Iranian Journal of Astronomy and Astrophysics
Vol. 3, No. 1, Spring 2016

DOI:10.22128/ijaa.2016.34

Iranian Journal of Astronomy and Astrophysics

Estimating the continuum of quasars using the artificial neural networks

Fatemeh Jafari[1] · Alireza Aghaee[2] · Seyed Masoud Barakati[3]

[1] Department of Physics, Faculty of Sciences, University of Sistan and Baluchestan, Zahedan, Iran; email: fjafari@staff.usb.ac.ir

[2] Department of Physics, Faculty of Sciences, University of Sistan and Baluchestan, Zahedan, Iran; email: aghaee@phys.usb.ac.ir

[3] Department of Electronics Engineering, Faculty of Electrical and Computer Engineering, University of Sistan and Baluchestan, Zahedan, Iran; email: smbaraka@ece.usb.ac.ir

Abstract. A lot of absorption lines are in the bluewards of Lyα emission line of quasar which is well-known as Lyα forest. Most of absorption lines in this forest belong to the Lyα absorption of the neutral hydrogen in the inter-galactic medium (IGM). For high redshift quasars and in the continuum with low and medium resolution, there are no many regions without absorption, so that, the quasar continuum in the forest is not obvious. Determination of the continuum in the forest is essential to study material distribution in the IGM, which is conductible through these absorption lines. One way to find this continuum is to predict it using longer wavelengths of the Lyα emission line of quasar, redwards of quasar continuum. Principal component analysis (PCA) method was proposed by researcher to estimate the bluewards of 50 low redshift quasars with 9% mean absolute error (error range was $3-30\%$). In this article, the whole continuum is predicted using only the redwards of the quasar continuum and ten random data of the forest by an artificial neural network (ANN). Five different training algorithms are used to train the ANN. The simulation results show that mean absolute error for the Lyα forest is decreased to 5.27% (with error range between $1.63-9.05\%$). These results verify the capability of the ANN to predict the quasar continuum in the Lyα forest as compared with the statistical methods.

Keywords: Quasar, Lyα, Forest, Neural Networks, continuum

1 Introduction

The observational continuum of the quasar has absorption lines most of which are due to the absorption of the neutral hydrogen Lyα existing in the line of sight of the quasar. Therefore, they are located within the wavelengths lower than the quasar emission line of Lyα. For the high redshift quasars, the number of these absorption lines is very high, so that the quasar continuum in this area (Lyα forest) is unidentifiable and a major limitation is how well the quasar continuum is known. An accurate determination of quasar continuum is also crucial for constraining the nature of the sources producing the UV background and the period of reionization in the Universe history. Other issues related to the physical state of the gas in the Lyα forest are also dependent on the continuum fitting; for example, the search for any deviation from Voigt-profile fitting which could have important consequences on our knowledge of the gas kinematics and thermal state. In spite of the many works published in the literature [17, 8, 15, 16, 4, 5, 2, 19, 11, 6, 1, 13, 14, 3], no consensus has been reached yet on what is the best method to determine the continuum. One robust used method is the principal component analysis (PCA), which is a technique used to emphasize variation

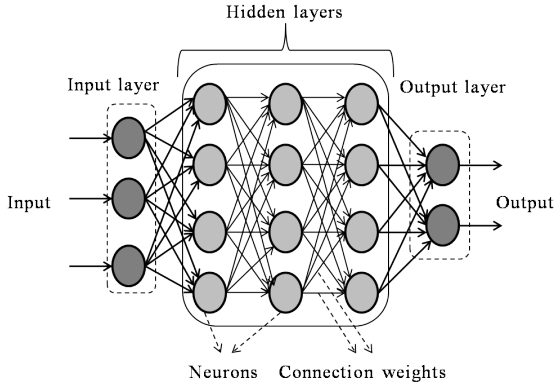

Figure 1: Typical representation of a feedforward MLP network with three inputs and two outputs.

and bring out strong patterns from a dataset. It can be used to estimate the continuum of quasars with their redwards [8, 19, 1]. The artificial neural network is a mathematical method designed based on inspiration by the biological neural network and learning concept of that. A feedforward (FF) multi-layer perceptron (MLP) is one of the most applicable neural networks. In this paper, the quasar continuum in the Lyα forest area is predicted using a feedforward multi-layer perceptron. The neural network is trained using data of the continuum in the wavelengths higher than the Lyα emission line of quasar (in redwards of the continuum which is easy to determine) and some limited random data of wavelengths lower than it (in the bluewards of the continuum). In the first try, the MLP is trained using 15 random data from the bluewards of the continuum; then, the number is reduced to 10 random data; subsequently, using the trained MLP, the whole continuum of the quasar is predicted.

2 Multi-Layer Perceptron (MLP)

A MLP is typically composed of a set of parallel and distributed processing units, called nodes or neurons. These are usually hierarchically arranged into layers, starting with the input layer and ending with the output layer. In between, a number of internal layers, also called hidden layers, provide most of the network computational power. The nodes appropriately interconnected by means of unidirectional (or bi-directional in some cases) weighted signal channels, called connections or synaptic weights [10]. Figure 1 shows a MLP network with three inputs, two outputs, and three hidden layers. Also, Figure 2 shows how a neuron works.

Learning is accomplished in general by developing algorithms that allow the system to learn for itself from a set of input/output training data. One major goal of learning algorithms is to combine the main features of a computing machine with those of human expertise to infer as many correct decisions as possible. An increasing number of systems are being designed today to have the very distinctive feature of learning. This is done by adjusting their parameters in response to unpredictable changes in their dynamics or their operating environment without the need for an explicit knowledge of a system model or rules that guide its behavior[10].

Through the selection of an adequate number of layers and neurons, which are often

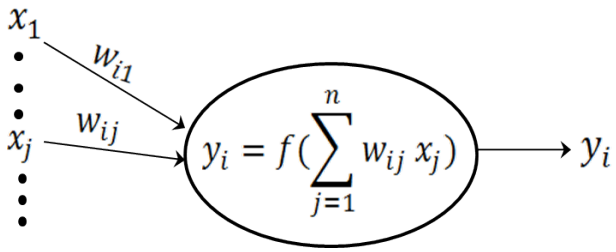

Figure 2: A typical neuron takes one or more input with some weights; then, it calculates output by activation function.This output is input for neurons of next layer with new weights.

limited in number, these networks are able to provide a non-linear mapping with the desired accuracy. The back-propagation (BP) training algorithm is a supervised learning method for multi-layered feedforward neural networks. It is essentially a gradient descent local optimization technique which involves backward error correction of network weights[12]. One try to adjust weights to close to target is one epoch.

There are a lot of article to improve training network. Most of methods to train MLP are based on first or second order Taylor series.The gradient descent method is an algorithm based on first order Taylor series and the Newton method is based on second order. The conjugate gradient method is an algorithm. It based on first order Taylor series so, it does not need compute second order of gradient; also, it searches in several direction conjugate to decrease error and this improves it better than gradient descent.

3 Sampling of quasar continuum

As aforementioned, a MLP artificial neural network is used to predict the quasar continuum. In this study, the MLP is trained using 50 sample quasars data that was observed by Hubble Space Telescope Faint Object Spectrograph (HST FOS). To have a comparison with previous researches, the used data are the same as which those used by Suzuki et al. [19, 18]. These 50 quasar continuum are a subset of the 334 high resolution HST FOS continuum and with S/N $>$ 10 per pixel, because weak emission line features can not be extracted in low S/N spectra. The redshift range of the continuum is from 0.14 to 1.04 with a mean of 0.58. For this range of redshift, the quasar continuum can be easily extracted because of low number of absorption lines in the forest. The quasar continuum is needed to verify the method of determining quasar continuum. The average S/N is 19.5 per 0.5 [A°] binned pixel[18]. Suzuki predicted the range of 1020 to 1600 angstrom of the continuum for the 50 quasars (the 50 samples studied in the paper) through principal component analysis (PCA) and using range of 1216 to 1600 angstrom (reawards) of the continuum which the mean absolute error was 9% and its range was from 3% to 30%[19].

Also, they defined five class for quasars. The first, second and third principal component spectra (PCS) account for 63.4, 14.5 and 6.2% of the variance respectively, and the first seven PCS take 96.1% of the total variance. Using the first two standardized PCS coefficients, they introduce five classifications: Class Zero and Classes 1-4. These classifications are shown in Figure 3. The first two PCS coefficients account for 77.9% of the variance and represent the overall shape of the quasar continuum.The first PCS carries Lyα, Lyβ and high ionization emission line features (OVI, NV, SiIV, CIV) that are sharp and strong. The second PCS has low ionization emission line features (FeII, FeIII,SiII, CII) that are broad and rounded[18].

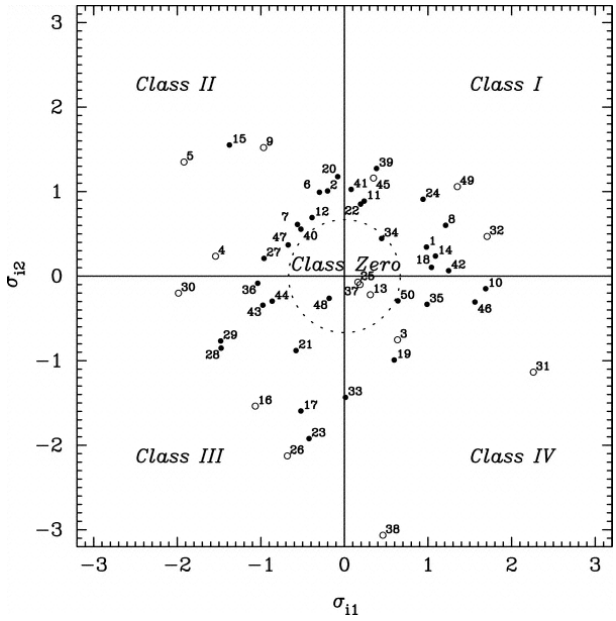

Figure 3: [18].The distribution of standardized first two PCS coefficients for 50 quasars in five classes.

4 Applying neural network to predict the quasar continuum

The strategy for finding the proper MLP neural network configuration can be explained as follow. In the first try, a MLP with one input, one output, and three hidden layers are chosen to predict the quasar Q003+1553. The proper results of first prediction, is a confirmation to apply the same configuration for some other quasars. In each stage, if the prediction faced with an unreasonable error, the configuration of the MLP will be modified based on changing the number of layers, neurons, and activation functions. The MLP configuration found with aforementioned strategy is consists of six hidden layers with 225, 225, 150, 70, 25, and 1 neurons, respectively. The activation function for the first hidden layer, other hidden layers, and output layer are tansig, logsig, and linear functions, respectively. After choosing the proper MLP configuration, it will be trained to predict the different classes of quasar continuum that having been classified in [18]. Five different training algorithmsare employed to train the MLP and the results are compared to find the best training algorithm for each class of the continuum. A proper structure MLP can be used for all quasars of a class because of similar shapes of quasars continuum in each class. The data for training, testing, and validation are extracted from five different classes of quasars continuum, including the data from the redwards of the continuum (769 data) and 15 random data in the bluewards of the continuum; Out of the extracted data, 76% data is used for training, 12% data for validation, and 12% data for testing of the MLP. The BP algorithm is used to train the MLP; to implement BP algorithm, different methods are possible in the MATLAB environment, consist of: Gradient descent (using Traingdx function), Newton (using Trainoss function), and Conjugate gradient (using functions, Traincgp, Traincgf, and Trainscg functions). To provide an accurate comparison between prediction of different quasars, the continuum prediction for all quasars is carried out using 15 selected data. In

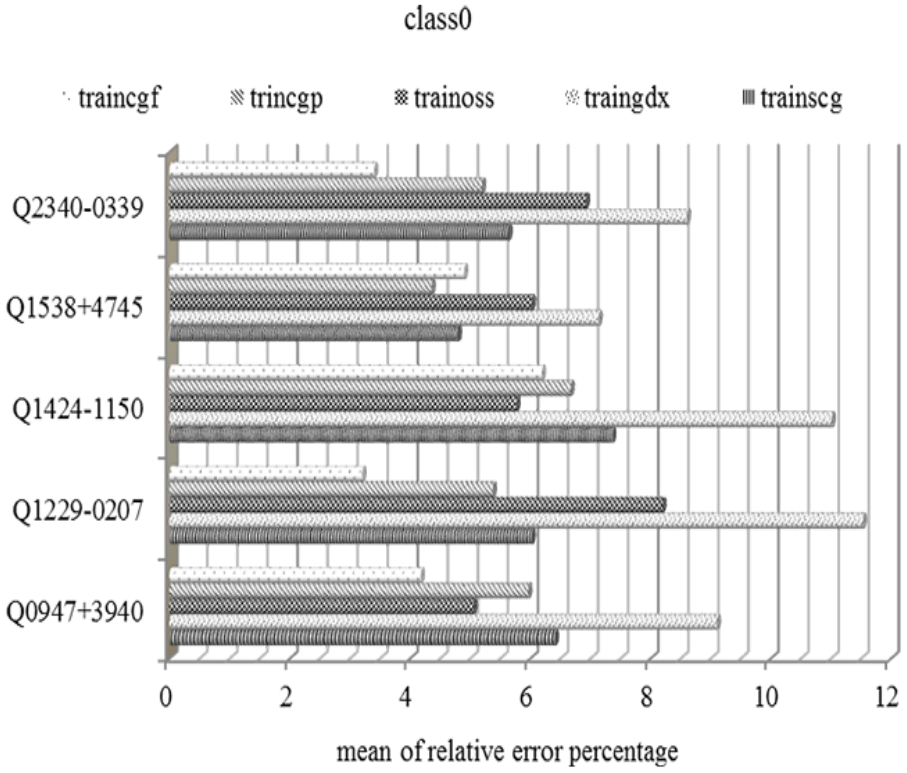

Figure 4: Results for class 0; vertical axis represents the mean of relative error percentage for each quasar and for each training algorithm; horizontal axis shows quasars name.

the begging of training stage, to increase the speed of training, the data are normalized to one. During the network training, if the average square of the error related to the training and validation data does not descend with an accurate rate, the algorithm will be restarted. The desired results may be obtained by performing the learning method after several iterations. In order to avoid the proximity effect of quasar, the Lyα forest of each quasars has been selected from the wavelength range of 2000 [km/s] (after the Lyβ emission line of the quasar) until 5000 [km/s] (before the Lyα emission line). The mean of relative error percentage for each algorithm is calculated for the selected region of the Lyα forest. By comparing the errors, the accurate algorithm for each class is selected. The errors of testing, validation, and training are analyzed to achieve an accurate trained network. The obtained results for the classes 0 to 4 are shown in Figures 4 to 8, respectively. As shown in the Figures, the class 0, Traincgf algorithm provides a considerably better response than the other algorithms. Both of Trainscg and Traincgf algorithms have a lower relative mean error for the class 1. In the classes 2 and 3, the Traincgf algorithm yields a better response; in class 4, the Traincgp algorithm has the least mean error, although the result is close the Traincgf algorithm.

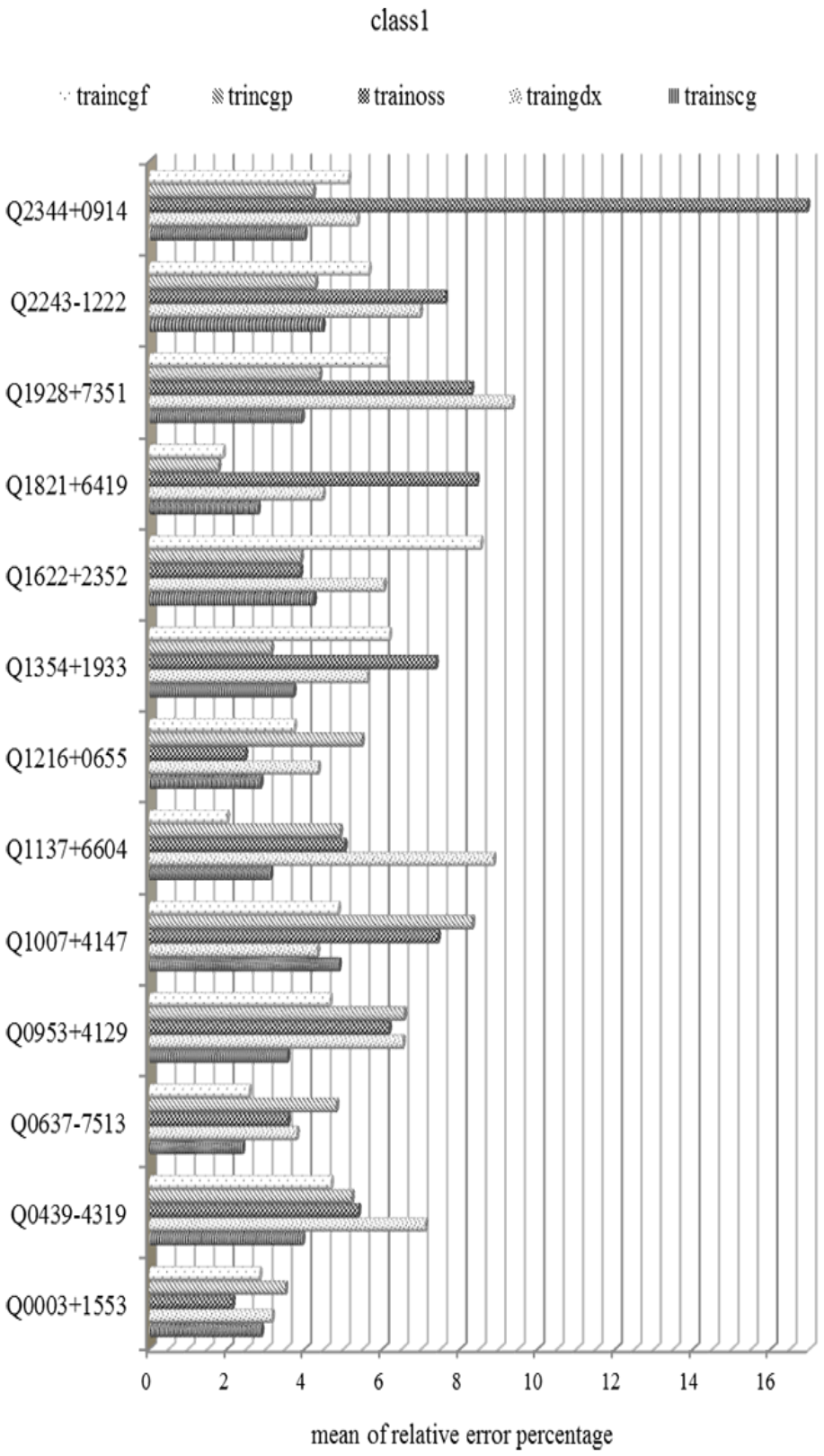

Figure 5: Results for class 1; vertical axis represents the mean of relative error percentage for each quasar and for each training algorithm; horizontal axis shows quasars name.

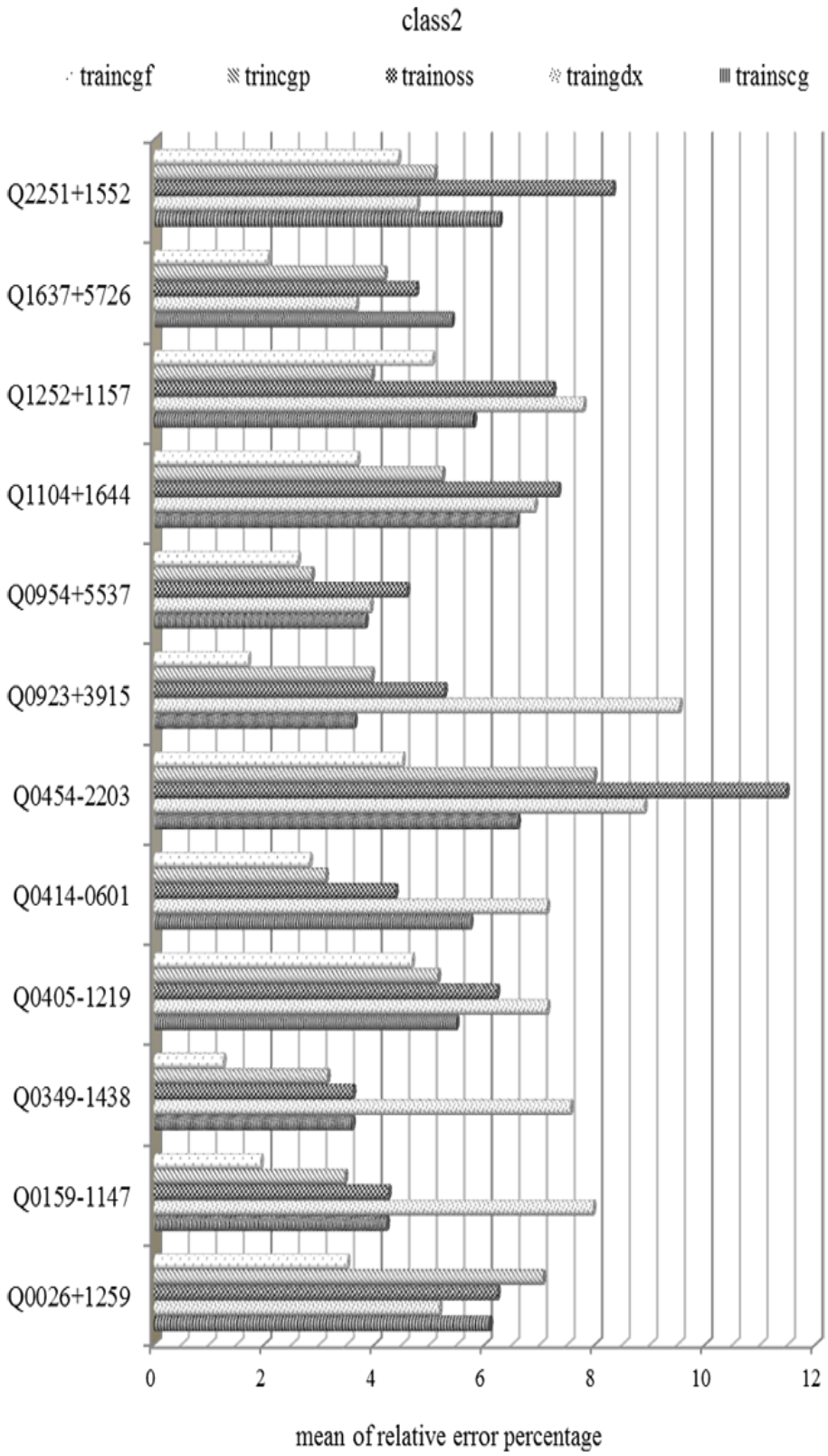

Figure 6: Results for class 2; vertical axis represents the mean of relative error percentage for each quasar and for each training algorithm; horizontal axis shows quasars name.

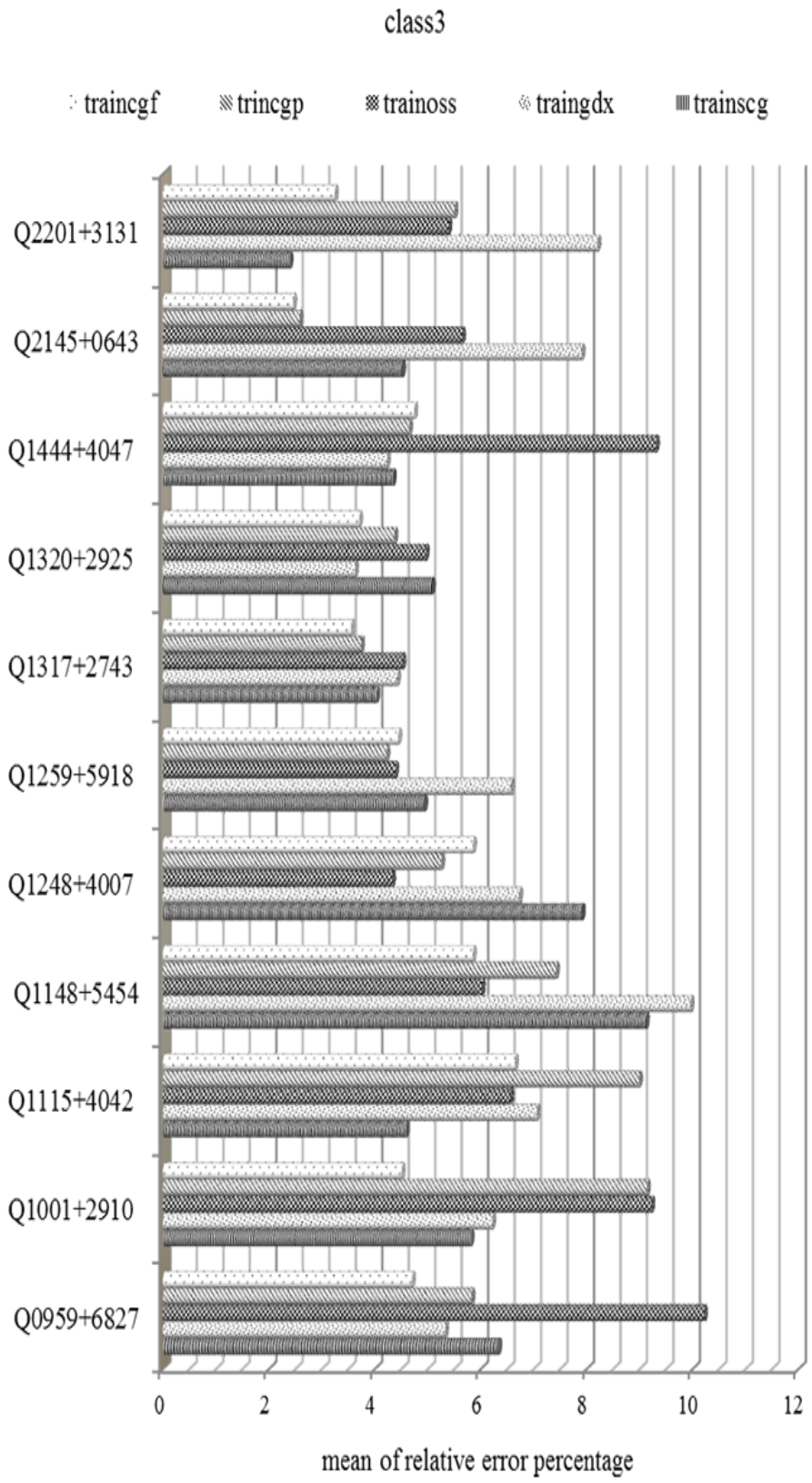

Figure 7: Results for class 3; vertical axis represents the mean of relative error percentage for each quasar and for each training algorithm; horizontal axis shows quasars name.

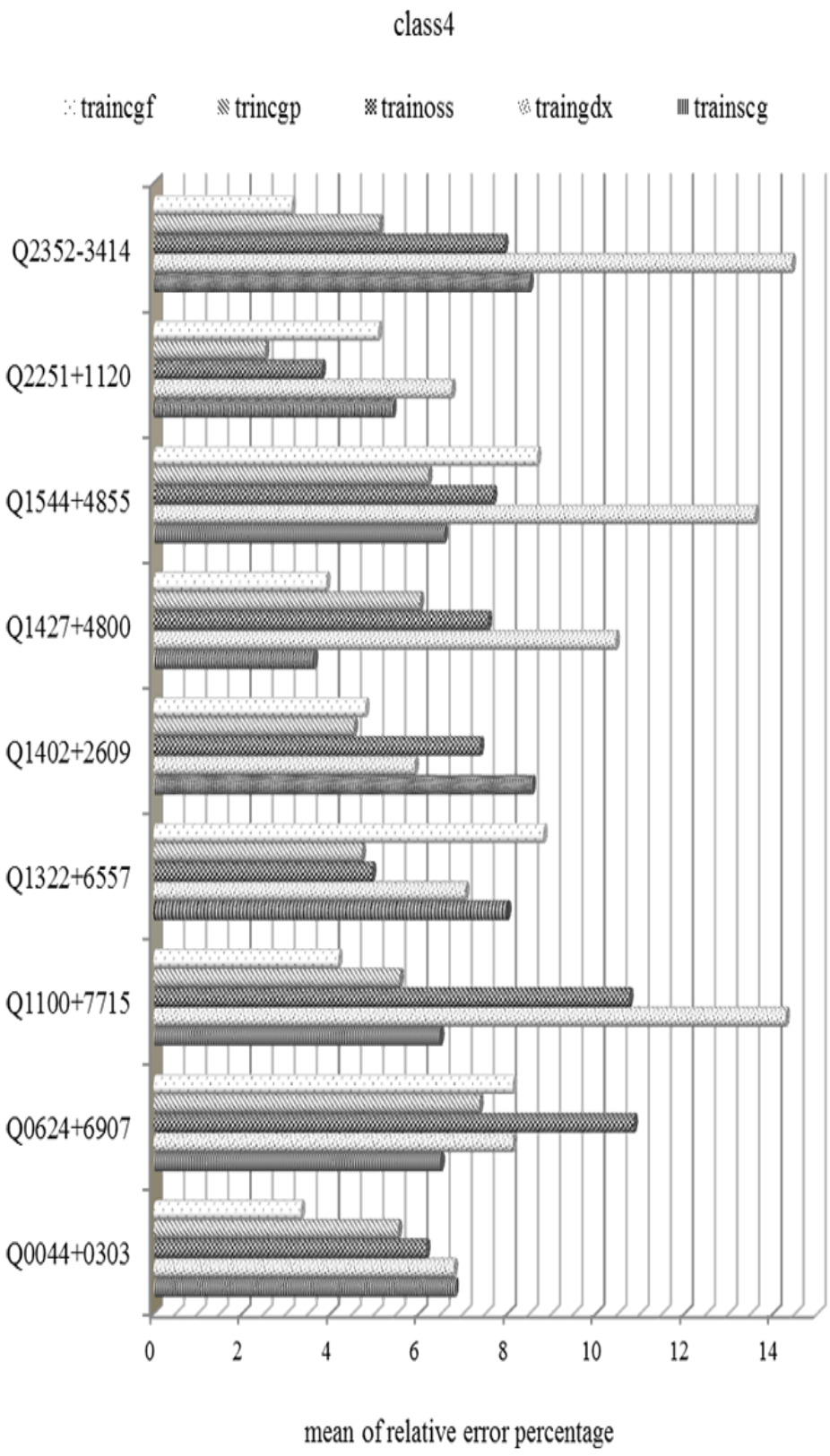

Figure 8: Results for class 4; vertical axis represents the mean of relative error percentage for each quasar and for each training algorithm; horizontal axis is shown quasars name.

5 Predicting the quasar continuum using least random data from the bluewards of the continuum

In previous section, a neural network was trained using 15 data from bluewards of the quasar continuum. In this section, alongside decreasing the number of the random data to 10, a MLP neural network structure with less hidden layers is used. Additionally, in training stage, the data are presented to the MLP for a maximum of 1200 times, while in previous network, this number was 3000. For the training of the network, Traincgf algorithm provides a better response for total 50 quasars is used, as it is shown in the Figures 4 to 8. The neural network that is used for this section, contains five hidden layers, with 225, 150, 75, 25 and 1 neurons. The activation functions are Tansig for the first layer, Logsig for the other ones, and the linear for the output layer. Of all the data, i.e., the data on the bluewards of the continuum (including 769 data) and of the 10 random data on the bluewards, 85% is used for the training and 15% for the validation of the network. To provide an accurate comparison of prediction results for different quasars, first 10 random data are selected and, then, the continuum is predicted for all the quasars using these 10 data. Before starting the network training, data are normalized to one. The target is to have the average square of the error in the training and validation lower than 5×10^{-6} and as close as to 10^{-6}. Training algorithm is run for many times for every quasar to achieve our target. Subsequently, the percentage of the relative mean error is calculated for the selected Lyα forest of each quasars. The result of mean of relative error percentage for every quasar is shown in Figure 9, where the mean of relative error percentage for the Lyα forest is between $1.63\% - 9.05\%$. Real and estimated continuum of the quasar Q1928+7351 are shown in Figure 10.

6 Predicting the deviation of the quasar continuum compared to average of continuum in some other quasars

At this stage, the average data of the first 45 quasars continuum are considered and the deviations in the continuum of five other quasars are predicted using the MLP neural network introduced in the section 5. In this case, the amount of the deviation from the average of the continuum of 45 other quasars (from the redwards of the continuum) and the deviation in the continuum of the quasar in 10 random data (from the bluewards of the continuum) with the amount of the average are presented to the MLP, so that 85% of data is used for the training and 15% for the validation. Then, the deviation from the average is predicted for the whole continuum and the relative mean error percentage is calculated within the range of 1023 to 1195.5 angstroms. The results are shown in Figure 11, where mean of relative error percentage for every quasar is less than 4%.

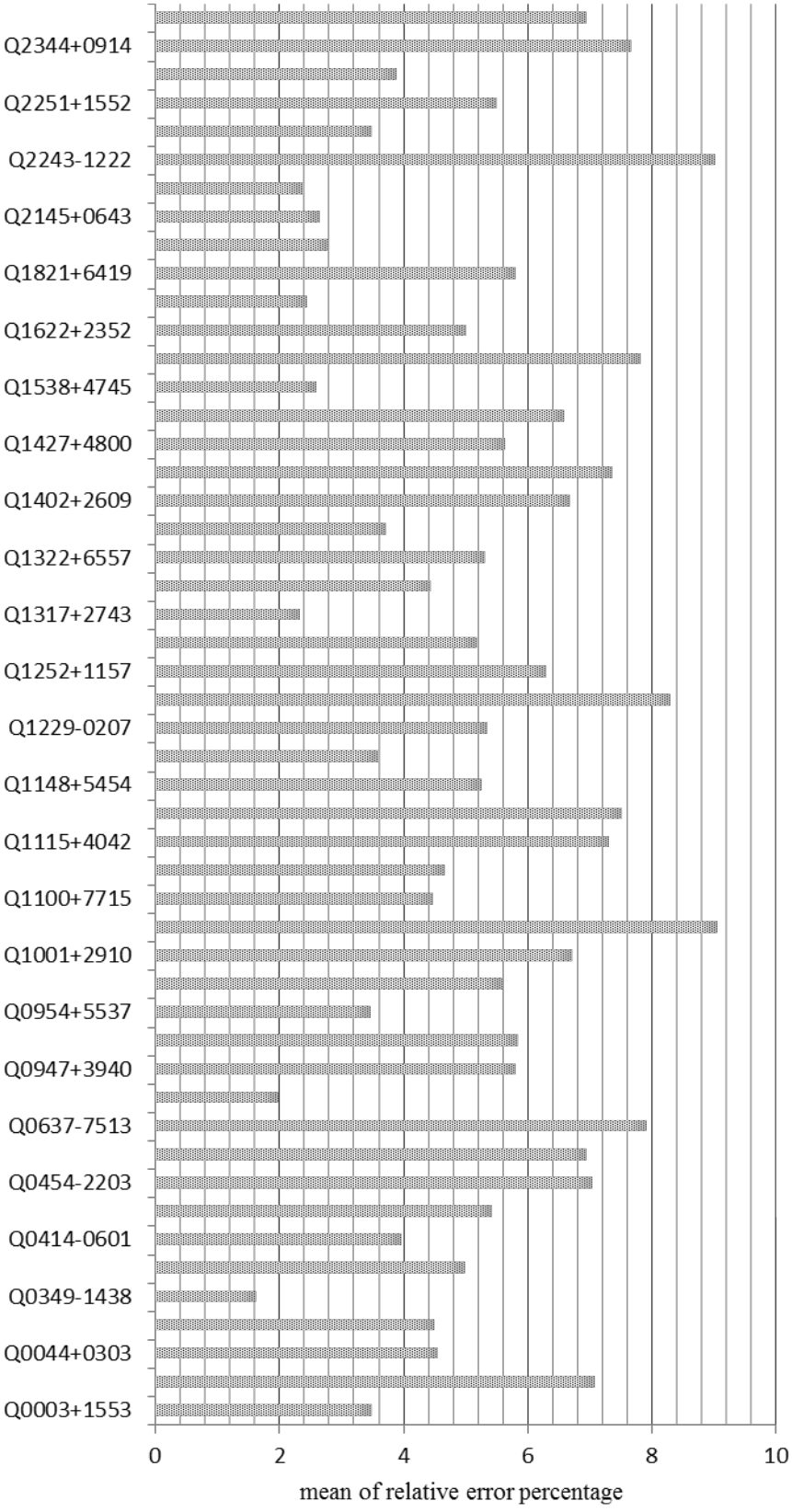

Figure 9: Horizontal axis represents the mean of relative error percentage for each quasar; vertical axis shows quasars name.

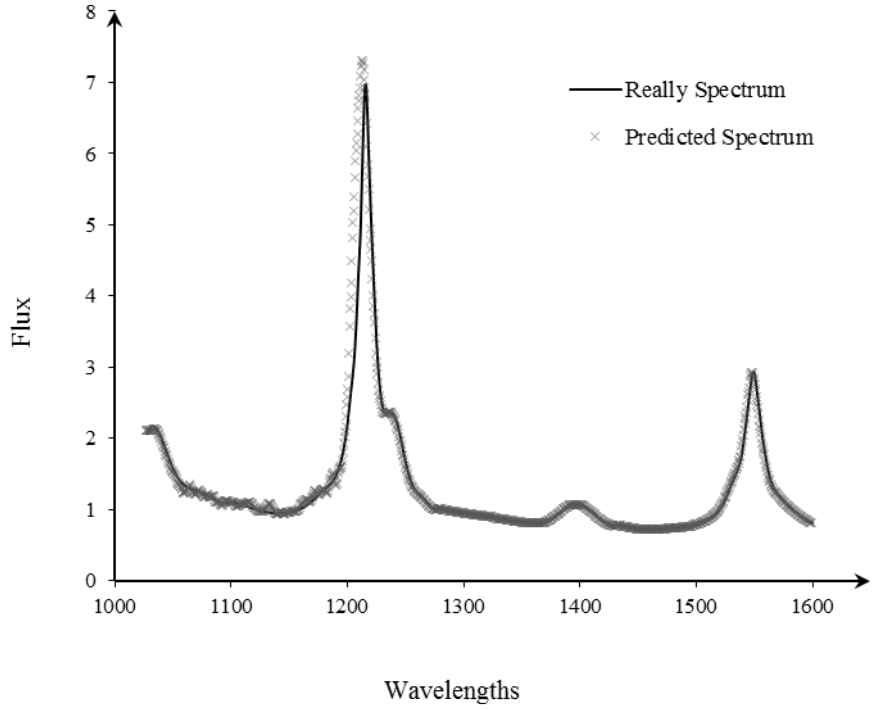

Figure 10: Actual and estimated neural networks for quasar continuum Q1928+7351 in the region 1020 to 1600 [A°].

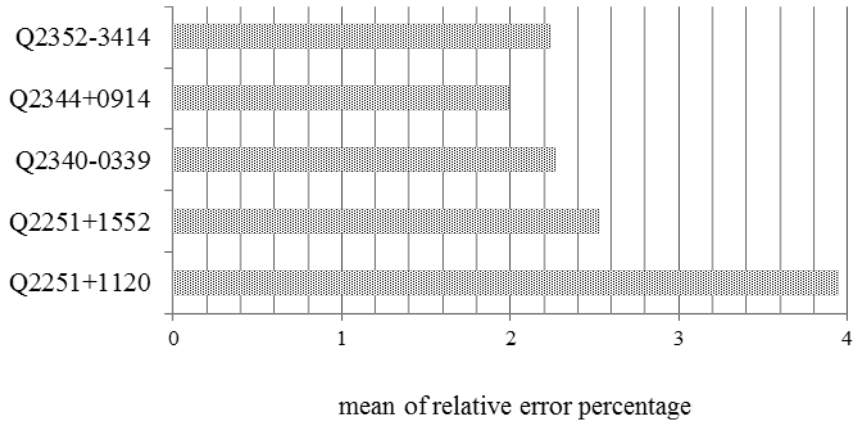

Figure 11: Horizontal axis represents the mean of relative error percentage for each quasar; vertical axis shows quasars name.

7 Conclusion

The quasar continuum in the Lyα forest is predicted using a multi-layer perceptron neural network. The neural network is trained for the wavelengths higher than the Lyα, using data obtained from a quasar continuum and also some random data from the wavelengths lower than the Lyα. The mean error for 50 quasars is 5.27% and error range varies from 1.63% to 9.05%. The mean error of PCA method for this sample of quasars [19] is 9% and error range varies from 3% to 30%. This comparison demonstrates that the neural network using some data from the blueward, can predict the quasar continuum fairly good. It is an open research for future works.

References

[1] Aghaee A., Petitjean P., Srianand R., Stalin C. S., Guimaraes R., 2010, J. Astrophys. Astr. 31, 59

[2] Aracil B., Petitjean P., Pichon C., Bergeron J., 2004, Astron. Astrophys. 419, 811

[3] Becker G. D., Hewett P. C., Worseck G., Prochaska J. X., 2013, MNRAS 430, 2067

[4] Becker R. H., Fan X. , White R. L., et al. , 2001, Astron. J. 122, 2850

[5] Bernardi M., Sheth R. K., Subba Rao M., et al., 2003, Astron. J. 125, 32

[6] Faucher-Giguère C.-A., Prochaska J. X., Lidz A., Hernquist L., Zaldarriaga M., 2008, Astron. J. 681, 831

[7] Fletcher R., Reeves C. M., 1964, Comput. J. 7, 149

[8] Francis Paul J., Hewett Paul C., Foltz Craig B., Chaffee Frederic H., 1992, Astron. J. 398, 476

[9] Hagan M. T., Demuth H.B., Beale M., 1996, pws publishing, first published

[10] Karray F. O., Silva C. D., 2004, New York, Harlow, Pearson Education Limited,First published

[11] Kirkman D., Tytler D., Lubin D., Charlton J., 2007, MNRAS 376, 1227

[12] Nawi N. M., Ransing R. S., Ransing M. R., 2007, Int. J. Comput. Intell. 4, 46

[13] Pâris I., Petitjean P., Rollinde E., et al. 2011, Astron. Astrophys. 530, 15

[14] Pâris I., Petitjean P., Aubourg É., et al., 2012 Astron. Astrophys. 548, 28

[15] Press W. H., Rybicki G. B., Schneider D. P., 1993, Astron. J. 414, 64

[16] Rauch M., 1998, Annu. Rev. Astron. Astrophys. 36, 267

[17] Schmidt M., Schneider D. P., Gunn J. E., 1991, APS Confer. Ser. 21, 39

[18] Suzuki N., 2006, Astrophys. J. Suppl. Ser. 163, 110

[19] Suzuki N., Tytler D., Kirkman D., O'Meara J. M., Lubin D., 2005, Astron. J. 618, 592