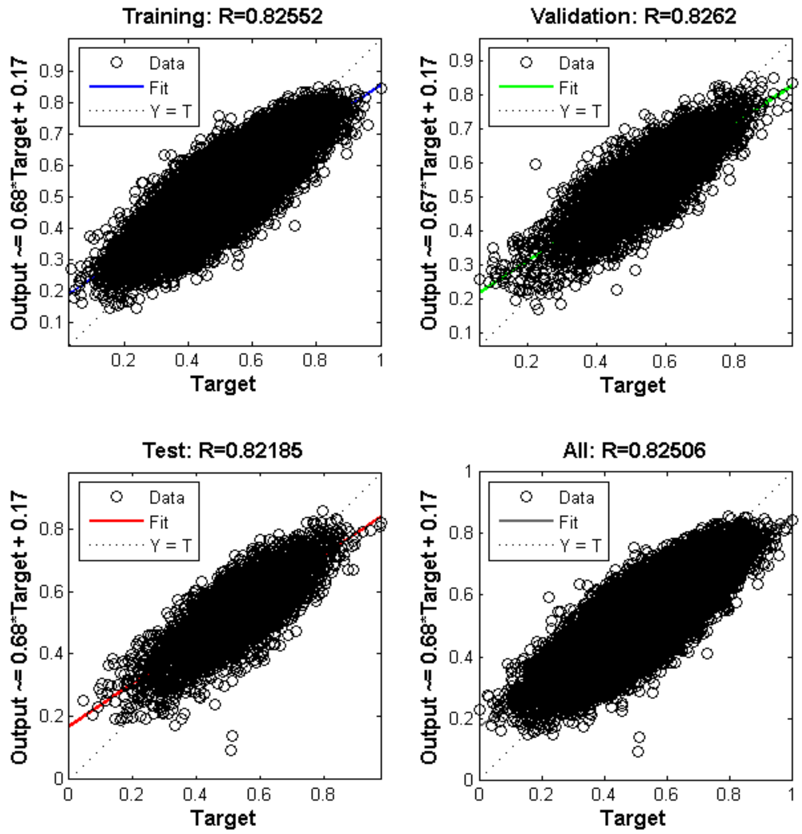

Figure 3. Correlation coefficients of the testing and training data sets for the first lactation yield

Performance curve of the validation data (Figure 4) demonstrated that mean squared error (MSE) is declining until epoch 68 and is fixed afterwards. If this process does not stop, the network instead of learning will memorize data and prediction accuracy will be low. Hence, this epoch (MSE=0.0051) was selected as the best validation performance.

Therefore, for the network with 3 layers, the best se-

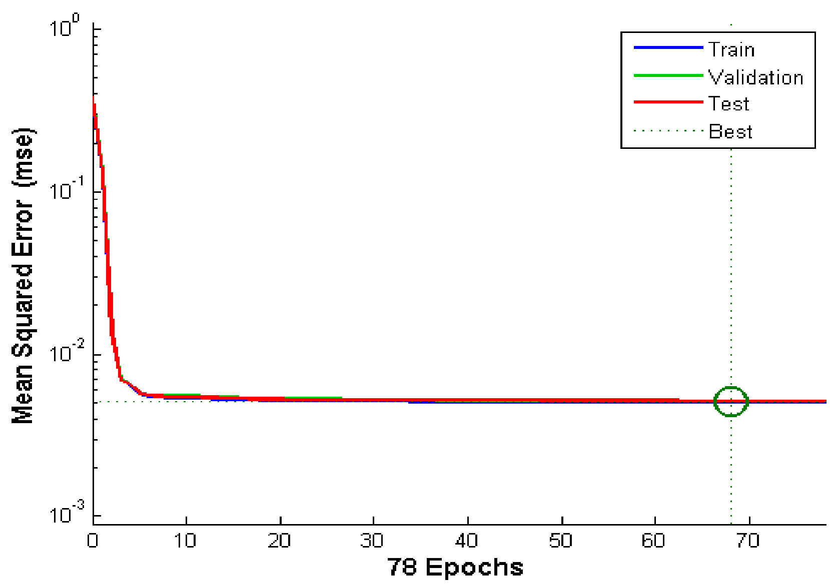

Figure 4. Performance curve of the validation data for the first hidden layer. The best validation performance is at epoch 68 (MSE=0.0051)

lected structure for milk production in the first lactation-contains input layers with 6 neurons (equal input parameters), hidden layer with 16 neurons, output layer with 1 neuron (equal output parameter) and with 68 epoch. After determining the number of neurons and epoch in the first hidden layer, the network with 4 layers was examined. For defining the number of neurons and epoch in the second hidden layers, adding neurons to the second hidden layers was started as far as the number of neurons in the second layer do not exceed the first hidden layer. Results showed that the highest correlation coefficient of the training and testing data sets corresponded to 6 neurons in the second hidden layer (Figure 5).

Figure 6 shows that for the second hidden layer of

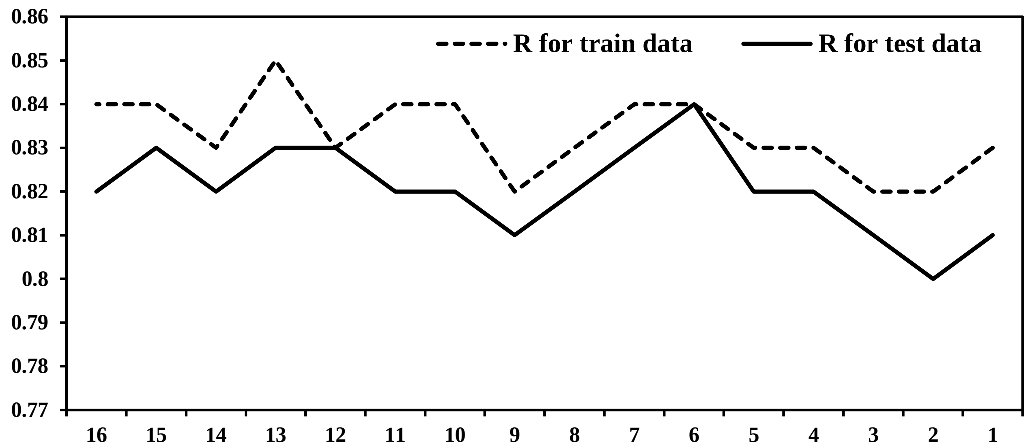

Figure 5. Number of neurons in the second hidden layer on correlation coefficients of the testing and training data set for prediction of the breeding value for milk production in the first lactation

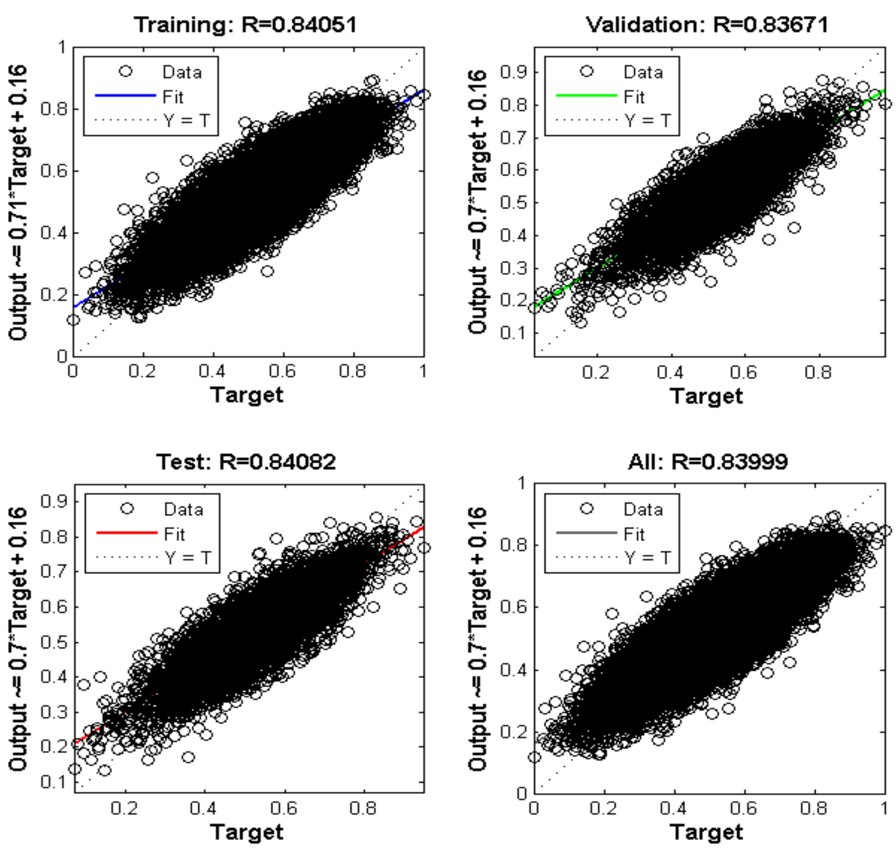

Figure 6. Correlation coefficients of the testing and training data set related to the first lactation for the second hidden layer

the network with 4 layers, the correlation coefficient (R) was the highest (0.84) and the root mean square error (RMSE) was the lowest (0.06). Performance curve of the validating data set for the second hidden layer (Figure 7) demonstrated that mean squared error (MSE) is declining until epoch 154 and is fixed afterwards, hence this epoch (MSE=0.0047) was selected as the best validation performance. Therefore, for the network with 4 layers, the best selected structure for the milk production trait in the first lactation contained input layers with 6 neurons (equal input parameters), first hidden layer with 16 neurons and with 68 epochs, second hidden layer with 6 neurons and 154 epochs and output layer with 1 neuron (equal output parameter).

Determining the best training function

In this section prediction of inbreeding values using different training functions was performed. Correlation coefficient (R) and root mean square error (RMSE) for the training and testing data sets were estimated (Table 2). Results showed that trainlm function has the least RMSE and it was better than other functions for predict-

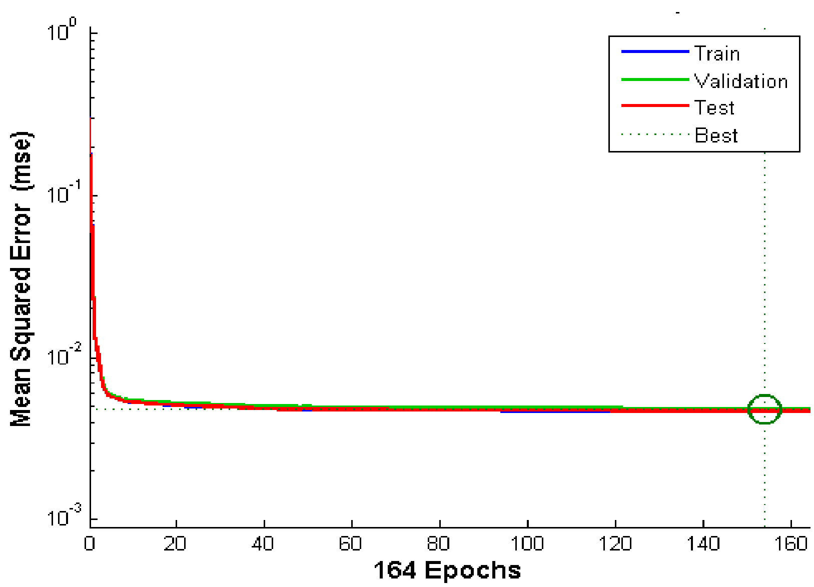

Figure 7. Performance curve of the validating data set for the second hidden layer. The best validation performance is at epoch 154 (MSE=0.0048).

Table 2. Evaluation of multilayer perceptron network in terms of the training function type for predicting the breeding values for milk production trait in the first lactation

Training function of network*	Train set	Test set	RMSE
Trainlm	0.84	0.84	0.06
Traincgp	0.81	0.80	0.07
Traincgb	0.80	0.80	0.07
Traincgf	0.80	0.79	0.07
Trainoss	0.80	0.80	0.07
Trainscg	0.80	0.79	0.07
Trainrp	0.80	0.80	0.07
Traingda	0.50	0.51	0.11
Traingdx	0.73	0.74	0.08
Traingd	0.38	0.39	0.15
Traingdm	0.57	0.58	0.13

*Trainlm: Levenberg-Marquardt back propagation, Traincgp: Conjugate Gradient with Polak-Ribiere updates, Traincgb: Conjugate Gradient with Powell-Beale restarts, Traincgf: Conjugate Gradient with Fletcher-Reeves updates, Trainoss: One step secant, Trainscg: Scaled Conjugate Gradient, Trainrp: Resilient BackPropagation, Traingda: Gradient Descent with adaptive (variable) learning rate, Traingdx: Gradient Descent with momentum and adaptive learning rate, Traingd: Gradient Descent, Traingdm: Gradient Descent with momentum tion of breeding values for the milk production trait in the first lactation.

Hence, multilayer perceptron network was the best network for predicting breeding values for the milk production trait in the first lactation. This network had 4 layers (1 input layer, 2 hidden layers and 1 output layer). In this network tangent sigmoid operates as motion function of hidden layers and linear acts as motion function of output layer. The first hidden layer had 16 neurons, the second hidden layer had 6 neurons and trainlm function as training function.

Discussion

The neural network models, trained using the training data sets to predict the breeding values for the milk production trait demonstrated that the network with 4 layers in which there are 1 input layer with 6 neurons (equal input parameters), the first hidden layer with 16 neurons and with 68 epochs, the second hidden layer with 6 neurons and with 154 epochs and output layer with 1 neuron (equal output parameter), predicted the breeding value for milk production trait in the first lactation of Holstein cows. Gorgulu (2012), using artificial neural networks for prediction of 305-day milk yield in Brown Swiss cows, showed that predicted 305-d mean milk production was very close to the observed values, with correlation coefficient (R) values between 0.74 and 0.82 for the artificial neural networks. However, 305-d milk yield prediction by multiple linear regression was lower than the observed 305-d milk yield. He proposed that artificial neural network module provided a better prediction for the 305-d milk yield than conventional regression models. Roush et al. (2006) compared the Gompertz non-linear regression model and neural network modeling for prediction of body weight in broilers and showed that neural network modeling resulted in the lowest bias.

Neural network models had also the potential for detecting minor and major pathogens that cause bovine mastitis (Hassan et al., 2009) and estimating the prevalence of clinical mastitis cases with milk production traits (Yang et al., 2000). Grzesiak et al. (2003) used multiple regression and artificial neural networks methods to predict the 305-day lactation milk yield. Their reported correlation coefficient (R) and rootmean square error (RMSE) were 0.88 and 0.08, respectively for the artificial neural networks in line with our results. Ehret et al. (2015) examined different non-linear network architectures, as well as several genomic covariate structures as network inputs in order to assess their ability to predict milk traits in three dairy cows data sets using large-scale SNP data. For training, they used a regular ized back propagation algorithm and the average correlation between the observed and predicted phenotypes in a 20 times 5-fold cross-validation used to assess predictive ability. They concluded that artificial neural network is a powerful method for non-linear genome-enabled predictions in animal breeding. However, to produce stable and high-quality outputs, variable selection methods are highly recommended, when the number of markers vastly exceeds the sample size.

Conclusions

Our results demonstrated that for predicting the breeding values for the milk production trait in Iranian Holstein cows artificial neural networks are replaceable with multiple regression models, because they had higher R^2 and Pearson's correlation coefficients and lower standard deviation and mean square error. Although both artificial neural networks and multiple regression models can excellently predict the breeding values for the milk production trait, but artificial neural network provides more precise estimates, and may be used as an alternative technique for predicting the breeding values for milk production trait.

References

Ehret, A., Hochstuhl, D., Gianola, D., Thaller G., 2015. Application of neural networks with back-propagation to genome-enabled prediction of complex traits in Holstein-Friesian and German Fleckvieh cattle. *Genetics Selection Evolution* 47, 22.

Eyduran, E., Yavuzsonmez, A., 2010. Use of factor scores in Multiple regression analysis for estimation of body weight by several body measurements in Brown Trouts (Sahno trutta fario). *International Journal of Agricultural Biology* 12, 66-76.

Ghazanfari, S., Nobari, K., Tahmoorespur, M., 2011. Prediction of egg production using artificial neural network. *Iranian Journal of Animal Science* 1, 11–16.

Gorgulu, O., 2012. Prediction of 305-day milk yield in Brown Swiss cattle using artificial neural networks. *South African Journal of Animal Science* 42, 280-287.

Grzesiak, W., Lacroixr, R., Wojcik, J., Blaszczyk, P., 2003. A comparison of neural network and multiple regression predictions for 305-day lactation yield using partial lactation records. *Canadian Journal of Animal Science* 83, 307–310.

Hagan, M.T., Demuth, H.B., Beale, M.H., 1996. 2nd Ed. Neural Network Design. PWS Publishing, Boston, USA.

Hassan, K.J., Lopez-Benavidest, M.G., Samarasinghe, S., 2009. Use of neural networks to detect minor and major pathogens that cause bovine mastitis. *Journal of Dairy Science* 92, 1493–1499.

Haykin, S., 1999. Neural networks: a comprehensive foundation. 2nd Ed. McMaster University, Ontario, Canada.

Javanmard, A., Mohammadabadi, M.R., Zarrigabayi, G.E., Gharahedaghi, A.A., Nassiry, M.R., Javadmansh, A., Asadzadeh, N., 2008. Polymorphism within the intron region of the bovine leptin gene in Iranian Sarabi cattle (Iranian Bos taurus). *Russian Journal of Genetics* 44, 495-497.

Kharrati Koopaei, H., Mohammadabadi, M.R., Ansari Mahyari, S., Tarang, A.R., Potki, P., Esmailizadeh, A.K., 2012. Effect of DGAT1 variants on milk composition traits in Iranian Holstein cattle population. *Animal Science Papers and Reports* 30, 231-240.

Lacroix, R., Wade, K.M., Kok, R., Hayes, J.F., 1995. Prediction of cow performance with a connectionist model. *Transactions of the American Society of Agricultural Engineers* 38, 1573–1579.

Mitchell, T.M., 1997. Machine Learning. 1st Ed. McGraw-Hill, New York, USA.

Mohammadabadi, M.R., Torabi, A., Tahmourespoor, M., Baghizadeh, A., Esmailizadeh, A.K., Mohammadi, A., 2010. Analysis of bovine growth hormone gene polymorphism of local and Holstein cattle breeds in Kerman province of Iran using polymerase chain reaction restriction fragment length polymorphism (PCR-RFLP). *African Journal of Biotechnology* 9, 6848-6852.

Mohammadi, A., Nassiry, M.R., Mosafer, J., Mohammadabadi, M.R., Sulimova, G.E., 2009. Distribution of BoLA-DRB3 allelic frequencies and identification of a new allele in the Iranian cattle breed Sistani (Bos indicus). *Russian Journal of Genetics* 45, 198-202.

OSIT., 2008. Office of Statistics and Information Technology, Ministry of Jihad-e-Agriculture. *Agricultural Statistics* 2, 91-93.

Pasandideh, M., Mohammadabadi, M.R., Esmailizadeh, A.K., Tarang, A., 2015. Association of bovine PPARGC1A and OPN genes with milk production and composition in Holstein cattle. *Czech Journal of Animal Science* 60, 97-104.

Raja, T.V., Ruhil, A.P., Gandhi, R.S., 2012. Comparison of connectionist and multiple regression approaches for prediction of body weight of goats. *Neural Computing and Applications* 21, 119-124.

Ramzan, S., Khan, S.I., 2010. Dimension reduction and remedy of multicolinearity using latent variable regression methods. *World Applied Science Journal* 8, 404-410.

Roush, W.B., Dozier, W.A., Branton, S.L., 2006. Comparison of Gompertz and neural network models of broiler growth. *Poultry Science* 85, 794–797.

Ruhil, A.P., Raja, T.V., Gandhi, R.S., 2013. Preliminary study on prediction of body weight morphometric of goats through ANN models. *Journal of the Indian Society of Agricultural Statistics* 67, 51-58.

Salehi, F., Lacroix, R., Wade, K.M., 1998. Improving dairy yield predictions through combined record classifiers and specialized artificial neural networks. *Computers and Electronics in Agriculture* 20, 199–213.

Shahinfar, S., Mehrabani-Yeganeh, H., Lucas, C., Kalhor, A., Kazemian, M., Weigel, K.A., 2012. Prediction of breeding values for dairy cattle using artificial neural networks and neuro-fuzzy systems. *Computational and Mathematical Methods in Medicine* 4, 1-9.

Silva, G.N., Tomaz, R.S., Sant'Anna, I.C., Nascimento, M., Bhering, L.L., Cruz, C.D., 2014. Neural networks for predicting breeding values and genetic gains. *Scientia Agricola* 71, 494-498.

Yang, X.Z., Wade, K.M., Lacroix, R., 2000. Investigation into the production and conformation traits associated with clinical mastitis using artificial neural networks. *Canadian Journal of Animal Science* 80, 415–426.

Communicating editor: *Omid Dayani*

پیش بینی ارزش‌های اصلاحی صفت تولید شیر در گاوهای هلشتاین ایران با استفاده از شبکه‌های عصبی مصنوعی

س. پورحمیدی[۱]، م. ر. محمدآبادی*[۱]، م. اسدی فوزی[۱] و ح. نظام آبادی پور[۲]

[۱]بخش علوم دامی، دانشکده کشاورزی، دانشگاه شهید باهنر کرمان، کرمان، ایران.

[۲]بخش مهندسی برق، دانشگاه شهید باهنر کرمان، کرمان، ایران.

***نویسنده مسئول، پست الکترونیک:** mmohammadabadi@yahoo.ca

چکیده شبکه‌های عصبی مصنوعی، الگوریتم‌های آموزش و روش‌های ریاضی که از توانایی فرآوری اطلاعات توسط مغز انسان تقلید می‌کنند می‌توانند داده‌های غیرخطی و پیچیده را به کار گیرند. هدف این پژوهش پیش بینی ارزش‌های اصلاحی صفت تولید شیر در گاوهای هلشتاین ایران با استفاده از شبکه‌های عصبی مصنوعی بود. از داده‌های مربوط به ۳۵۱۶۷ گاو هلشتاین ایران که بین سال‌های ۱۹۹۸ تا ۲۰۰۹ در مرکز اصلاح نژاد دام ایران رکوردبرداری شده بودند استفاده شد. ارزش‌های اصلاحی صفت تولید شیر با استفاده از مدل حیوانی تک صفته ASReml برآورد شدند. برای این که آموزش شبکه عصبی مصنوعی به نحو مطلوب انجام شود، از کل داده‌ها ۷۰ درصد برای آموزش، ۱۵ درصد برای تست و ۱۵ درصد برای اعتبارسنجی استفاده شدند. مدل به کار رفته در این تحقیق پرسپترون چند لایه (MLP) و الگوریتم مورد استفاده پس انتشار خطا بود. مدل پرسپترون سه لایه شامل یک لایه ورودی، یک لایه مخفی و یک لایه خروجی بود و مدل پرسپترون چهار لایه شامل یک لایه ورودی، دو لایه مخفی و یک لایه خروجی بود. موثرترین پارامترها برای صفات ورودی شامل پدر، گله، سال زایش، دو بار دوشش در روز، فصل زایش و سن به ماه بودند. ارزش‌های اصلاحی صفت تولید شیر به عنوان متغیر خروجی استفاده شدند. برای شبکه چهار لایه بهترین ساختار انتخاب شده برای صفت شیردهی اول شامل لایه ورودی با ۶ نرون، اولین لایه مخفی با ۱۶ نرون و ۱۸ تکرار، لایه مخفی دوم با ۶ نرون و ۱۵۴ تکرار و لایه خروجی با یک نرون بود. نتایج نشان داد که توانایی مدل شبکه عصبی مصنوعی بالاتر است و به ارزش‌های اصلاحی برآوردشده نزدیک‌تر هستند، بنابراین می‌توان از شبکه‌های عصبی مصنوعی به جای روش‌های معمول برای پیش بینی ارزش‌های اصلاحی صفت تولید شیر استفاده کرد.