Avrupa Bilim ve Teknoloji Dergisi
Sayı 18, S. 465-475, Mart-Nisan 2020

Araştırma Makalesi

www.ejosat.com ISSN:2148-2683

European Journal of Science and Technology
No. 18, pp. 465-475, March-April 2020

Research Article

Artificial Neural Network (ANN) Approach for Dynamic Viscosity of Aqueous Gelatin Solutions: A Soft Computing Study

Barış Demirbay[1*], Buse Karakullukçu[2], Fatma Gülay Acar[3]

[1]İstanbul Technical University, Institute of Science and Technology, Physics Engineering Program, 34469, İstanbul, Turkey (ORCID: 0000-0002-5454-7437)
[2]Kocaeli University, Faculty of Engineering, Department of Mechanical Engineering, 41001, Kocaeli, Turkey (ORCID: 0000-0002-3655-0931)
[3]İstanbul Technical University, Faculty of Science and Letters, Department of Physics Engineering, 34469, İstanbul, Turkey (ORCID: 0000-0001-6831-7635)

(İlk Geliş Tarihi 28 Ocak 2020 ve Kabul Tarihi 9 Mart 2020)

(**DOI:** 10.31590/ejosat.680773)

ATIF/REFERENCE: Demirbay, B., Karakullukçu, B. & Acar, F. G. (2020). Artificial neural network (ANN) approach for dynamic viscosity of aqueous gelatin solutions: a soft computing study. *Avrupa Bilim ve Teknoloji Dergisi*, (18), 465-475.

Abstract

In this research, we present a multi-layered feed-forward neural network (ANN) model developed for prediction of dynamic viscosity of aqueous gelatin solutions using experimental data collected from a number of measurements. In ANN architecture, shear stress, shear strain, torque of spindle, the angular velocity of spindle together with mass concentrations of gelatin solutions were introduced as input neurons, whereas dynamic viscosity of aqueous gelatin solutions was assigned as a single output neuron to be predicted. Developed ANN model was trained using backpropagation algorithm optimized with Bayesian regulation. Optimal geometry of the hidden layer was first studied to search out the ANN architecture which yields the most accurate performance results. Mean squared error (MSE), mean absolute error (MAE), root-mean-squared error (RMSE), determination of coefficient (R^2), the variance accounted for (VAF) and regression analyses were used as performance assessment parameters for suggested network models. Sensitivity analysis was carried out to investigate the most effective input neuron strongly influencing the performance of the developed ANN model. As a result, the use of 8 neurons in the hidden layer has shown excellent performance results yielding the least MSE and the highest R^2 values compared to other suggested ANN models. Upon sensitivity analysis, the shear rate was found to be the most effective input neuron significantly affecting network performance. ANN-based predicted dynamic viscosity values were found to be in excellent agreement with measured viscosity values, demonstrating the robustness as well as the accuracy of the developed ANN model. Developed ANN model can, therefore, be effectively used to predict the dynamic viscosity of aqueous polymer solutions using the same input and output parameters in specific data range reported in this paper with statistical details.

Keywords: Artificial intelligence, Machine Learning, Artificial neural network, Bayesian regulation, Regression, Relative importance, Statistical analysis, Gelatin, Biopolymer, Rheology, Dynamic viscosity.

Jelatin Çözeltilerinin Dinamik Viskozitesine Yapay Sinir Ağı (YSA) Yaklaşımı: Esnek Hesaplama Çalışması

Öz

Bu araştırmada, bir dizi ölçümden toplanmış deneysel veriyi kullanarak jelatin çözeltilerinin dinamik viskozitesini tahmin etmek üzere geliştirilen çok katmanlı ileri beslemeli bir yapay sinir ağı modeli (YSA) sunuyoruz. YSA yapısında, kayma gerilmesi, kayma oranı, mil torku, mil açısal hızı ile birlikte jelatin çözeltilerinin kütle konsantrasyonu giriş nöronları olarak tanıtılırken, jelatin çözeltilerinin dinamik viskozitesi tahmin edilmek üzere tek bir çıkış nöronu olarak kullanılmıştır. Geliştirilen YSA modeli, Bayesian regülasyonu ile optimize edilmiş geri yayılım algoritması kullanılarak eğitilmiştir. İlk olarak, en doğru performans sonuçlarını veren YSA yapısını bulmak üzere gizli katmanın optimal geometrik yapısı çalışılmıştır. Önerilen ağ modelleri için ortalama karesel hata (MSE), ortalama

[*] Corresponding Author: İstanbul Technical University, Institute of Science and Technology, Physics Engineering Program, 34469, İstanbul, Turkey ORCID: 0000-0002-5454-7437, demirbay@itu.edu.tr

mutlak hata (MAE), ortalama kare hatalarının karekökü (RMSE), determinasyon katsayısı (R^2), varyans (VAF) ve regresyon analizleri performans değerlendirme parametreleri olarak kullanılmıştır. Geliştirilen YSA modelinin başarı performansını etkileyen en etkin giriş nöronunu araştırmak amacıyla duyarlılık analizi yapılmıştır. Sonuç olarak, gizli katmanda 8 nöronun kullanılması, önerilen diğer YSA modellerine kıyasla en düşük MSE ve en yüksek R^2 değerlerini vererek en yüksek başarı performansını göstermiştir. Duyarlılık analizinin sonucu olarak, kayma oranı oluşturulan sinir ağının başarı performansını etkileyen en etkin giriş nöronu olarak bulunmuştur. Tahmin edilen dinamik viskozite değerlerinin, ölçülen dinamik viskozite değerleriyle büyük bir uyum içinde olması, geliştirilen YSA modelinin doğruluğunu ve güvenilirliğini ispatlamıştır. Bu nedenle geliştirilen YSA modeli, bu araştırmada istatistiksel detayları verilen veri aralığındaki giriş ve çıkış parametrelerini kullanarak, polimer çözeltilerinin dinamik viskozitesini tahmin etmek için efektif bir kullanım sağlamaktadır.

Anahtar kelimeler: Yapay Zeka, Makine Öğrenmesi, Yapay Sinir Ağları, Bayesian Regülasyonu, Regresyon, Duyarlılık Analizi, İstatistiksel Analiz, Jelatin, Biyopolimer, Reoloji, Dinamik Viskozite.

1. Introduction

Gelatin is a water-soluble natural biopolymer that can be obtained by the chemical degradation or thermal denaturation of collagen taken from animals. Chemical structure of gelatin is essentially composed of 14% hydroxyproline, 16% proline and 26% glycine, respectively (Wang et al., 2019). Production of gelatin increases with a rapid trend since it can be easily extracted from the skin, tendons, cartilage and bones of different animals, particularly fish and porcine (Char et al., 2019). It has a wide array of applications not only in the food industry but also in clinical studies. Due to its biocompatibility, biodegradability, as well as low cost, much attention has recently been focused on the applications of gelatin mainly in biomedical research methods including tissue engineering, drug delivery and wound dressing (Gullapalli, 2010). Particularly, gelatin is well known with its natural chemical cross-linking ability and it is, therefore, being used as a cross-linker to trap pharmaceutical molecules in drug delivery systems (Foox & Zilberman, 2015). In such systems, for transportation of drugs to target regions in the human body, viscosity of substance as well as carrier medium plays a significant role to ensure homeostasis.

Depending on the age, breed and partial hydrolyzation of the animal from which the collagen is obtained, the resultant gelatin may have different physicochemical and rheological properties that affect its efficiency and workability as well. In this respect, viscosity is one of the most significant parameters to measure both the grade and quality of gelatin-based substances. The viscosity of such natural chemical cross-linkers including gelatin can be controlled by varying a degree of cross-linking and sample concentration. However, the viscosity measurements of gelatin-based solutions, in the presence of other additives or at higher concentrations, are not that simple since the flow characteristics can quickly change over time due to formation of macromolecules which are strongly bound to each other by long peptide chains (Osorio, Bilbao, Bustos, & Alvarez, 2007). It is also known that the rheological characterization of polymer-based solutions requires overtime work as much as a joint effort of researchers from many different backgrounds. In that sense, the applications of machine learning methods, particularly artificial neural networks (ANNs), are recently of utmost importance to researchers to model the most challenging, long-time running experimental problems using earlier measured data and to predict potential results for unknown data sets instead of performing new measurements.

Today, the implementation of ANNs to estimate measurement parameters including the viscosity of materials from different classes has become one of the best soft computing methods due to the efficiency and reliability of ANNs. In accordance with the literature review, there are several successful neural networks have been developed to predict the viscosity of polymer solutions using a different set of parameters (Aminian, 2017; Fatehi, Raeissi, & Mowla, 2017; Hemmat Esfe & Abbasian Arani, 2018). Other than these contributions, there are only a few ANN studies addressing the dynamic viscosity of aqueous biopolymer solutions. However, the dynamic viscosity of gelatin solutions has not been studied by means of ANNs yet. For this reason, the main goal of this study is to develop a multi-layered feed-forward neural network (FFNN) to forecast the dynamic viscosity of aqueous gelatin solutions using a large set of experimental data gathered from rheological measurements. Information regarding ANN present with further details in the following section.

2. Artificial Neural Network (ANN)

Artificial neural network (ANN) is a mathematical analog which can examine linear and non-linear complex relationships between input and output data by learning algorithm (Haykin, 1994). To develop a simple ANN model for any particular system, one needs to define a topology of the network, the training algorithm as well as activation functions, respectively. In each ANN architecture, the simplest processing element is called an artificial neuron, which imitates the behavior and the functions of biological neurons in the human brain where large amounts of information can be stored and processed simultaneously by neurons (Akkoyun, Yildiz, & Kaya, 2019). ANN topology basically consists of a number of artificial neurons, similar to biological neurons. ANN has an input layer and an output layer with different numbers of neurons, which are interconnected to each other by one or more hidden layers. Among the various ANN architectures, the most well-known is a feed-forward neural network (FFNN), which has three distinct layers with several neurons: the input layer, the hidden layer(s) and the output layer (Pal & Chakraborty, 2005). In order to illustrate a simple topology of ANN, the ANN architecture used in this research article is present in Figure 1.

In FFNN, interconnection between each neuron is restricted just to one direction, meaning that information flows from the input layer to the output layer by passing through the hidden layer(s) by using neurons in each layer (Asteris, Roussis, & Douvika, 2017). The first layer in the network topology is known as the input layer where data from the outside world is kept by input neurons. Data stored in the input layer is then passed through the neurons in the second layer, called the hidden layer, to be processed using an

activation function. The hidden layer plays a very significant role in prediction performance since it allows the neural networks to examine the logic of complex relationships between the inputs and the outputs. Based on the complexity of the prediction problems, more than one hidden layer can also be used in the network topology. However, in most of the engineering problems, one hidden layer with an optimal number of neurons gives excellent prediction results. The last layer of the network is the output layer which receives processed information from the neurons in the hidden layer(s) and yields a response of network topology to the inputs. In FFNN, inter-connected nodes between each neuron have weights that strongly affect the output response of the network model.

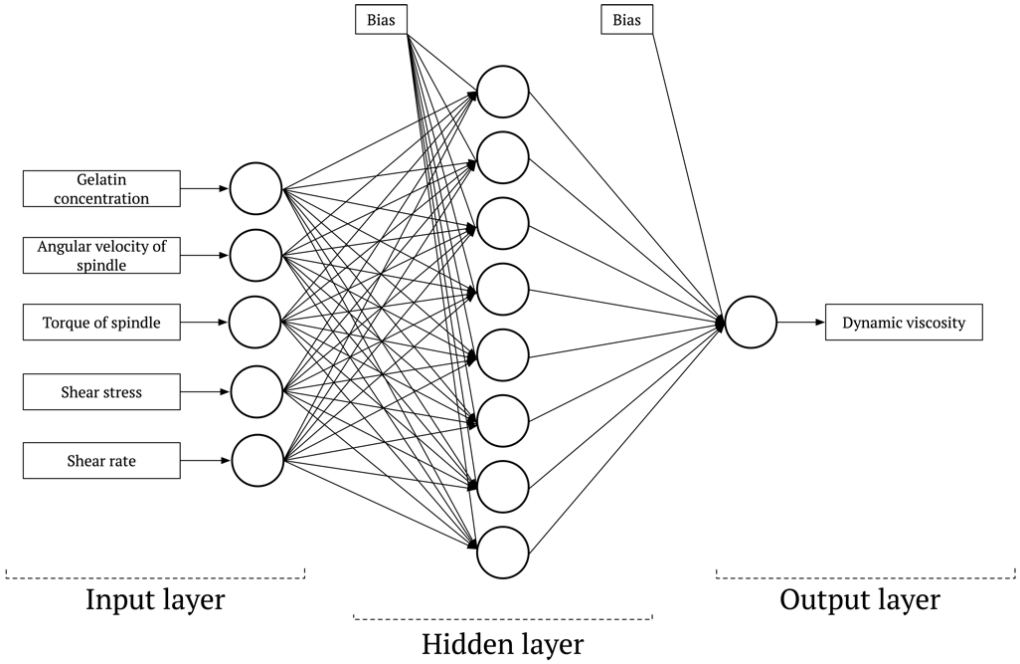

Figure 1. A basic topology of artificial neural network (ANN)

In order to achieve successful prediction results using the ANN approach, the most important step is to train neural network optimally. There are many different types of training algorithms available in the literature used for training ANNs. The most common and effective algorithm is backpropagation, which involves two passes. In the forward pass, data in the input layer are fed to ANN. Signals in the input layer propagate towards the network topology in order to predict an output value in the output layer based on the weights, biases (only connected to the neurons both in the hidden and the output layers) and processing functions of neurons. Predicted output value is then compared with the real output value of the assigned input-output data set when the generated signal reaches the output layer. To evaluate the performance of the backpropagation algorithm, the error between real and predicted output values is computed, and the error signals are then sent backward from the output layer to the input layer by passing through the hidden layer(s) (Erdil & Arcaklioglu, 2013). One pass through the set of training patterns by updating the weights is called epoch. Weights and bias values are updated simultaneously at the end of each epoch. As also seen from Figure 2, training takes place repeatedly through gradient descent on the sum of squares of the errors for all training patterns until these errors are fully minimized (Goh, 1995). After training process is successfully done, all final weights and bias values are stored in ANN memory and the neural network is then able to reproduce target output values that should close enough to real output values.

In order to better minimize the error by updating interconnection weights, an effective optimization method must be carried out in the backpropagation algorithm. There are many different back-propagation algorithms such as Levenberg-Marquardt, Bayesian regulation, scaled conjugate gradient and so forth, which have been used in training of different neural network structures. The trained ANN model has to be tested by a different data set, which are not introduced to neural network during the training process. A perfectly trained network model is supposed to give nearly similar training and testing errors. For this reason, after training and testing phases, neural network is usually validated with a different data set in order to prevent over-fitting when standard optimization methods such like Levenberg-Marquardt algorithm are applied in backpropagation algorithm (Ghatak & Robi, 2018). Recent studies have shown that the optimization of backpropagation using Bayesian regulation is more robust compared to other standard optimization methods as it eliminates the need for lengthy cross-validation (Burden & Winkler, 2009). It is also known that the larger the weights, the higher the error. Bayesian regulation backpropagation minimizes the sum of squared errors together with weights between neurons and bias values to allow the network to generalize very well. Unlike Levenberg-Marquardt optimization, Bayesian regulation optimization does not require an additional data set for validation step and there is no risk of over-fitting (Demirezen & Fung, 2019; Kayri, 2016). Therefore, in this study, Bayesian regulation has been used as an optimization method in back-propagation algorithm during training of developed network model.

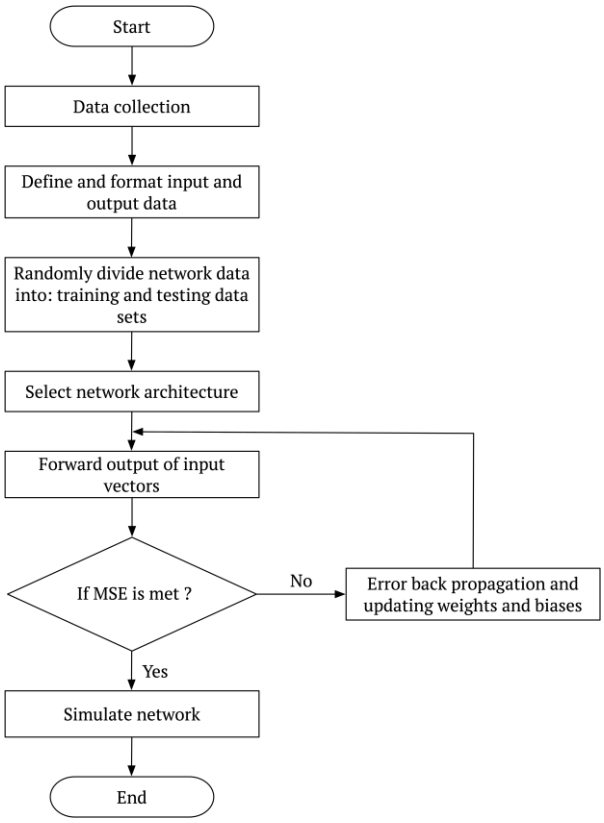

Figure 2. Flow-chart illustration of back-propagation algorithm

3. Materials & Methods

3.1. Preparation of aqueous gelatin solutions

Gelatin (bloom, purchased from Doğa drug industry) powder was used as a main substance to prepare all aqueous solutions. Different amounts of gelatin powders were weighed using a weighing vessel and then added into glass beakers filled with distilled water (dH_2O). All aqueous solutions were homogeneously mixed on a magnetic stirrer with a constant speed of 1200 rpm at 40 °C for 30 min. In total, 6 different aqueous gelatin solutions, each having the same sample amount of 18 ml and different mass concentrations in the range between 0.66 wt/vol% and 5 wt/vol%, were prepared in separate glass beakers and cooled down to the room temperature. The sample preparation process has been done at room temperature about 21 ± 2 °C and each sample solution was freshly prepared just before measurements on the same day. All experiments were performed in cross-linked materials research laboratory in the department of physics engineering at Istanbul Technical University (ITU).

3.2. Dynamic viscosity measurements

Dynamic viscosity measurements of prepared aqueous gelatin solutions were performed using a rotational rheometer (Fungi-Lab premium series) equipped with a heat control unit. All solutions were separately dropped into a removable (stainless-steel) sample container of rheometer, which has 18 ml volume capacity, where measuring spindle is attached and rotates to detect viscous forces of fluid against rotation direction of the spindle. During the measurements, prepared solutions within container were isolated from an external environment using a special component called solvent trap, which restricts the evaporation of the sample. Using a rotational rheometer, the dynamic viscosity of solutions can be determined by detecting shear stress and shear rate values. Shear stress is mainly associated with measured torque and also geometrical shapes of the spindle whereas the shear rate is calculated from the system's geometry as well as the angular velocity of the spindle. In our experiments, coaxial cylinder spindle (LCP) was used to measure the dynamic viscosity of solutions at room temperature (294.15 K). In measurements, not only dynamic viscosity was measured, but also other data sets including shear stress, shear rate, torque of spindle as well as the angular velocity of spindle were recorded to be used in the ANN development. Information regarding the ANN model is given with further details in the following section.

3.3. Neural network model

ANN architecture is designed to predict the dynamic viscosity of gelatin solutions by using several different rheological measurement parameters and mass concentrations of prepared aqueous sample solutions. In the developed ANN model, mass concentration of gelatin solution (C), velocity of spindle (w), torque of spindle (M), shear rate ($\dot{\gamma}$) and shear stress (τ) were used as input

neurons, whereas dynamic viscosity (η) was used as a single output neuron. Statistical details of input and output variables are present in Table 1.

Table 1. Statistical details of input and output neurons employed in ANN model

Network parameters	Variables	Units of variables	Properties of data set used in network topology			
			Minimum	Maximum	Mean	Standard Deviation
Input	C	wt/vol%	0.66	5	82.2059	42.1648
Input	w	rad/s	10	220	102.8431	62.6567
Input	M	N.m	2.3481	98.8399	42.7886	28.3242
Input	$\dot{\gamma}$	s^{-1}	12.23	269.06	125.7772	76.6292
Input	τ	N/m^2	0.1723	7.8513	3.1390	2.0777
Output	η	cP	1.3427	13.2926	2.8513	2.2896

Many different numbers of dynamic viscosity measurements were performed using aqueous gelatin solutions at different mass concentrations. In total, 204 separate experimental data were collected to be used in the ANN development. All collected experimental data were then randomly divided into two different sub-datasets: a training data set, to develop the network topology, and an independent testing data set, to be predicted and then compared with measurement data. 70% of the data (143 data set) were used in network training and the remaining 30% (61 data set) were used in ANN testing, respectively. Design of ANN architecture, development of network model, and all calculations regarding performance indices were computed in Matlab® (MathWorks, version 2017-b) environment and all figures and plots were prepared using Python. A single hidden layer in ANN topology is often sufficient to get high prediction quality when optimal geometry is maintained by varying the number of neurons in that of the hidden layer. For this reason, different number of neurons has been employed in the hidden layer to find out successfully predicted dynamic viscosity values. Since each input and output variable used in ANN development have different orders of magnitude, all experimental data must be normalized to be at a comparable range. Therefore, both input and output data were normalized separately in forms of single vectors in the range between 0 and 1, by using the equation (1):

$$x_i^* = \frac{x_i - x_{min}}{x_{max} - x_{min}} \tag{1}$$

The selection of activation functions that are used in the hidden and the output layers remarkably affects the performance of the developed ANN model. Therefore, several different activation functions, such as logistic-sigmoid, hyperbolic sigmoid and linear functions have been tested to investigate which one of them yields the best performance of the network model. Over-fitting is a major problem typically arising during the training stage. The use of suitable optimization methods, such as Levenberg-Marquardt, Bayesian regulation and so forth, in back-propagation can eliminate the risk of over-fitting. Other than Levenberg-Marquardt, the utilization of Bayesian regulation optimization in backpropagation algorithm does not require additional data set used for network validation (Demirezen & Fung, 2019) and there is no danger of over-fitting. For this reason, the Bayesian regulation backpropagation algorithm was used in network training. The learning rate and maximum number of epochs were set to 0.3 and 1000, respectively, as the network model has shown very successful prediction results using these hyperparameter values. All input and output neurons were denormalized after network training to measure the success of the developed neural network models. For interpretation of network performance, several different performance assessment parameters including mean squared error (MSE), mean absolute error (MAE), root-mean-squared error (RMSE), coefficient of determination (R^2), the variance accounted for (VAF) and regression analysis were employed. Equations of MSE, MAE, RMSE and R^2 were listed as below (Erzin & Turkoz, 2016; Khalaj, 2013; Nazari, Hajiallahyari, Rahimi, Khanmohammadi, & Amini, 2019):

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{2}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{3}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{4}$$

$$R^2 = 1 - \frac{\sum_i^n (y_i - \hat{y}_i)^2}{\sum_i^n \hat{y}_i{}^2} \tag{5}$$

where y_i is measured output data, $\hat{y}_i$ is ANN predicted output data and n is the total number of data.

4. Results and Discussion

An ANN model was developed to forecast the dynamic viscosity of gelatin solutions. Network architecture has 5 neurons in the input layer, 1 neuron in the output layer together with 1 hidden layer with different number of neurons, as mentioned in previous section. As a first step, the effect of hidden layer geometry, i.e. the number of neurons in the hidden layer as well as different activation functions used in the hidden and the output layer on training performance of developed network model was studied. MSE has been chosen as an error function in training phase since Bayesian regulation backpropagation algorithm can minimize sum of squared errors. MSE, R^2 and other details of all optimal network models were listed in table 2.

Table 2. Performance evaluation of the best network models having different transfer functions and hidden neurons

Number of hidden neurons	Transfer functions used in		μ	Epoch	Training		Testing	
	Hidden layer	Output layer			R^2	MSE	R^2	MSE
5	Tan-sigmoid	Tan-sigmoid	0.005	195	0.99991	5.28×10^{-6}	0.99869	1.78×10^{-4}
6	Log-sigmoid	Tan-sigmoid	0.5	509	0.99998	1.34×10^{-6}	0.99983	3.71×10^{-4}
6	Tan-sigmoid	Tan-sigmoid	0.005	422	0.99904	4.07×10^{-6}	0.99894	0.0106
7	Tan-sigmoid	Linear	0.5	314	0.99999	7.21×10^{-7}	0.99912	1.32×10^{-4}
7	Tan-sigmoid	Tan-sigmoid	0.005	549	0.99998	1.37×10^{-6}	0.99077	0.0075
8*	Tan-sigmoid	Tan-sigmoid	0.05	546	0.99999	6.27×10^{-7}	0.99918	1.26×10^{-4}
8	Tan-sigmoid	Linear	0.05	352	0.99998	8.45×10^{-7}	0.99957	3.91×10^{-4}
9	Tan-sigmoid	Tan-sigmoid	0.05	574	0.99999	6.52×10^{-7}	0.99752	1.65×10^{-4}
9	Log-sigmoid	Tan-sigmoid	0.05	474	0.99999	7.16×10^{-7}	0.99672	0.0011
10	Tan-sigmoid	Tan-sigmoid	0.05	685	0.99999	4.91×10^{-7}	0.99183	0.0018
11	Log-sigmoid	Tan -sigmoid	0.5	441	0.99999	3.62×10^{-7}	0.99172	7.32×10^{-4}

*Represents a suggested ANN model yielding the best performance results

As clearly shown in Table 2, when ANN model having 8 neurons in the hidden layer is trained by using a hyperbolic-tangent activation function in both hidden and output layer, it yields the highest R^2 and the least MSE values for both training and testing data sets, compared to other network models having different numbers of neurons in the hidden layer and different activation functions. For the best ANN model, MSE and R^2 values were found to be 6.27×10^{-7} and 0.9999, respectively.

Apart from MSE values, as shown in Figure 3(a), regression between measured viscosity and ANN predicted viscosity values, which yielded excellent R^2(=0.9999) value, quite close to unity, is a great indication of equally updated weights between interconnected nodes. It is clearly shown in Figure 3(a) that a uniform distribution of training data points around the linear least-square fit line in the regression plot is another reason why ANN reflects an excellent R^2 value.

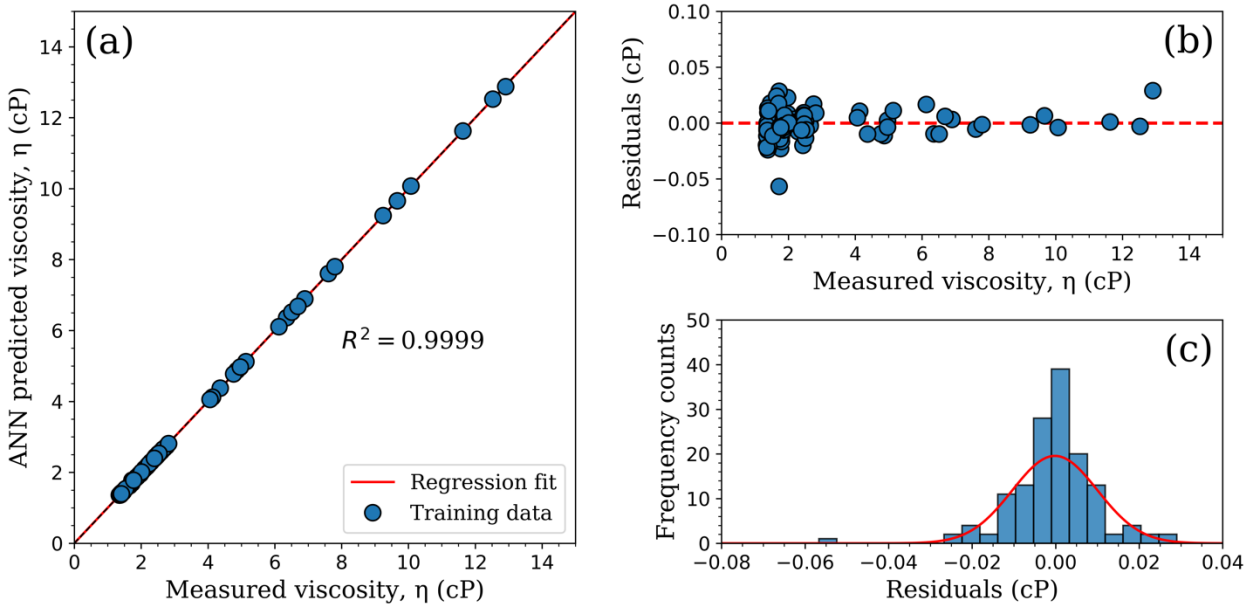

Figure 3. Performance assessment of ANN training: (a) regression plot, (b) residuals, the difference between ANN predicted viscosity and measured viscosity values, (c) residual distribution histogram

To further investigate how the developed network model is well-trained, residuals and their distribution were additionally analyzed. For each training data point, residuals (the difference between measured viscosity and ANN predicted viscosity values) were calculated. Residuals were found to be mainly centered around zero with very small homogeneous deviations (see Figure 3(b)), meaning that there is no systematic error stemming from the ANN method. The randomized symmetrical distributions of residuals on both sides of zero indicate that the developed ANN model is unbiased with a good prediction capability. The main source of prediction error might be due to the nature of data sets which cannot be entirely attributed to prediction capability of neural network itself, however, partially the data set used in the network model. The residual distribution of the training data set was then studied with a histogram plot to find out the range of the error. Figure 3(c) reveals that residuals obtained by the ANN method show a very symmetrical Gaussian distribution and most of the training data points used in ANN training found very close to zero, which is in a good agreement with results shown in Figure 3(b). It was also observed that residuals obtained using the ANN model were very small, varying between -0.0567 and 0.0289. Residuals of training data were found to be 0.0856 in total.

Table 3. Weights and biases for the best ANN topology having 8 hidden neurons

Hidden neuron	Weights						Bias	
	Input neurons					Output neuron	Hidden layer	Output layer
	C	w	M	$\dot{\gamma}$	τ	η		
1	-0.17429	-2.83420	5.96910	-2.83420	5.96930	13.0190	7.83870	3.039
2	-0.19178	-1.76520	-0.01034	-1.76450	-0.00365	2.1228	-3.10890	
3	0.06877	0.07517	0.09080	0.08348	0.22441	4.2771	0.27651	
4	-0.07328	0.35505	-1.09480	0.35502	-1.09850	-0.9783	-1.33850	
5	-0.00030	2.02180	-0.25705	2.02330	-0.25461	4.297	3.18920	
6	-0.06137	-0.13843	-0.11151	-0.12491	-0.00559	11.312	-0.88134	
7	0.15642	-2.37500	0.53605	-2.37400	0.53530	2.7375	-3.54640	
8	0.80464	-20.46600	4.18640	-20.4660	4.18660	10.569	-35.991	

In order to determine which of the input neurons employed in ANN training has a domestic effect on prediction performance of dynamic viscosity, relative importance-based sensitivity analyses were performed. The relative importance of each input neuron on that of dynamic viscosity was studied using inter-connection weights and biases at each neuron presented in Table 3 as well as Garson's algorithm (Garson, 1991). Garson suggested a method of partitioning ANN weights in order to determine the sensitivity of each input neuron in network topology as given below:

$$I_j = \frac{\sum_{m=1}^{m=Nh}\left(\left(w_{jm}^{ih} \div \sum_{k=1}^{Ni}\left|w_{km}^{ih}\right|\right)\cdot\left|w_{mn}^{ho}\right|\right)}{\sum_{k=1}^{k=Ni}\left\{\sum_{m=1}^{m=Nh}\left(\left(\left|w_{jm}^{ih}\right| \div \sum_{k=1}^{Ni}\left|w_{km}^{ih}\right|\right)\cdot\left|w_{mn}^{ho}\right|\right)\right\}} \tag{6}$$

where I_j is the relative importance of the jth input parameter on the output parameter, N_i and N_h are the number of input and hidden neurons, respectively, w are connection weights, the subscripts i, h and o refer to input, hidden and output layers, respectively, and subscripts k, m and n refer to input, hidden and output neurons, respectively. Essentially, the sensitivity analyses are applied to the developed network models to be trained with a new set of random weights to check the robustness of the developed ANN model. In our analysis, the best ANN model has been trained five times with various random starting weights to elucidate how the relative importance of each input neuron responds after each training step.

Table 4. Sensitivity analysis using relative importance of each input neuron employed in the network topology

Trial Number	Relative importance (%)				
	C	w	M	$\dot{\gamma}$	τ
1	4.9212	30.6567	16.9452	30.4652	17.0117
2	9.8811	31.4406	13.5993	31.4481	13.6310
3	7.2235	24.6285	23.2589	24.7425	20.1465
4	7.8629	25.4671	20.5122	25.5287	20.6291
5	4.6052	29.5418	18.0748	29.6451	18.1331
Average	6.8977	28.3469	18.4781	28.3659	17.9103
Ranking	**5**	**2**	**3**	**1**	**4**

The analysis results were listed in Table 4 and the calculated average relative importance (%) was additionally provided with a histogram plot including error bars that obtained from standard deviation (see Figure 4(a)). As shown in Table 4, it is clear that the order

of relative importance of each input neuron does not change after each training step with new random weights, demonstrating the robustness and reliability of the developed ANN model. The contribution of input neurons on network prediction varied from 4.6052% to 31.4481%. Among all input neurons, it is shown that $\dot{\gamma}$ has the most significant effect on the predicted dynamic viscosity values, followed by w, M, τ and C.

The reason why $\dot{\gamma}$ has a domestic contribution, as opposed to C, on the prediction of η by suppressing all other input neurons is a large distribution of $\dot{\gamma}$ data set used in ANN training. Distribution of all normalized input neurons obtained from the best ANN model is present in histogram plots shown in Figures 4(b)- 4(f). As clearly seen in Figure 4, most of the data points of $\dot{\gamma}$ distributed in a very broad range and distribution of data set of C (shown in Figure 4(b)) is very narrow due to repetitive nature of C. Since only 6 different aqueous gelatin concentrations were used as input neurons in the development of the ANN model, the relative importance of that of C was found to be weaker on prediction performance of developed ANN model, compared to other input neurons.

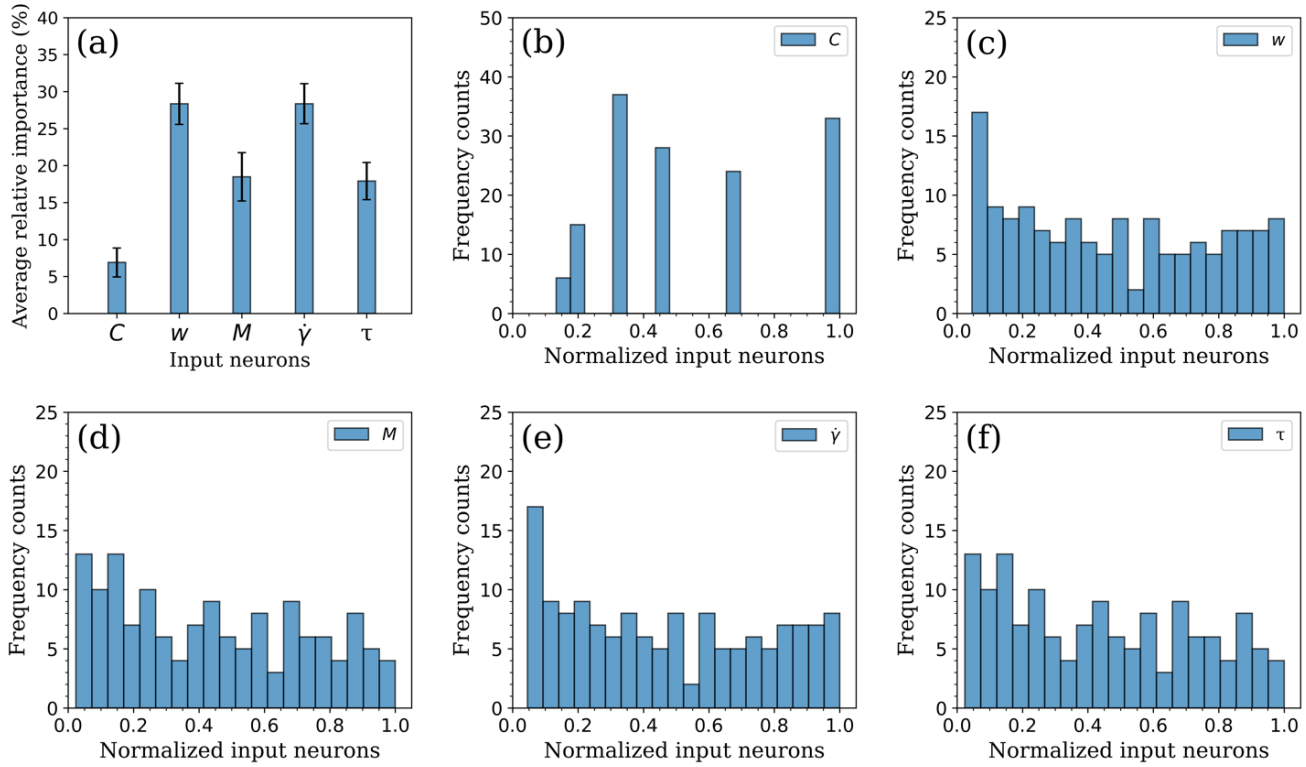

Figure 4. (a) Sensitivity analysis results and distributions of normalized input variables: (b) mass concentration of aqueous gelatin solutions, (c) the angular velocity of spindle, (d) torque of spindle, (e) shear rate, and (f) shear stress

After ANN training stage, a second data set, which has not introduced to neural network during training, have been used to test the performance of the developed ANN model with additional performance evaluation parameters to demonstrate the versatility of network architecture. In ANN testing, as an alternative to MSE, MAE and RMSE were considered as additional error functions to measure the accuracy of the ANN model. Because MSE itself may not be universal in testing the ANN model.

For the testing data set, MSE, MAE, RMSE and R^2 values were found as 1.26 x 10^{-4}, 0.0898, 0.1496 and 0.99918, respectively. Regression between measured viscosity and ANN predicted viscosity values for testing data set was also studied and a very fine R^2(=0.99918) value, which is almost unity, was obtained, as found for training data set. As shown in Figure 5(a) that testing data points dispersed very uniformly around the linear least-square fit line in the regression plot, indicating equally updated weights between interconnected nodes, as observed for ANN training.

Table 5. Additional performance indices of the best ANN model

Data set	Number of data	RMSE (cP)	MAE (cP)	VAF (%)
Training	143	0.0102	0.0068	99.9977
Testing	61	0.1496	0.0898	99.7738

In order to interpret how errors are greatly minimized in ANN testing, the residuals and their distribution were further investigated with a histogram plot. For each testing data point, the residuals were calculated and they randomly dispersed around zero with very

small homogeneous deviations on both sides of zero (see Figure 5(b)), showing that developed ANN model is unbiased and it is able to predict dynamic viscosity with high performance. The main reason why slightly higher errors arouse in ANN testing is possibly due to the number of data set used in testing (61 data points) is less than the amount of training data set (163 data points). Small prediction errors possibly as a result of a nature of data sets used in ANN testing, as mentioned and also found for ANN training. As shown in Figure 5(c), the majority of data points used in ANN testing accumulated around zero with a very small divergence, which is in a good agreement with the results shown in Figure 5(b). It was also observed that residuals found by using the ANN method varied from -0.756 to 0.125. Residuals of testing data were found to be 0.881 in total.

As distinct from MSE, MAE, RMSE and R^2 values employed in performance evaluation of ANN testing phase, Variance Accounted For (VAF) between predicted and measured data points, represented by Equation 7, was additionally computed to interpret the performance of the developed network model (Erzin & Cetin, 2013).

$$VAF = \left[1 - \frac{var(y - \hat{y})}{var(y)}\right] \times 100 \tag{7}$$

where *var* represents variance, y is measured value and $\hat{y}$ is the predicted value. If VAF and RMSE values are close enough to 100% and 0, respectively, the performance of the developed ANN model can be treated as excellent (Erzin & Turkoz, 2016). Computed VAF values together with other performance indices including RMSE and MAE are present in Table 5.

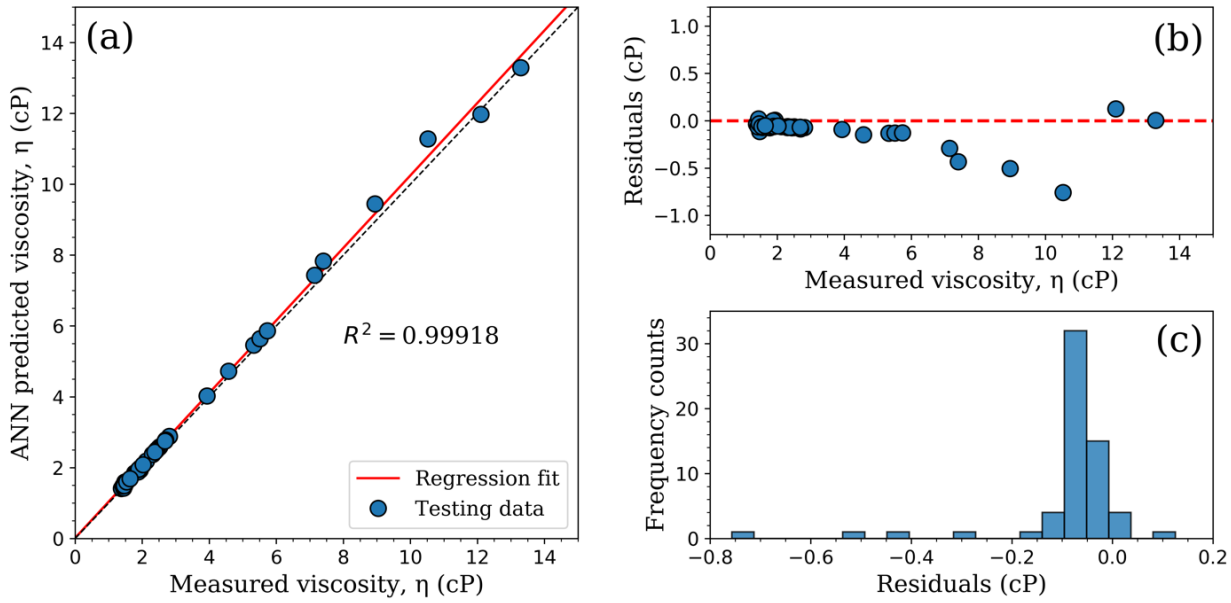

Figure 5. Performance assessment of ANN testing: (a) regression plot, (b) residuals, the difference between ANN predicted viscosity and measured viscosity values, (c) residual distribution histogram

Table 6. Details of the best ANN topology

Network type	Feed-forward neural network
Number of neurons in input layer	5
Number of hidden layers	1
Number of neurons in hidden layer	8
Number of neurons in output layer	1
Transfer function in hidden layer	Hyperbolic-tangent (tan-sig)
Transfer function in output layer	Hyperbolic-tangent (tan-sig)
Training rule	Bayesian regularization back-propagation algorithm
Training termination	Minimum mean squared error (MSE)
Momentum factor (μ)	0.05
Learning rate	0.3
Maximum number of epochs used in training	1000
Epoch number achieved minimum MSE	546
Performance goal	0.00000001

It can be seen from Table 5 that VAF and RMSE values for both training and testing were found to be quite close to 100% and 0, meaning that developed ANN model exhibits a very high prediction performance and these results can be easily verified with other performance indices obtained. All error parameters calculated in both training and testing phases (see Tables 2 and 5) were found very close to each other. It is known that in the case of over-fitting, testing errors and inter-connection weights would be much higher. In our analysis, as shown in Table 3, inter-connection weights and biases that we found in ANN training are very small, demonstrating that the data is not over-fitted. All details of the best ANN topology are additionally given in Table 6.

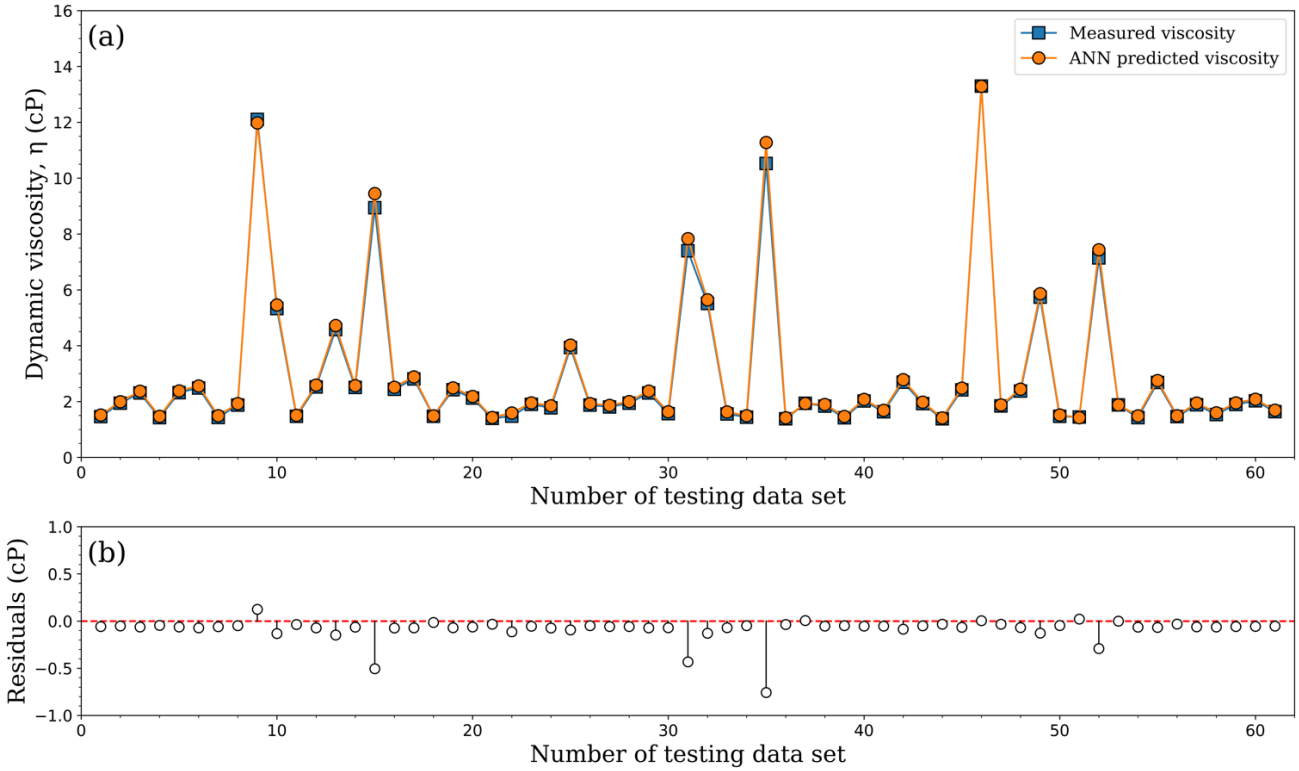

Figure 6. Prediction performance of developed ANN model using testing data set: (a) a comparison of all measured viscosity values with ANN predicted viscosity values, (b) residual analysis

As a final step of ANN performance analysis, ANN predicted viscosity values were compared with measured values. Figure 6(a) reveals that all ANN predicted η values were found quite close to measured η values, proving the usefulness and efficiency of the developed network model. The difference between these values was successfully minimized as most testing data points are approaching zero with random distributions around zero due to high performance and the robustness of the developed ANN model (see Figure 6(b)). To conclude, when all performance assessment parameters including MSE, MAE, RMSE, R^2 and VAF, found in both training and testing phases, are taken into account, it is very clear that η of aqueous gelatin solutions could be effectively predicted using ANN approach. One should also say that similar network models can be designed in order to forecast η value of other biopolymer solutions by using the same input and output parameters in the specific data range we reported in this work.

5. Conclusion

In this study, a multi-layered feed-forward neural network (FFNN) was used to predict the dynamic viscosity of aqueous gelatin solutions having different mass concentrations. For this purpose, the experimental results of viscosity measurements were used in the development of the FFNN model. In network topology, five different input neurons ($C, w, M, \dot{\gamma}$ and τ) and one output neuron, η, were used, respectively. Employing 8 neurons in the hidden layer has yielded the highest R^2 value, quite close to unity, and the least MSE value in both training and testing phases. Apart from MSE and R^2, other performance indices, such as MAE, RMSE and VAF were used to evaluate the prediction performance of developed FFNN models. MSE, R^2, MAE, RMSE, and VAF values were calculated as 6.24×10^{-7}, 0.99999, 0.0068, 0.0102 and 99.9977, respectively, for training data set, and calculated as 1.26×10^{-4}, 0.99938, 0.0898, 0.1496 and 99.7738, respectively, for testing data set. Using regression plots of training and testing data sets, the residuals between measured η values and ANN predicted η values were obtained and they randomly dispersed around zero with very small homogeneous deviations on both sides of zero since developed ANN model is unbiased and shows excellent performance. In addition, as a result of sensitivity analysis, $\dot{\gamma}$ was found as the most significant input neuron remarkably affecting the η value of aqueous gelatin solutions. All taken together, we can conclude that ANN can be employed as a very powerful soft computing method for the prediction of dynamic viscosity of aqueous gelatin solutions. Our findings have demonstrated the applicability of ANN in the prediction of η for a specific input-output data range.

6. Acknowledgment

Experimental data used in this work were provided from a part of undergraduate thesis written by Barış Demirbay.

7. References

Akkoyun, S., Yildiz, N., & Kaya, H. (2019). Neural Network Estimation for Attenuation Coefficients for Gamma-Ray Angular Distribution. Physics of Particles and Nuclei Letters, 16(4), 397-401. doi:10.1134/s1547477119040034

Aminian, A. (2017). Predicting the effective viscosity of nanofluids for the augmentation of heat transfer in the process industries. Journal of Molecular Liquids, 229, 300-308. doi: https://doi.org/10.1016/j.molliq.2016.12.071

Asteris, P. G., Roussis, P. C., & Douvika, M. G. (2017). Feed-Forward Neural Network Prediction of the Mechanical Properties of Sandcrete Materials. Sensors, 17(6), 1344.

Burden, F., & Winkler, D. (2009). Bayesian Regularization of Neural Networks. In D. J. Livingstone (Ed.), Artificial Neural Networks: Methods and Applications (pp. 23-42). Totowa, NJ: Humana Press.

Char, C., Padilla, C., Campos, V., Pepczynska, M., Díaz-Calderón, P., & Enrione, J. (2019). Characterization and Testing of a Novel Sprayable Crosslinked Edible Coating Based on Salmon Gelatin. Coatings, 9(10), 595.

Demirezen, G., & Fung, A. S. (2019). Application of artificial neural network in the prediction of ambient temperature for a cloud-based smart dual fuel switching system. Energy Procedia, 158, 3070-3075. doi: https://doi.org/10.1016/j.egypro.2019.01.992

Erdil, A., & Arcaklioglu, E. (2013). The prediction of meteorological variables using artificial neural network. Neural Computing and Applications, 22(7), 1677-1683. doi:10.1007/s00521-012-1210-0

Erzin, Y., & Cetin, T. (2013). The prediction of the critical factor of safety of homogeneous finite slopes using neural networks and multiple regressions. Computers & Geosciences, 51, 305-313. doi: https://doi.org/10.1016/j.cageo.2012.09.003

Erzin, Y., & Güneş, N. (2011). The prediction of swell percent and swell pressure by using neural networks. Mathematical and Computational Applications, 16(2), 425-436.

Erzin, Y., & Turkoz, D. (2016). Use of neural networks for the prediction of the CBR value of some Aegean sands. Neural Computing and Applications, 27(5), 1415-1426. doi:10.1007/s00521-015-1943-7

Fatehi, M.-R., Raeissi, S., & Mowla, D. (2017). Estimation of viscosities of pure ionic liquids using an artificial neural network based on only structural characteristics. Journal of Molecular Liquids, 227, 309-317. doi: https://doi.org/10.1016/j.molliq.2016.11.133

Foox, M., & Zilberman, M. (2015). Drug delivery from gelatin-based systems. Expert Opinion on Drug Delivery, 12(9), 1547-1563. doi:10.1517/17425247.2015.1037272

Garson, G. D. (1991). Interpreting neural-network connection weights. AI Expert, 6(4), 46-51.

Ghatak, A., & Robi, P. S. (2018). Prediction of creep curve of HP40Nb steel using artificial neural network. Neural Computing and Applications, 30(9), 2953-2964. doi:10.1007/s00521-017-2851-9

Goh, A. T. C. (1995). Back-propagation neural networks for modeling complex systems. Artificial Intelligence in Engineering, 9(3), 143-151. doi: https://doi.org/10.1016/0954-1810(94)00011-S

Gullapalli, R. P. (2010). Soft gelatin capsules (softgels). Journal of Pharmaceutical Sciences, 99(10), 4107-4148. doi:10.1002/jps.22151

Haykin, S. (1994). Neural Networks: A Comprehensive Foundation: Prentice Hall PTR.

Hemmat Esfe, M., & Abbasian Arani, A. A. (2018). An experimental determination and accurate prediction of dynamic viscosity of MWCNT(%40)-SiO2(%60)/5W50 nano-lubricant. Journal of Molecular Liquids, 259, 227-237.

Kayri, M. (2016). Predictive Abilities of Bayesian Regularization and Levenberg–Marquardt Algorithms in Artificial Neural Networks: A Comparative Empirical Study on Social Data. Mathematical and Computational Applications, 21(2), 20.

Khalaj, G. (2013). Artificial neural network to predict the effects of coating parameters on layer thickness of chromium carbonitride coating on pre-nitrided steels. Neural Computing and Applications, 23(3), 779-786. doi:10.1007/s00521-012-0994-2

Nazari, A., Hajiallahyari, H., Rahimi, A., Khanmohammadi, H., & Amini, M. (2019). Prediction compressive strength of Portland cement-based geopolymers by artificial neural networks. Neural Computing and Applications, 31(2), 733-741. doi:10.1007/s00521-012-1082-3

Osorio, F. A., Bilbao, E., Bustos, R., & Alvarez, F. (2007). Effects of Concentration, Bloom Degree, and pH on Gelatin Melting and Gelling Temperatures Using Small Amplitude Oscillatory Rheology. International Journal of Food Properties, 10(4), 841-851. doi:10.1080/10942910601128895

Pal, S. K., & Chakraborty, D. (2005). Surface roughness prediction in turning using artificial neural network. Neural Computing & Applications, 14(4), 319-324. doi:10.1007/s00521-005-0468-x

Wang, Y., Guo, Z., Qian, Y., Zhang, Z., Lyu, L., Wang, Y., & Ye, F. (2019). Study on the Electrospinning of Gelatin/Pullulan Composite Nanofibers. Polymers, 11(9), 1424.