International Journal on
"Technical and Physical Problems of Engineering"
(IJTPE)
Published by International Organization of IOTPE

ISSN 2077-3528
IJTPE Journal
www.iotpe.com
ijtpe@iotpe.com

June 2020 | Issue 43 | Volume 12 | Number 2 | Pages 58-65

USING ARTIFICIAL INTELLIGENCE FOR DETECTION OF LYMPHATIC DISEASE AND INVESTIGATION ON VARIOUS METHODS OF ITS CLASSIFICATIONS

S. Sarabi [1] M. Asadnejad [1] S.A. Tabatabaei Hosseini [2] S. Rajebi [3]

1. Department of Biomedical Engineering, Tabriz Branch, Islamic Azad University, Tabriz, Iran
stu.m.asadnejad@iaut.ac.ir, stu.s.sarabi@iaut.ac.ir
2. Department of Advanced Medical Sciences, Tabriz University of Medical Science, Tabriz, Iran
tabatabaeihosseini@tbzmed.ac.ir
3. Department of Electrical Engineering, Seraj Higher Education Institute, Tabriz, Iran, s.rajebi@seraj.ac.ir

Abstract- Lymph metastases and malign lymph are those diseases which late diagnosis can cause leukemia, but timely diagnosis has a great effect on their treatment. In this paper, data collected from the lymphogram are used to classify them by using four different methods include the Bayesian method, *K*-Nearest Neighbour method (KNN), Perceptron Neural Network and RBF Neural Network, Where Bayesian and KNN are known as old methods and Neural Network is a new method. In the KNN method changing the value of *K* and in Perceptron Neural Network, changing the number of neurons and the type of transfer function and in RBF Neural Network, changing the number of Radial Basis functions changes the value of *CCR*. Given the available data and classifications performed by each method and the comparison of the times and the obtained *CCR*s, the RBF Neural Network method can perform the most desirable answer for the data classification and the KNN method performs classification as quickly as possible.

Keywords: Lymph, Lymphography, Classification, Bayesian, *K*-Nearest Neighbor, Artificial Neural Network.

1. INTRODUCTION

Lymph contains specific white blood cells called lymphocytes, which plays an essential role in immunizing the body. Lymph is a fluid that flows through lymphatic system. The lymph system is composed of lymphatic vessels, lymph nodes, and the like. The function of this system is like intravenous system, which causes fluid to return from the tissue to central circulation system [1, 2].

Lymphography is a medical imaging technique in which a radioactive substance is injected into the body to create contrast. Then, X-rays are used to illustrate the structures of the lymphatic system (lymph nodes, lymphatic channels, etc.). The lymphogram figures issued to locate the vessels and identify the obstruction [3, 4].

Metastases and malign lymph can be cited for problems that can occur for the lymphatic system. The features of lymphogram can be used to classify these problems. The 18 features have been investigated, some of these features are as follows:

- lymphatics: normal, arched, deformed, displaced
- block of affair: no, yes
- by pass: no, yes
- extravagates: no, yes
- early uptake in: no, yes
- changes in lymph: bean, oval, round
- special forms: no, chalices, vesicles
- no. of nodes in: 0-9, 10-19, 20-29, 30-39, 40-49, 50-59, 60-69, >=70

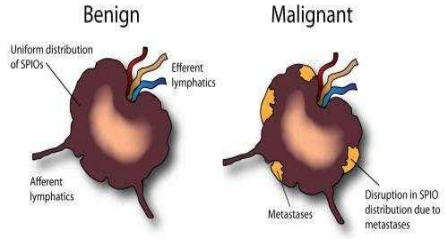


Figure 1. Lymph metastases

In this paper, the features from the lymphogram obtained by the University Medical Centre, Institute of Oncology, Ljubljana, Yugoslavia, have been used.

2. METHODOLOGY

To implement a classifier between benign and malignant lymph, by using the above mentioned collected data, various methods include: KNN, Bayesian and Neural Network (Perceptron and RBF) have been used. The data are categorized in two classes of metastases and malign lymph and the accuracy of each method is compared with each other.

2.1. Bayesian Method

Bayesian is a statistical model that by assuming the existence of a Gaussian distribution between the data, measures the probability of placing the test data in each class. Then, the class with the highest probability is considered as the test data class [5, 6].

The Gaussian function, along with its basic parameters, is depicted in Figure 2.

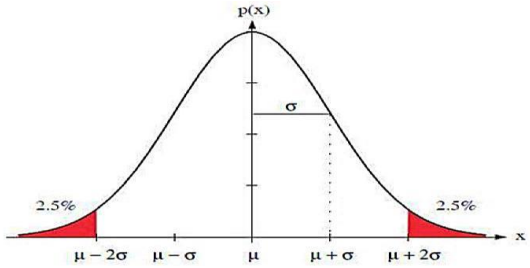


Figure 2. Gaussian function form along its basic parameters

Given the Gaussian Equation (1), the probability of the test data in each class is measured, and then the class with most P is considered as class on the test data.

$$P = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{(x-\mu)^2}{\sigma^2}} \quad (1)$$

where, P is probability of test data in class σ^2 is variance and μ is mean of the training data from that class.

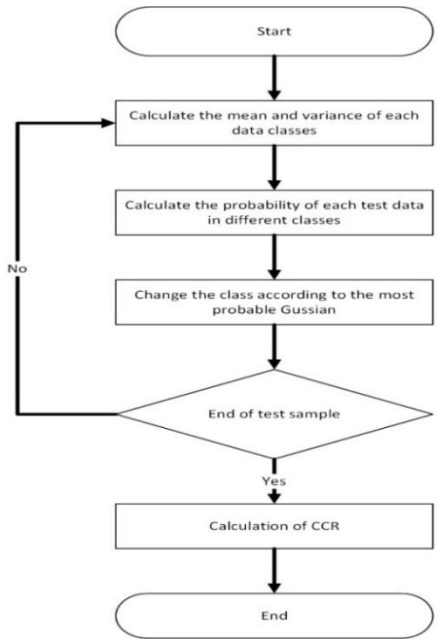


Figure 3. Flowchart of Bayesian classifier

In accordance with the flowchart shown in Figure 3, the Bayesian classifier was implemented and evaluated on lymph data. The obtained *CCR* has been reported as follows: *CCR*=9.09% and elapsed time was 0.48 seconds. (The computer used for this simulation has a 2.9 GHz Core i7 CPU and 8 GB of RAM). Also, diffusion matrix has been simulated as following. This matrix is used to indicate the number of correct answers in each class, in which the main diameter members indicate the correct answers and the other members indicate the wrong answers.

$$\begin{bmatrix} 1 & 10 \\ 10 & 1 \end{bmatrix}$$

The diffusion matrix of this algorithm shows that in each of the two classes, only one data is correctly classified and the other 10 are classified in wrong class.

The reason why the results are not desirable is that our data form is not fully Gaussian shaped. To improve the results, whitening of the data is suggested.

2.2. *K*-Nearest Neighbor Method

KNN is a way to find class of a test data by using the closest points in metric squares. In this method, the Euclidean distance of the test data is compared with the train data. Data with less distances are chosen as neighbours, then from class types of selected neighbours, class of the test data is reported [7, 8].

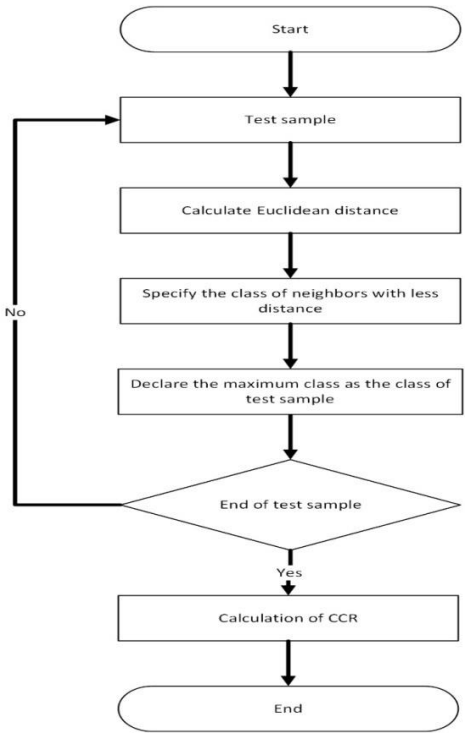


Figure 4. Flowchart of *K*-Nearest Neighbour classifier

In accordance with the flowchart shown in Figure 4, the KNN classifier was implemented and evaluated on lymph data. The obtained *CCR* has been reported as *CCR*=90.91% and Elapsed time for this classification is 0.03 seconds. Also, diffusion matrix has been simulated as following.

$$\begin{bmatrix} 9 & 2 \\ 0 & 11 \end{bmatrix}$$

Creating changes in K (the number of nearest neighbours) causes changes in the obtained *CCR*. Effect of changing in the number of K is summarized in Table 1.

Table 1. Effect of changing of K in performance of KNN classifier

K	*CCR* (%)	Time (s)	Diff
3	77.27	0.017	$\begin{bmatrix} 9 & 2 \\ 3 & 8 \end{bmatrix}$
5	90.90	0.039	$\begin{bmatrix} 9 & 2 \\ 0 & 11 \end{bmatrix}$
7	81.81	0.021	$\begin{bmatrix} 8 & 3 \\ 1 & 10 \end{bmatrix}$

By studying the *CCR*s in different values of *K*, it can be seen that *K* has a great influence on the value of the obtained results. Therefore, it is possible to achieve the desired result by changing the value of *K*, which in this paper the best result is for $K = 5$. In the KNN algorithm (with $K = 5$) in the first class, there are 9 correct answers and in the second class, all of answers are classified correctly.

2.3. Artificial Neural Network

The main philosophy of ANN is to model the processing of the human brain to approximate computational methods with the biological processing method [9, 10].

Several neural networks have been introduced for using in classification. In this paper Multi-Layer Perceptron and Radial Basis Function Network have been used to classify between lymphography dataset. These two methods have been studied and implemented in the following.

2.3.1. Perceptron Neural Network

Perceptron is a type of neural networks suggests an algorithm for learning under the supervision classifications. This classification can decide if an input represents a vector of numbers belonging to a particular class. In the Perceptron neural network, a vector of weights with real values is produced in the input vector, and the bias task is to move the decision boundary from the origin, and its value does not depend on the inputs, the relationship is as follows [11, 12].

$$w.x = \sum_{i=1}^{m} w_i . x_i + b \tag{2}$$

In which *b* is bias, *x* is the input vector and *w* is the vector of weight with real values.

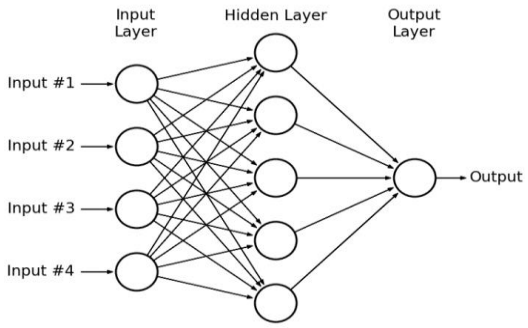


Figure 5. Perceptron neural network form

In the Perceptron neural network, different results can be obtained by changing the number of layers, the number of neurons, and the type of transfer function.

According to Figure 6, the simple single-layer perceptron neural network with different transfer functions has been simulated to classify the data.

The transfer function is one of the factors that can affect the *CCR*. Figure 7 shows four types of transfer functions.

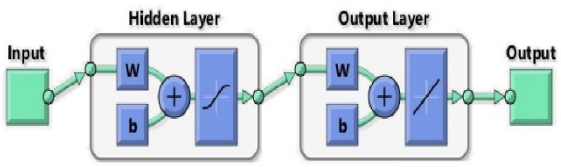


Figure 6. A double-layer perceptron neural network form

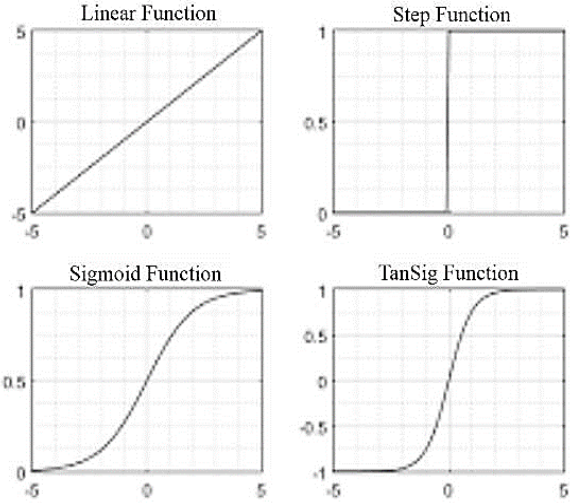


Figure 7. Form of some transfer functions

The (tan sig) and (log sig) are considered as two types of transfer functions, which their relationships are as follows.

$$\tan\mathrm{sig}(x) = 1/(1+e^{(-2x)}) - 1 \tag{3}$$

$$\log\mathrm{sig}(x) = 1/(1+e^{(-x)}) \tag{4}$$

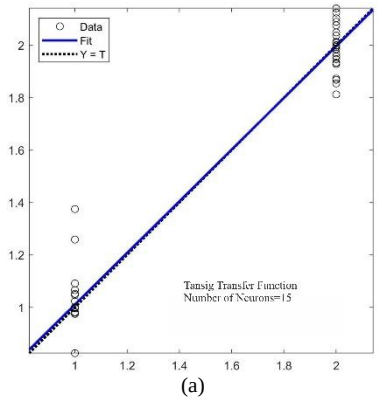
The regression line and the values of *CCR* of the multi-layer Perceptron neural network with the (tan sig) transfer function and different number of neurons in the hidden layer are shown in Figure 8.

According to the *CCR* values shown in Table 2, the best answer is in the range of $10<NN<20$. The values of *CCR* and the Regression line shown in Table 3 and Figure 9 are obtained by changing the transfer function from (tan sig) to (log sig).

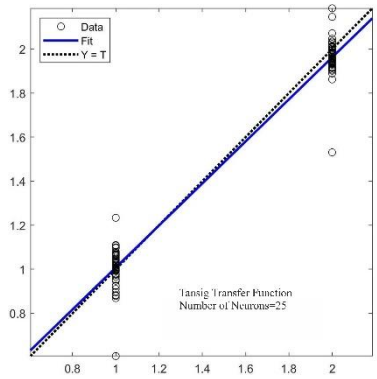
According to Tables 2 and 3, as well as Figure 9, it can be concluded that changing the number of neurons in the (log sig) transfer function does not have much effect on the *CCR*, but increasing the number of neurons in (tan sig) transfer function decreases the *CCR*. Changing the transfer function does not have much impact on the average elapsed time.

2.3.2 Radial Basis Function Neural Network

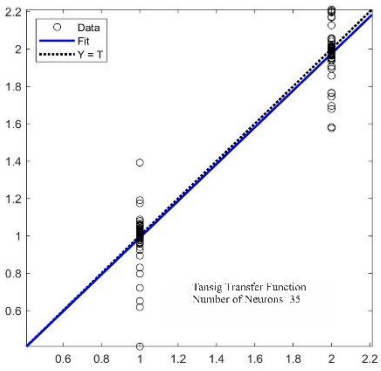
The RBF neural network is composed of neurons that use Radial Basis Functions. This network uses Gaussian functions to separate classes. In other words, the separator line is created by a linear combination of Gaussian functions. The variance and the mean of Gaussian functions are arranged to model the separator line [13, 14].



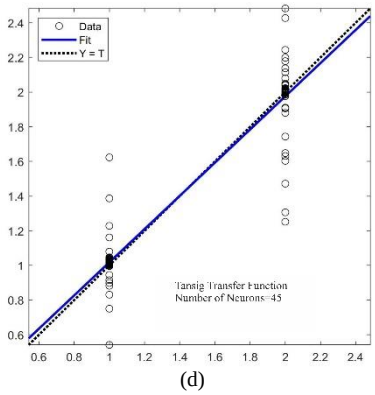
(a)



(b)



(c)



(d)

Figure 8. Regression line of different values of *CCR*: (a) *NN*=15; (b) *NN*=25; (c) *NN*=35; (d) *NN*=45

Table 2. Different results of neuron numbers of (tan sig) transfer function

Figure 8	a	b	c	d
NN	15	25	35	45
CCR (%)	99.59	98.82	96.34	95.53

the average elapsed time is 1.29 seconds

Table 3. Different results of neuron numbers of (log sig) transfer function

NN	15	25	35	45
CCR (%)	99.21	99.72	99.54	99.23

the average elapsed time is 1.28 seconds

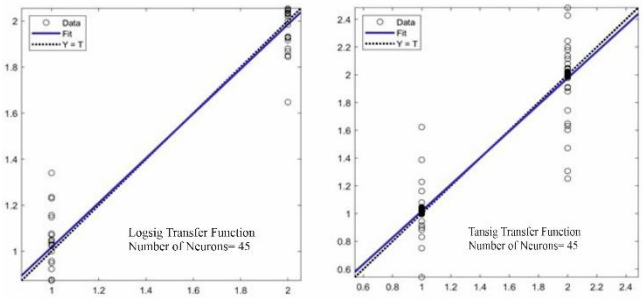


Figure 9. Regression line for (log sig) and (tan sig) transfer functions with *N*=45

The Regression line and the chart of different values of *CCR* with different number of Radial Basis Function are shown in Figures 12 and 13.

According to Figure 11 neurons with Radial Basis Functions in the hidden layer has been used for simulation.

According to Figure 13, with increasing the number of Radial Basis Function, the increase in *CCR* occurs, and it is seen that in the case the number of Radial Basis Function of 40, *CCR* has reached its highest value. The amount of *CCR* has remained constant since then.

According to Figure 14, it is seen that with the increase of the number of Radial Basis Function, the error rate has decreased and in the number of Radial Basis Function of 40 it has reached its lowest error.

According to Table 4, the results obtained are improved by increasing the number of Radial Basis Function. When reaches 40, the *CCR* reaches its maximum, and then its value stays constant. Also, with the changes in number of Radial Basis Function, the Elapsed time changes. The greater the amount of Radial Basis Function, the longer the Elapsed time will be.

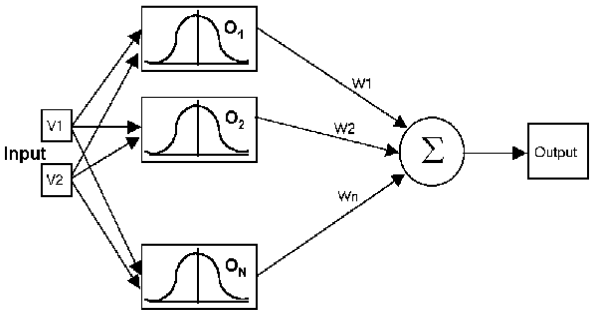


Figure 10. RBF Neural Network

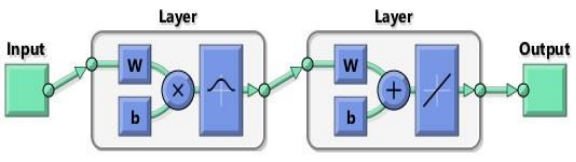


Figure 11. Radial Basis Function form

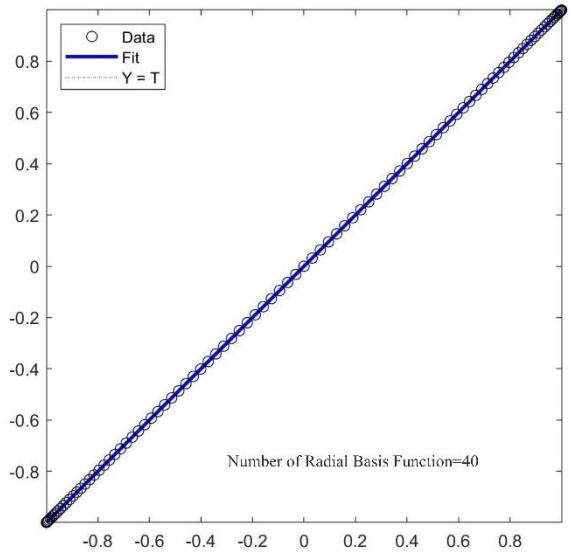


Figure 12. Regression line of Radial Basis Function of 40

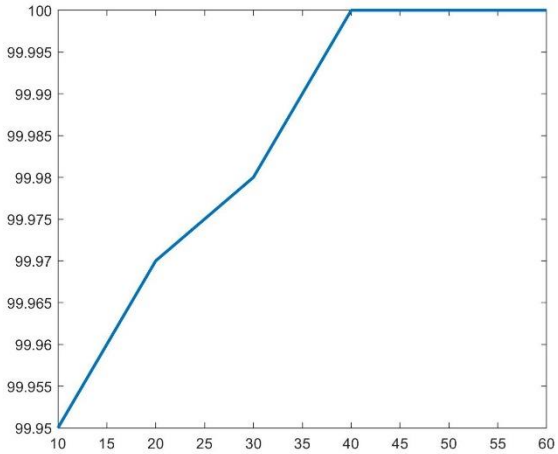


Figure 13. The chart of different values of *CCR*

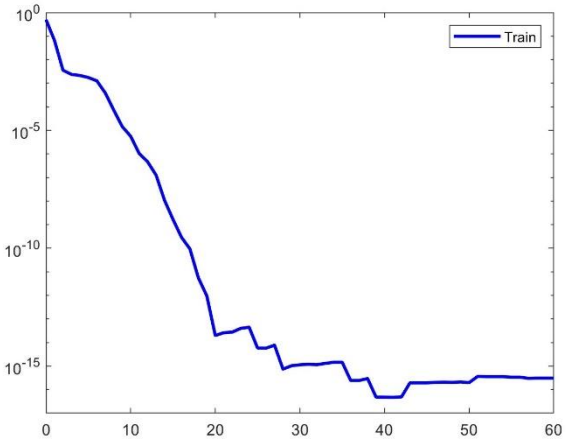


Figure 14. Performance

Table 4. Results from different numbers of Radial Basis Function

NRBF	10	20	30	40	50	60
Time (s)	0.98	1.14	1.35	1.88	2.08	2.17
CCR (%)	99.95	99.97	99.98	100	100	100

2.3.3. Support Vector Machine

Support vector machine is a supervised machine learning approach that can be used in regression and classification to classify data into two categories. SVMs enhance the accuracy of prediction while avoiding data overfitting using machine learning theories [17, 18].

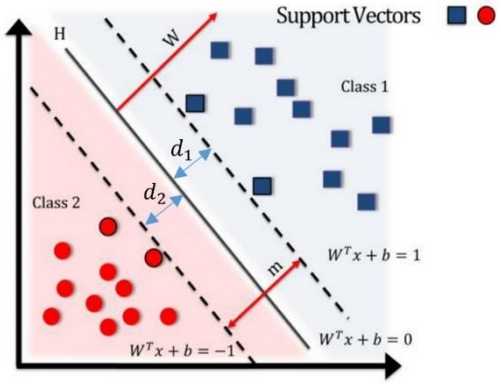


Figure 15. SVM classification plot [18]

A support-vector machine builds a hyperplane (*H*) to maximize the boundary between the two classes, which can be used for classification, regression, or other tasks such as detecting outlier samples.

By assuming that our dataset is linearly separable, the selection of two hyperplanes which separate the data, have to be done in a way that the margin or distance between hyperplanes (*m*) reach the maximum value.

The equation of hyperplane can be written:

$$w^T.x+b=0 \tag{5}$$

$$w^T.x+b\geq 1 \tag{6}$$

$$w^T.x+b\leq -1 \tag{7}$$

$$y(w^T.x+b)\geq 1 \tag{8}$$

where, *x* represents the Input vector of each class, *w* is the normal vector to the hyperplane and *b* depicts the bias vector.

According to Figure 15 and by considering *d* as the distance between each hyperplane we have:

$$d_1=d_2 \tag{9}$$

$$m=d_1+d_2 \tag{10}$$

$$m=2\left\|\frac{-(b-1)}{w}-\frac{-b}{w}\right\| \tag{11}$$

The goal is maximizing the distance between negative and positive hyperplanes *m*. In order to Maximizing the margin, the norm of *w* have to be minimized (12).

$$m=\frac{2}{\|w\|} \tag{12}$$

To be confident of classification accuracy, a relaxing factor $\xi_i\geq 0, i=1,2,...,n$ is introduced, so optimization problem can be written as follow:

$$\min\frac{1}{2}\|w\|^2+c\sum_{i=1}^{n}\xi_i,\ \xi_i\geq 0$$
$$y(w^T.x+b)\geq 1-\xi_i,\ i=1,2,...,n,\ c\geq 0 \tag{13}$$

In which c is the penalty factor. To achieving a low training and testing errors to generalize the most appropriate classifier to unseen data, the parameter c must be acted as a penalty factor which controls the trade-off between errors of the SVM on training data and margin maximization. Small c leads constraints to be easily ignored while large c makes them hard to ignore, setting c equal to infinity ($c=\infty$) causes to the hard margin SVM and enforces all constraints [18].

Finding the optimal hyperplane requires to choose the best value of w in Equation (14). To achieve this goal, Lagrange multipliers are used Equation (15).

$$w=\sum_{i=1}^{n}\alpha_i y_i x_i \tag{14}$$

By using terms α and w, Lagrange function can be written as below:

$$L(a)=\sum_{i=1}^{n}\alpha_i-\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_i\alpha_j y_i y_j K(x_i^T.x_j)$$
$$\sum_{i=1}^{n}\alpha_i y_i\ ,\ 0\le\alpha_i\le c\ ,\ i=1,2,\ldots,n \tag{15}$$

In which $K(x_i^T.x_j)$ represents a kernel function and equals to $K(x_i^T.x_j)=\varphi^T(x_i).\varphi(x_j)$, where φ is a nonlinear function and is usually unknown which the inner product operation can be replaced with so-called kernel function that satisfies Mercer's condition [17].

The final classification results of data points will be calculated by decision function of y which written as below and eventually the model have to build a hyperplane to separate two classes completely.

$$y=\text{sign}\left(\sum_{i=1}^{n}\alpha_i y_i K\left(x_i^T,x_j\right)+b\right) \tag{16}$$

The most popular kernel functions are: linear, polynomial, Gaussian radial basis function (RBF) [11].

$$K\left(x_i,x_j\right)=\left(x_i^T.x_j\right) \tag{17}$$

$$K\left(x_i,x_j\right)=\left(1+x_i^T.x_j\right)^p \tag{18}$$

$$K\left(x_i,x_j\right)=\exp\left(-\frac{1}{2\sigma^2}\|x_i-x_j\|^2\right) \tag{19}$$

In this paper, the classification accuracy of three above mentioned kernel functions with different penalty factors c have been measured.

According to the testing results in Figure 16 , it can be seen that the *CCR* value changes in response to the type of kernel function in a way that, in Linear kernel function, *CCR* value equals to 95%, in Polynominal kernel function it equels to 90% and it is 82% in RBF kernel function.It can indicated that by choosing different values for penalty factor *c*, there is no changes occur in *CCR* value but the elapsed time changes respectively.

The results and elapsed time for each function are plotted as follow.

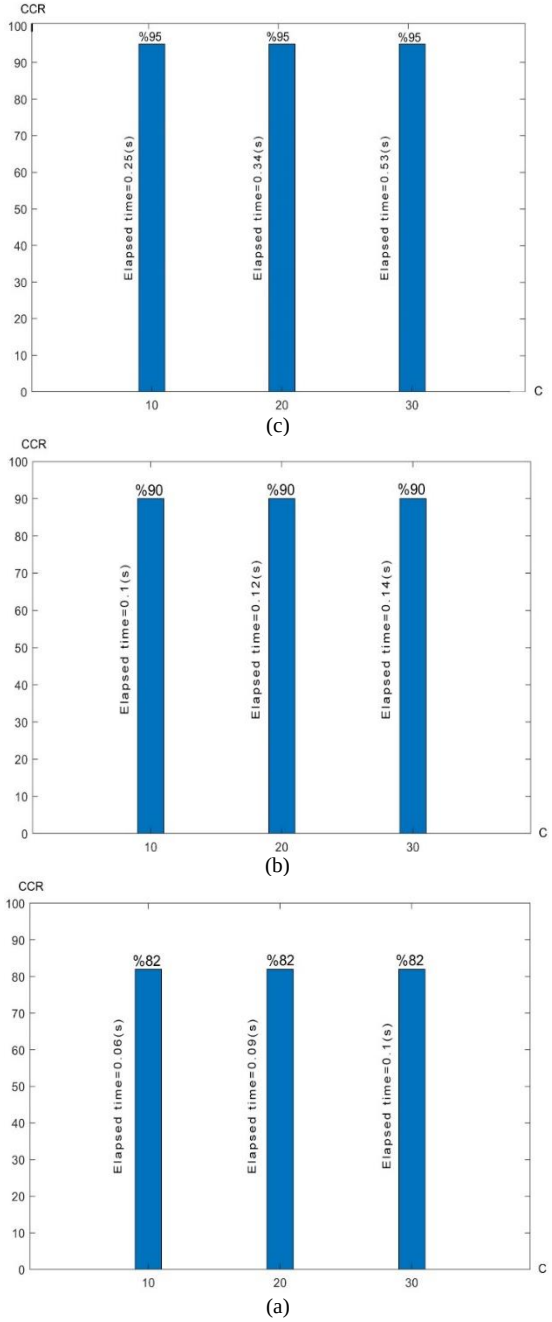


Figure 16. Accuracy bar chart of (a) RBF kernel function, (b) Polynominal kernel function, (c) Linear kernel function

3. CONCLUSIONS

The correct classification of the data obtained by the lymphogram has a great influence on detecting metastases and malign lymph. Data are classified by using Bayesian method, *K*-Nearest Neighbour method, Perceptron Neural Network and RBF Neural Network.

In the Bayesian method, the result is not desirable due to low amount of data and the form of data which is not fully Gaussian shaped. The solution obtained from the KNN method is better than the Bayesian method. In this method, it was observed that changing the value of *K* changes the value of the *CCR*. Because of the old method of KNN, the Neural Network method is used to obtain more favorable responses. Two methods of Perceptron

Neural Network and RBF Neural Network for data classification have been used. The Perceptron method changes the value of *CCR* by changing the number of neurons and the type of transfer function. In the RBF Neural Network, Radial Basis functions are used, which increases the value of *CCR* by increasing the number of Radial Basis functions.

Table 5. Time and *CCR* of each method

Method	Bayesian	KNN	Perceptron	RBF
CCR (%)	9.09	90.91	99.72	100
Time(s)	0.48	0.03	1.28	1.88

According to the Table 5, the best *CCR* obtained from the RBF method, but in comparison with the time, the KNN method is obtained in the shortest possible time.

NOMENCLATURES

A. Acronyms

KNN *K*-Nearest Neighbor
CCR Correct Classification Rate
RBF Radial Basis Functions
NRBF Number of Radial Basis Functions
NN Number of Neurons
Diff Diffusion Matrix
ANN Artificial Neural Network

B. Symbols / Parameters

P: Probability of each test data
σ^2: Variance
μ: Mean
w: Weight
b: Bias

ACKNOWLEDGEMENTS

This lymphography domain was obtained from the University Medical Centre, Institute of Oncology, Ljubljana, Yugoslavia. Thanks go to M. Zwitter and M. Soklic for providing the data.

REFERENCES

[1] L. Shao, T. Ouchi, M. Sakamoto, S. Mori, T. Kodama, "Activation of Latent Metastases in the Lung after Resection of a Metastatic Lymph Node in a Lymph Node Metastasis Mouse Model", Biochem. Biophys. Res. Commum, No. 460, pp. 543-548, 2015.

[2] L. Shao, S. Mori, Y. Yagishita, T. Okuno, Y. Hatakeyama, T. Sato, T. Kodama, "Lym-Phatic Mapping of Mice with Systemic Lymphoproliferative Disorder: Usefulness as an Inter-Lymph Node Metastasis Model of Cancer", J. Immunol. Methods, No. 389, pp. 69-78, 2013.

[3] A. Guermazi, P. Brice, C. Hennequin, E. Sarfati, "Lymphography: An Old Technique Retains its Usefulness", Radiographics, Vol. 23, No. 6, pp. 1541-1458, 2003.

[4] "Guyton and Hall Textbook of Medical Physiology", Saunders, pp. 186-187, 2010.

[5] S. Chandrasekaran, M.F. Chan, J. Li, M.R.King, "Super Natural Killer Cells that Target Metastases in the Tumor Draining Lymph Nodes", Biomaterials, doi:10.1016/j.biomaterials.

[6] A. Darwiche, "Modelling and Reasoning with Bayesian Network", Cambridge University Press, Cambridge, 2009.

[7] L.J. Yan, N. Cercone, "Bayesian Network Modeling for Evolutionary Genetic Structures", Computers and Mathematics with Applications, No. 59, pp. 2541-2551, 2010.

[8] F. Latifoglu, K. Polat, S. Kara, S. Gunes, "Medical Diagnosis of Atherosclerosis from Carotid Artery Doppler Signals Using Principal Component Analysis (PCA), K-NN Based Weighting Pre-Processing and Artificial Immune Recognition System (AIRS)", J. Biomed. Inform., No. 41, pp. 15-23, 2008.

[9] M. O'Farrell, E. Lewis, C. Flanagan, N. Jackman, "Comparison of K-NN and Neural Network Methods in the Classification of Spectral Data from an Optical Fibre-Based Sensor System Used for Quality Control in the Food Industry", Sens. Actuators B: Chemical, No. 111-112C, pp. 354-362, 2005.

[10] J. Schmidhuber, "Deep Learning in Neural Networks: An Overview", Neural Networks, No. 61, pp. 85-117, 2015.

[11] D. Graupe, "Principles of Artificial Neural Networks", World Scientific Publishing Co. Pte. Ltd., Vol. 6, 2nd Edition, 2007.

[12] R. McDonald, K. Hall, G. Mann, "Distributed Training Strategies for the Structured Perceptron", Association for Computational Linguistics, pp. 456-464, 2010.

[13] M. Olazaran, "A Sociological Study of the Official History of the Perceptrons Controversy", Social Studies of Science, Vol. 26, No. 3, pp. 611-659, 1996.

[14] H. Mohamed, A. Negm, M. Zahran, O.C. Saavedra, Assessment of Artificial Neural Network for Bathymetry Estimation Using High Resolution Satellite Imagery in Shallow Lakes: Case Study El Burullus Lake", International Water Technology Conference, 2015.

[15] M.D. Buhmann, "Radial Basis Functions: Theory and Implementations", Cambridge University, 2003.

[16] R.P. Borghate1, S.K. Gosavi, "Frequency Offset Compensation in OFDM System Using Neural Network", International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, Vol. 3, Issue 7, July 2014.

[17] X. Yan, M. Jia, "A Novel Optimized SVM Classification Algorithm with Multi-Domain Feature and its Application to Fault Diagnosis of Rolling Bearing", Neurocomputing, 2018.

[18] N. Leal, E. Leal, G. Sanchez, "Marine Vessel Recognition by Acoustic Signature", ARPN Journal of Engineering and Applied Sciences, Vol. 10, No. 20, November 2015.

BIOGRAPHIES

Sepehr Sarabi was born in Tabriz, Iran, 1995. He received the B.Sc. degree from Tabriz Branch, Islamic Azad University, Tabriz, Iran in 2018. Currently, he is a M.Sc. student of Biomedical Engineering at the same university. His research interests areas are pattern recognition and image processing. He is a main member of the Scientific society of Biomedical Engineering.

Milad Asadnejad was born in Tabriz, Iran, 1996. He received the B.Sc. degree from Tabriz Branch, Islamic Azad University, Tabriz, Iran in 2018. Currently, he is a M.Sc. student of Biomedical Engineering at the same university. His research interests areas are pattern recognition and signal processing. He is a member of the Scientific Society of Biomedical Engineering.

Seyed Amir Tabatabaei Hosseini was born in Tabriz, Iran in 1991.He received his B.Sc. degree in 2018 in Biomedical Engineering. Currently, he is a M.Sc. student of Biomedical Engineering at Tabriz University of Medical Science (Tabriz, Iran). His research interests are in the area of machine learning and image processing.

Saman Rajebi was born in Tabriz, Iran in 1981. He received his B.Sc. degree in Electrical Engineering in 2003 and M.Sc. and Ph.D. degrees in Communication Engineering in 2006 and 2017, respectively. He is a Fulltime Teaching Staff of Seraj Higher Education Institute, Tabriz, Iran and is the member of Executive Committee of ICTPE conferences. His research interests areas are radio frequency electronics, microwave applications, antenna designing and neural networks.