TEST Engineering & Management

January - February 2020
ISSN: 0193 - 4120 Page No. 15505 - 15512

An Artificial Neural Networks Model for Determining Feasible Side-effects of Antibacterial Drugson Domestic Animals

***K. S. Senthilkumar[1], Kamashi Kumar[2]**

[1]School of Arts and Sciences, Department of Computers and Information Technology

[2]School of Veterinary Medicine, Department of Anatomy, Physiology & Pharmacology

St. George's University, Grenada, W.I.

Article Info
Volume 82
Page Number: 15505 - 15512
Publication Issue:
January-February 2020

Article History
Article Received*: 18 May 2019*
Revised*: 14 July 2019*
Accepted*: 22 December 2019*
Publication*: 28 February 2020*

Abstract
The potential side-effects of Antibacterial drugs are a major concern during the human and veterinary therapy. It became essential for the clinicians to be aware of the side-effects of these drugs for ensuring the safety of the treating animal. Adverse reactions to antimicrobial agents consist of toxic effects mediated through chemical changes in tissue cells, sensitivity reactions produced by antigen-antibody reactions, and biological effects caused by alteration of the bacterial flora in the body. Artificial Neural network (ANN) has been successfully applied to find solutions for various real world problems in many fields including healthcare. Medical diagnostic systems, biochemical analysis, and scanned image analysis are some of them where the medical practitioners couldn't address alone. Neural networks are good at discovering and learning patterns, especially in large volumes of quantified data. The main aim of this research is to develop an accurate, cost-effective and easy-to-use decision support system to support the clinicians. Considering these facts, this research was developed to associate the antimicrobial use for treating the respective bacterial infections to their probable side-effects by using supervised neural network. The experimental results provide substantial evidence to identify the side effects of a specific drug.

Keywords: *Supervised Neural Network, Antibacterial drug, soft computing.*

1. Introduction

Artificial Neural network (ANN) is a soft computing model, inspired by biological system in human brain. Training in biological systems involves making adjustments to the synaptic connections between the neurons and in ANN training is possible by adjusting the weights between the neurons. Both artificial and natural paradigms learn/train by examples and learning completes with the all training instances are run. Training the neural network is an critical research question in the field of ANN. The ANN's most common objective is reaching the global optima and the important shortcoming with ANN is the local minima. One of the popular supervised learning methods in ANN is called Backpropagation algorithm (BP). In this method, inputs and outputs are provided and the neural network is being trained based on them. Once the neural network is trained, it is expected to be able to classify or forecast unknown cases as well. We ran the training data multiple times using the learning algorithm without over fitting to get a better accuracy.

Antibacterial drugs are the drugs used for treating bacterial infections in human and veterinary medicine. The potential side-effects of these drugs are a major concern during the therapy. Hence, it is more essential for the clinicians to be aware of the side-effects of the antibacterial drugs for

January - February 2020
ISSN: 0193 - 4120 Page No. 15505 - 15512

ensuring the safety of the treating animal (Alison *et al*., 2012). Adverse reactions to antimicrobial agents consist of toxic effects mediated through chemical changes in tissue cells, sensitivity reactions produced by antigen-antibody reactions, and biological effects caused by alteration of the normal microbial flora in the gastrointestinal tract. Patients vary in their susceptibility to these side-effects because of genetically determined differences in enzyme and protein make-up, differences in immunologic reactivity, differences in environment, or because of the effects of disease. Almost any organ in the body may be involved, indicating that the metabolic processes affected by these agents are represented to a varying degree in the cells of the host (Wightman, 1965). A research study about the adverse drug effects estimated the prevalence of drug related adverse effects as 19% and to a maximum extent, allergic reactions constitute 79% of those adverse effects (Shehab et al., 2008).

Considering these facts, the present study was developed to associate the antimicrobial use for treating the respective bacterial infections to their probable side-effects. In this ANN pattern, the inputs used were the drug, route of administrationand clinical indication to derive the probable side-effect of the respective drug. The results were more satisfactory to identify the side effects of a specific drug.The remainder of the paper is structured as follows: Section 2 discusses somerelated work carried out using ANNinside-effects of antibacterial drugs problem, the proposed decision support scheme using ANN isdiscussed in Section 3; Section 4 contains the experimental results and analysis of the proposed model and finally, Section 5 presents conclusions and some directions for future work.

2. Related work

The ANN based topological method was developed to predict the properties of quinolone antimicrobial agents based on their chemical structures. This method uses neural networks in which training algorithms are used as well as artificial intelligence with a suitable set of topological descriptors are employed to determine the minimal inhibitory concentration (MIC) of quinolones. Analysis of the results showed that the experimental and calculated values are highly similar (José Jaén-Oltra et al., 2000). ANN technique was used on antimicrobial resistance pattern of avian pathogenic *Escherichia coli* samples to predict 38 different genes responsible for virulence factors(Daniela T. Rocha, 2015). A learning strategy using ANN approach was proposed to predict drug-target interactions on the basis of drug side effects and traits from genome-wide association studies and generate a novel and clinically applicable insight into the molecular determinants of drug action (Prinz et al., 2018).

A deep neural network model was developed for identifying the potential side-effects of drugs and for predicting the possible adverse drug reactions of a new drug. In this study, the detection performance was enhanced by distributing the representations of the target drugs in a vector space to capture the drug relationships using the word-embedding approach to process substantial biomedical literature (Chi-Shiang Wang, 2019).Rehman et al. proposed an ANN algorithm for improving the current working performance of Back-propagation algorithm by adaptively changing the momentum value but kept the gain parameter fixed for all nodes in the neural network. The performance of the proposed method is compared with the performances of the Gradient Descent Method with Adaptive Gain and Gradient Descent with Simple Momentum by using five classification problems. They have shown that the proposed algorithm can be used as an alternative approach for BPNN because it demonstrate better accuracy ratio on the chosen classification problems.Jiang and et al. reported the current status of Artificial Intelligence applications using ANN in healthcare such as cancer, neurology and cardiology. Fu et al. have proposed to demonstrate an ANN techniques to classify and predict positive wheelchair angle and recline settings for wheelchair users. A hybrid GA-based ANN technique is used to train the network to reduce the chance of converging to local optima and to improve the capability of classifying unseen data.

The ANN is used to find solutions for many real world problems. A single perceptron neural network model was proposed to represent the non-linear relationships among variables in Smart

January - February 2020
ISSN: 0193 - 4120 Page No. 15505 - 15512

Beam forming in array antennas problem by (Senthilkumar et al.) This Neural network model was programmed with the MATLAB software and designed to optimize the complex weights of a dipole array antenna to steer the beam to desired directions. A Perceptron network based successful implementation with a high accuracy in the maximum radiation direction of the desired beam, as well as optimization in the direction of null points is discussed by (Senthilkumar et al.) They compared the results of the array antenna beams obtained using the Perceptron Single Neuron Weight Optimization method (SNWOM) and the optimized beams obtained using the Least Mean Square (LMS) method. Two different types of desired antenna beams: one broad side beam and the other with the antenna radiating in four different desired directions were verified for the reliability and accuracy of the Perceptron based beamformer.

Haykin (1999) has published a detailed review of gradient descentbased different training algorithms. The Multi-Layer Perceptron (MLP), consists input neurons, hidden neurons, and output neurons. Each neuron in one layer is connected with all neurons in the next layer. Researchers in the area of ANN have used weight initialization techniquesto overcome local minima problem. Moreover, a number of search strategies have been proposed in themost recent research work by focusing on adaptation of relevant parameters and the numerical accuracy required for ANN, and their hardware implementation (Adam et al., 2014, Judd et al., 2016).

For this research we are concerned with supervised neural networks.

3. Design and implementation

A multilayer feed-forward neural networks based on the gradient descent method is proposed in this work. We tried to use gradient descent technique to minimize the output error function. The most of the ANNalgorithms are a combination of backpropagation learning algorithm and the feed-forward, with layered networks.To reach a great and accurate output always the weights of the ANN have to be optimized.

We consider the neural networks having four layers: the input, two hidden and the output layers with random initial weights. The data set contains 529 different records. Each example consists of three inputs and one output; each input corresponds to an individual drug, root and indication and the output is classifier that determines the side-effects. The mean squared error (MSE)metric is used to determine the network performance.MSE is calculated bysquare the average differences between each actual and the desired output.

Antimicrobial drug's data such as dose, clinical indications and side-effects are collected from the Small Animal Clinical Pharmacology book specific for dogs (Maddison et al., 2008).One of the most important procedures when forming a neural network is data normalization.We normalized the data using a *max-min normalization* technique, which changes the values of all input data from Rto [0, 1].

3.1 Hidden layer

The fixing of the number of hidden layers and number of neurons is an important for designing of Neural Network model. The random selection of hidden neuron may cause over fitting and under fitting problem in the network. Several researchers tried and proposed many approach for fixing this problem. The first parameter to be determined was the number of hidden layer and the number of units in hidden layer, which is basically a model selection problem. We did some experiments to determine these parameters and which show that changing the learning rate does not have a significant effect on the network performance. This experiment uses two fully-connected hidden layers, with nine nodes in each layer. Input layer contains 3 nodes and one node in the output layer.

3.2 learning rate

The next very important parameter is learning rate (η) and it is highly dependent on the training data. Choosing a proper learning rate could be a difficult part. The network will converge to the optimum quickly when using a higher learning rate than lower learning rate. We have chosen 0.45 as initial learning rate and 0.05 as minimum learning rate. When training the neural networks,

January - February 2020
ISSN: 0193 - 4120 Page No. 15505 - 15512

it is often useful to reduce learning rate as the training progresses. This activity is called adaptation and can make a huge difference in optimization speed and accuracy. We also tested with a modification to the learning rate, every 50 training epochs, we subtract a small amount (step size is 0.001) from the current learning rate to make sure to slow down as the MSE differential decreases. Deciding a good step size for modification is also a challenging part in ANN. For each training example, the weight deltas are immediately added to the weights after the error is calculated.

3.3 Activation function

The forward phase of the backpropagation algorithm can be described as follows: Assuming the input data $x=(x_1, x_2,\ldots,x_n)$, and the connecting weight between the neurons is $w=(w_1,w_2,\ldots,w_n)$. A typical neuron is depicted in figure 1.

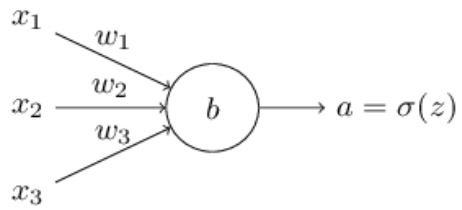

Figure 1. A Neuron

Each neuron processes the input data as describe in the equation (1):

$$z=\Sigma(w_i x_i)+ b \quad \text{----------} \quad (1)$$

Where b is the bias for neuron. The output of the equation (1) will be input to the function in equation (2).

$$a= \sigma (z) \quad \text{----------} \quad (2)$$

Where σ *(z)* is called activating function or transform function, which often takes the form of step function, and we used the sigmoid function displayed in equation (3). The sigmoid function returns a real-valued output.

$$\sigma(z)=1/(1+e^{-z}) \quad \text{----------} \quad (3)$$

Function (4) is used to calculate the Mean Square Error (MSE) by using the actual output o_i and the target output t_i and sum them to get the total error for each epoch.

$$MSE=\Sigma (o_i-t_i)^2 /n \quad \text{----------} \quad (4)$$

Where n is the number of records in the data set for training. In the backward phase of the backpropagation algorithm, weights in the network are updated, so that they cause the actual output to be closer to the target output. The training process was done for 10,000 - 160,000 epochs to train the network.

4. Results and discussion

The experiment result shows that, when we use two hidden layers with nine neurons, gave mostly good results in MSE and classification accuracy of output as shown in the Table 1 (n= Number of Neurons in hidden layer/s). Since the data in the dataset are all positive (labels), we normalize them between [0, 1] by using the formula A(i)=(A(i)-min(A))/(max(A)-min(A)), where A is the data set.The training data set was used to optimize the weights, and the test data was used to measure the performance of the proposed model. The big network will both be able to express more possible solutions and is less likely to get stuck in local minima. Decreasing the learning rate helped us to find a path out of what might seem like local minima. Training was carried out for 160000 epochs and the MSE wasbrought down to0.00013626. The resulting MSE in different epochs is depicted in figure 2. For epochs from 10,000 to 160,000 the Mean square error value during training and the accuracy of the model during testing is displayed in Table 1.

Table 1. Mean square error and Accuracy

Epochs	MSE	Accuracy
10000	0.002553	0.63681
20000	0.0013265	0.60731
30000	0.00093665	0.96573
40000	0.00076308	0.96997

January - February 2020
ISSN: 0193 - 4120 Page No. 15505 - 15512

50000	0.00059069	0.97215
60000	0.00031674	0.97733
70000	0.00019542	0.98135
80000	0.00017221	0.98282
90000	0.0001625	0.98345
100000	0.00015857	0.98369
110000	0.00015544	0.98394
120000	0.00015219	0.98419
130000	0.00014856	0.98437
140000	0.00014438	0.98442
150000	0.00013974	0.98452
160000	0.00013626	0.98457

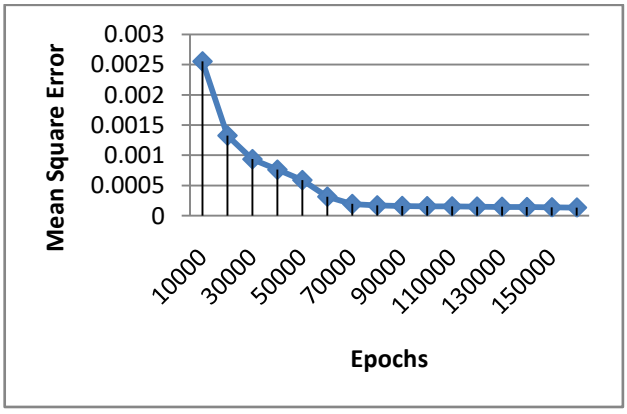

Figure 2. Mean square Error for each epochs during training

After training 160,000 epochs the accuracy was0.98457. The resulting Accuracy in different epochs is depicted in figure 3.For testing, 250 samples were used and the resultafter the 160,00th epochs is depicted in figure 4. The experiment result shows that although the actual value and target value are not identical for some samples, but the Mean square error is small enough, so the result is acceptable.

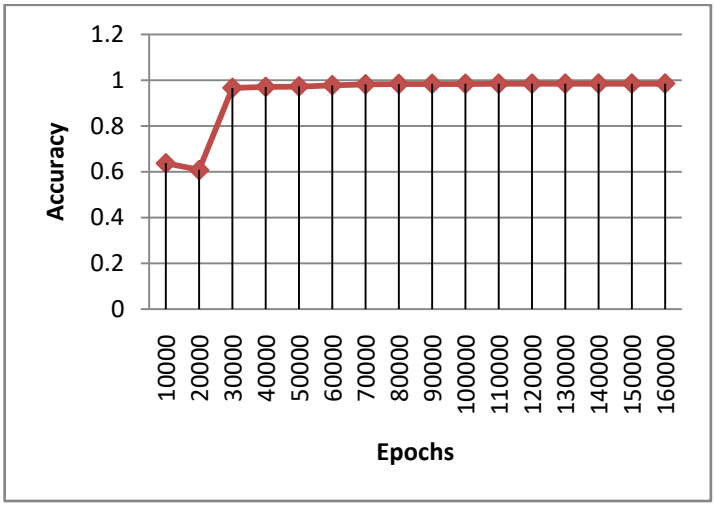

Figure 3. Accuracy for each epochs during testing

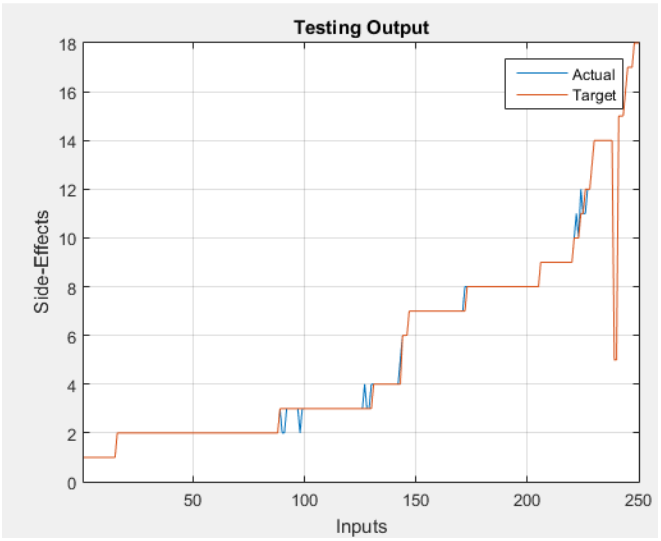

Figure 4. Actual output from the model against target results

5. Conclusion and future work

The inputs drug, species and clinical indication were normalized between the values 0 and 1 to make sure that the input and desired data are not in different ranges. Normalization of the data allowed us to use systematic weight initializations. The experimental results showed that two hidden layer with nine neurons in each layer is sufficient to solve the non-linearity among the data variables. The above experiments results provide substantial evidence to identify the side effects of a specific drug. Hence ANN study tool acts as a source for measuring the antibacterial drug's

January - February 2020
ISSN: 0193 - 4120 Page No. 15505 - 15512

adverse effects. However, factors like appropriate diagnosis of clinical disease, drug's formulation, dose administered, host health, duration of therapy could also influence the antibacterial drug's side effects.We are still training up this neural network model to get 100% accurate results on associating the antimicrobial drug to its related side-effects. Different ANN architectures with different activation functions will be considered to compare the result of this model.

Sample data (for dogs)

Drug	Route	Indications	Side-effects
Penicillin G sodium, potassium	IV, IM, SC	Gram Positive bacteria, anaerobes (abscess, wounds, pyothorax), periodontitis, gingivitis, ulcerative stomatitis, clostridial diseases, actinomycosis	GI disturbances, hypersensitivity reactions
Penicillin G procaine	IM, SC	Gram Positive bacteria, anaerobes (abscess, wounds, pyothorax), periodontitis, gingivitis, ulcerative stomatitis, clostridial diseases, actinomycosis	GI disturbances, hypersensitivity reactions
Penicillin V	PO	Gram Positive bacteria, anaerobes (abscess, wounds, pyothorax), periodontitis, gingivitis, ulcerative stomatitis, clostridial diseases, actinomycosis	GI disturbances, hypersensitivity reactions
Amoxicillin	PO, IV, IM, SC	Gram positive and negative bacteria, gingivitis, periodontitis, ulcerative stomatitis	GI disturbances, hypersensitivity reactions
Ampicillin	IV, IM, SC, PO	Gram positive and negative bacteria, gingivitis, periodontitis, ulcerative stomatitis	GI disturbances, hypersensitivity reactions
Ticarcillin-clavulanate	IV	Gram negative bacteria, Pseudomonas aeruginosa	GI disturbances, hypersensitivity reactions
Carbenicillin	IV, IM, SC	Gram negative bacteria, Pseudomonas aeruginosa	GI disturbances, hypersensitivity reactions
Ticarcillin	IV, IM, SC	Gram negative bacteria, Pseudomonas aeruginosa	GI disturbances, hypersensitivity reactions

January - February 2020
ISSN: 0193 - 4120 Page No. 15505 - 15512

Oxytetracycline, Tetracycline	IV, IM	Gram positive, Gram negative aerobic bacteria, *Rickettsia, Borrelia and Mycoplasma, conjunctivitis (eye ointment), antibiotic responsive diarrhea*	GI disturbances, hypersensitivity reactions, renal damage, liver damage, tooth mottling, growth retardation
Lincomycin	PO	Gram positive aerobes	Gastroenteritis
Lincomycin	IV, IM	Gram positive aerobes	Gastroenteritis
Erythromycin	PO	Gram positive, *Campylobacter jejuni, Mycoplasma*	Prokinetic, vomiting

Compliance with Ethical Standards:

Funding: This study was not funded by any funding organization. This research did not receive any specific grant from funding agencies in the public, commercial, ornot-for-profit sectors.

Conflict of Interest: we (authors) have no conflict of interest.

Ethical approval : All applicable international, national, and/or institutional guidelines for the care and use of animals were followed.

Ethical approval: This article does not contain any studies with human participants performed by any of the authors.

References

[1] Adam S. P., Karras D. A., Magoulas G. D., and Vrahatis M. N. (2014). Solving the linear interval tolerance problem for weight initialization of neural networks. Neural Networks, 54, 17-37.

[2] Alison E. Barnhill, Matt T. Brewer, Steve A. Carlson. (2012). Adverse Effects of Antimicrobials via Predictable or Idiosyncratic Inhibition of Host Mitochondrial Components, Antimicrobial Agents and Chemotherapy , 56(8), 4046–4051

[3] Chi-Shiang Wang,Pei-Ju Lin, Ching-Lan Cheng, Shu-Hua Tai, Yea-Huei Kao Yang, Jung-Hsien Chiang. (2019). Detecting Potential Adverse Drug Reactions Using a Deep Neural Network Model. Journal of Medical Internet Research, 21(2).

[4] Daniela T. Rocha, Felipe O. Salle, Gustavo Perdoncini, Silvio L.S. Rocha, Flávia B.B. Fortes, Hamilton L.S. Moraes, Vladimir P. Nascimento, Carlos T.P. Salle. (2015). Classification of antimicrobial resistance using artificial neural networks and the relationship of 38 genes associated with the virulence of Escherichia coli isolates from broilers. *Pesq. Vet. Bras*.,35(2).

[5] Fu, J., Genson, J., Jan, Y.-K., & Jones, M. (2011). Using Artificial Neural Network to Determine Favorable Wheelchair Tilt and Recline Usage in People with Spinal Cord Injury: Training ANN with Genetic Algorithm to Improve Generalization. *23rd IEEE International Conference on Tools with Artificial Intelligence*, 25-32.

[6] Haykin S., (1999) Neural Networks A Comprehensive Foundation. Prentice-Hall, Upper Saddle River, New Jersey.

[7] Jill E. Maddison, Stephen Page & David Church (Editors). (2017). Small Animal Clinical Pharmacology, (2nd edition, 2008). Saunders Ltd., ISBN: 978-0-7020-2858-8.

[8] Jiang F, Jiang Y, Zhi H, Dong Y, Li H, Ma S, Wang Y., Dong, Q., Shen H.,

January - February 2020
ISSN: 0193 - 4120 Page No. 15505 - 15512

Wang Y. (2017). Artificial intelligence in healthcare: past, present and future. Stroke Vasc Neurol. 2(4):230–43.

[9] José Jaén-Oltra, Ma Teresa Salabert-Salvador, Francisco J. García-March, Facundo Pérez-Giménez, and Francisco Tomás-Vert. (2000).Artificial Neural Network Applied to Prediction of Fluorquinolone Antibacterial Activity by Topological Methods.J. Med. Chem., 43(6), 1143-1148.

[10] Judd P., Albericio J., Hetherington T., Aamodt T. M., Jerger N. E., and Moshovos A., (2016,). Proteus: Exploiting numerical precision variability in deep neural networks, in Proceedings of the 2016 International Conference on Supercomputing, ICS '16, 23, 1-12, New York, ACM.

[11] Prinz J, Koohi-Moghadam M, Sun H, Kocher JA, Wang J. (2018). Novel Neural Network Approach to Predict Drug-Target Interactions Based on Drug Side Effects and Genome-Wide Association Studies.Hum Hered., 83(2), 79-91.

[12] Rehman, M. Z., & Nawi, N. M. (2011). Improving the Accuracy of Gradient Descent Back Propagation Algorithm (GDAM) on Classification Problems. International Journal onNew Computer Architectures and Their Applications, 838-847.

[13] Senthilkumar K. S., Pirapaharan K., Hoole P. R. P., Hoole H. R. H. (2016).Single Perceptron Model for Smart Beam Forming in Array Antennas, International Journal of Electrical and Computer Engineering (IJECE), 6(5), 2300 -2309.

[14] Senthilkumar K. S., Pirapaharan K., Lakshmanan G. A., Hoole P. R. P., and Hoole S. R. H, (2016 June). Accuracy of Perceptron Based Beam forming for Embedded Smart and MIMO Antennas, Proceedings of International Symposium on Fundamentals of Electrical Engineering (ISFEE2016), Bucharest, Romania, June 30-July 2, 1-6.

[15] Shehab N, Patel PR, Srinivasan A, Budnitz DS. (2008). Emergency department visits for antibiotic-associated adverse events. *Clin Infect Dis.* 47(6), 735-743.

[16] Wightman K. J. R. (1965). Adverse Reactions to Antimicrobial Therapy.Can Med Assoc Journal. 93(16), 870–873.