ISSN 2348 – 8034
Impact Factor- 5.070

GLOBAL JOURNAL OF ENGINEERING SCIENCE AND RESEARCHES

A COMPARATIVE ANALYSIS OF VARIOUS ALGORITHMS TO PREDICT THE SURVIVAL OF LIVER GRAFT TRANSPLANT

Aditi Gupta[*1] & Dr. Dinesh Kumar[2]
[*1]Research Scholar, Department of Computer Science and Engineering, DAV Institute of Engineering and Technology, Jalandhar.
[2]Associate Professor and Head Department of Information Technology, DAV Institute of Engineering and Technology, Jalandhar

ABSTRACT

The artificial learning models such as ARTIFICIAL NEURAL NETWORK, RADIAL BASIS FUNCTION, and ARTMAP have shown a promising application in the medical industry. The present work is a comparative analysis of above mentioned. This study was conducted during January- May 2018, in the CG-1 POST GRAD. Laboratory, Department of Computer Sciences, DAV Institute of Engineering and Technology, Jalandhar (144001), Punjab, India

The results of this investigation have indicated that among ARTIFICIAL NEURAL NETWORK, RADIAL BASIS FUNCTION, and ARTMAP the numeric values obtained from ARTIFICIAL NEURAL NETWORK werec omparatively better. Further, the analysis of the accuracy among the three selected algorithms was found 98.9708%, 97.2556%, and 58.1475% respectively. According to literature survey performed, it is evident that most studies in this regard have received lesser attention, especially in India. Based on our findings it seems that ARTIFICIAL NEURAL NETWORK could be the best mode while predicting the graft survivals during liver transplantation.

Keywords: *Artificial Neural Network,RADIAL BASIS FUNCTION,ART MAP,Liver Transplant(LT).*

I. INTRODUCTION

Most of the work in Artificial Neural Networks is being carried out across the globe[1]. However, the data from Indian medical field is yet to be tested for these algorithms[2]. Researchers face more challenging task in healthcare sectors to predict the diseases from the voluminous medical databases[3]. Nowadays data mining has become more essential in healthcaresectors[4]. Datamining technique includes classification, clustering, and association rule mining for finding frequent patterns applied to medical data for disease prediction[2].

Indatamining, classification techniques play a vital role in medical diagnosis and predicting diseases. In this research work, Naïve Bayes and Support Vector Machine (SVM) classifier algorithms are used for liver diseaseprediction[5]. The liver is the second largest internal organ in the human body, playing a major role in metabolism and serving several vital functions, e.g. decomposition of red blood cells, etc. It weighs around three pounds[3]. The liver performs many essential functions related to digestion, metabolism, immunity, and the storage of nutrients with in the body[3]. These functions make the liver an important organ, without which body tissues would quickly die due to lack of energy and nutrients[3].

For instance, several studies have derived an expression using stepwise logistic regression analysis to find out the probability of graft failure in the liver patient[6]. They evaluated their results with the help of Receiver Operating Characteristics(ROC) curve analysis using Labroc 1 software. But the authors did not succeed to provide accuracy in the prediction of survival after LT with lack of large datasets. Also, in this study ROC for prediction was used but a comparative study of ARTIFICIAL NEURAL NETWORK, Radial basis function, and ARTMAP for Data Mining and this proves that ARTIFICIAL NEURAL NETWORK data mining study is superiorto ROC[7].

While, every year 2, 00,000 people die of liver diseases, it is the most active organ of the body that performs multifarious critical functions in the body is, unfortunately, the most taken for granted organ[8].The causes for liver problems are many such as decompensated cirrhosis, progressive hepatitis 'B' or 'C' Alcohol damage, Fatty liver disease abnormality of billiard system etc. etc. In acute cases, a liver transplant is suggested for an end-stage liver disease. But the liver transplant is governed by many bodily factors such as Body mass Indies, blood group, creatinine, albumin and other Biochemical tests. In addition to this, the other factors that complicate the process are the availability of a donor, medical urgency of the patient and geographical proximity of donor, age, sex etc[3]. The computer technology comes in handy to provide accuracy speed, predicting survival rate post-surgery complication, possible remedies and prioritize patients[9].

In the light of above literature, we planned this study entitled "A Comparative Analysis of Various Algorithms to Predict the Survival of Liver Graft Transplant".

II. MATERIALS AND METHODS

The current study was conducted during Jan–May 2018, at the laboratory of CG-1 POST GRAD LAB, Department of computer sciences, D.A.V. Institute of engineering and technology, Jalandhar (144001), Punjab, India.

2.1 Process of data collection

The data used in this study were retrieved from the liver transplant dataset available at Indian Transplant Registry - www.transplantindia.com and the following archives: -

1. https://archive.ics.uci.edu/ml/machine-learning-databases/00225/
2. https://archive.ics.uci.edu/ml/datasets/ILPD+(Indian+Liver+Patient+Dataset).

2.2 Tools for comparative analysis

The current study was implemented in WEKA 3.8 TOOL. The results were evaluated using Multilayer Perceptron Artificial Neural Networks with 10-fold cross-validation. The whole data was divided into training data and test data which gives an accuracy of 100% by Multilayer Perceptron Artificial Neural Network model. Such approach is believed to reduce the post-transplantation mortality rate by using an intelligent system that can find correct donor-recipient pairs from a pool of donor-recipient data.

2.3 Artificial Neural Networks Algorithm

Firstly we initialized the weight, bias, and learning rate and then checked the stopping condition if it was false we performedbipolar or binary training vector pairs: t.
Set activation of each input unit i=1 to n:

$x_i = s_i$

Then we calculated the output response of each output unit j=1 to m: First, the net input was calculated as
$y_in_j = b_j +$

Then activations were applied over the net input calculate the output response:
$y_j = \; =$
Adjustments were made in weights and bias for j = 1 to m and i=1 to n.

If $t_j = y_j$ then

$w_{ij}\,(new) = w_{ij}\,(old) + \; t_j x_i$

$b_j\,(new) = b_j(old) + t_j$

Else, we have $w_{ij}\,(new) = w_{ij}\,(old),\; b_j(new) = b_j\,(old)$

Test for the stopping condition, i.e. if there was no change in weights then we stop the training process else started again from activation.

2.4 Radial basis function algorithm

We initialized the weight, bias, and learning rate. Thenwe checked the stopping condition if it was false we performed the input unit (x_i for all i=1 to n) to receive input signals and transmit to the next hidden layer. We then selected the centers for the radial basis function. The centers were selected from the set of input vectors. It may be noted that a sufficient number of centers were selected to ensure an adequate sampling of the input vector space.

We calculated the output from the hidden layer unit:

$V_i(x)_i =$ **….**

Where x_{ji} was the center of the RADIAL BASIS FUNCTION unit for the input variable: the width of ith RADIAL BASIS FUNCTION unit: x_{ij} the jth variable of the input pattern.

We calculated the output of the neural network:

$y_{net} =$**….**

We calculated the error and test for the stopping condition. The stopping condition was the number of epochs or to a certain extent of weight change.

2.5 ARTMAP Algorithm

Learning rate wasinitialized (vigilance parameter and error) and then we checked the stopping condition if it was false we set activations of all F_1(a) and F_1 units as follows

F_2 = 0 and F_1(a) = input vectors

The input signal from F_1 (a) to F_1 (b) layer was sent like

si=xisi=xi …..

For every inhibited F_2 node

yj=∑ibijxiyj=∑ibijxi the condition is **$y_j \neq -1$…..**

We the performed step 8-10 when the reset was true.

Find **J** for **$y_J \geq y_j$** for all nodes **j**

We again calculated the activation on F_1 (b) as follows

xi=sitJixi=sitJi**….**

Now, after calculating the norm of vector **x** and vector **s**, we checked the reset condition as follows −

If ||**x**||/ ||**s**|| < vigilance parameter **ρ**,then inhibitnode **J** and go to step 7

Else If ||**x**||/ ||**s**|| ≥ vigilance parameter **ρ**, then proceeded further.

Weight updating for node **J** was done as follows −

bij(new)=αxiα−1+||x||bij(new)=αxiα−1+||x||

tij(new)=xitij(new)=xi**…..**

The stopping condition for algorithm was checked as follows −

- No change in weight.
- Reset of units.
- The maximum number of epochs reached.

2.6 Determination of Accuracy of algorithms (ARTIFICIAL NEURAL NETWORK, RADIAL BASIS FUNCTION, and ARTMAP)

ARTIFICIAL NEURAL NETWORK was used for the comparison of performance and accuracy with RADIAL BASIS FUNCTION and ARTMAP in our study[10]. This was due to the stable nature and high training speed, ARTIFICIAL NEURAL NETWORK is superior to all other models and the accuracy is far better than ARTMAP. ARTIFICIAL NEURAL NETWORK is more sensitive to data noise and to the order of presentation of input patterns.

2.7 Parameters

- **KAPPA Statistic:** - Cohen's **kappa coefficient** (κ) is a **statistic** which measures inter-rater agreement for qualitative (categorical) items. It is generally thought to be a more robust measure than simple percent agreement calculation, as κ takes into account the possibility of the agreement occurring by chance[11].
- **MEAN ABSOLUTE ERROR**(MAE) is a measure of the difference between two continuous variables. ... Allocation Disagreement is MAE minus Quantity Disagreement. The **Mean Error** is given by It is also possible to identify the types of difference by looking at a plot[12].
- **ROOT-MEAN-SQUARE**deviation (RMSD) or **root-mean-square error** (**RMSE**) (or sometimes **root-mean-squared error**) is a frequently used measure of the differences between values (sample and population values) predicted by a model or an estimator and the values actually observed[13].
- The **absolute error** is the magnitude of the difference between the exact value and the approximation. The **relative error** is the **absolute error** divided by the magnitude of the exact value. The percent **error** is the **relative error** expressed in terms of per 100[12].
- The **Relative absolute error** (and analogically **Root relative squared error**) is calculated as the Mean **absolute error** divided by the **error** of the ZeroR classifier (a classifier, that ignores all predictors and simply selects the most frequent value)[12].
- TP RATE: The fundamental prevalence-independent statistics are sensitivity and specificity. Sensitivity or True Positive **Rate** (TPR), also known as recall, is the proportion of people that tested positive and are positive (True Positive, **TP**) of all the people that actually are positive (Condition Positive, CP = **TP** + FN)[14].
- FP Rate: where **FP** is the number of false positives, TN is the number of true negatives and N=**FP**+TN is the total number of negatives. ... The **false positive rate** (or "false alarm **rate**") usually refers to the expectancy of the **false positive** ratio[14].
- Precision = TP / (TP+FP)[15].
- Recall = TP / (TP+FN)[15].
- The **F measure** (**F1 score** or **F score**) is a **measure** of a test's accuracy and is defined as the weighted harmonic mean of the precision and recall of the test[15].
- MCCC: Matthews Correlation Coefficient Interpretation[16].
- ROC: In statistics, a receiver operating characteristic curve, i.e. *ROC curve*, is a graphical plot that illustrates the diagnostic ability of a binary classifier system[15].

III. RESULTS

The following results have also been summarized in Table 1.

3.1The comparative analysis

Comparative analysis for Kappa Statistics in the ARTIFICIAL NEURAL NETWORK, RADIAL BASIS FUNCTION, and ARTMAP was found to be 0.979, 0.94552, and 0 respectively. The Mean Absolute Error for the above three was found to be 0.0116, 0.0184, and 0.328 respectively. The Root Mean Squared Error was found to be 0.0803,0.1353, and 0.04046 respectively. The Relative Absolute Error was noted 3.5371%, 5.6213%, and 100% respectively. The Root Relative Absolute Error19.8605%, 33.434%, and 100% respectively.

3.2 The accuracy of algorithms (ARTIFICIAL NEURAL NETWORK, RADIAL BASIS FUNCTION, and ARTMAP)

The accuracy value for Artificial Neural Network (Figure 1), Radical Basis Function (Figure 2) and ARTMAP (Figure 3) was found to be 98.9708%, 97.2556%, and 58.1475% respectively.

ISSN 2348 – 8034
DOI- 10.5281/zenodo.1287716
Impact Factor- 5.070

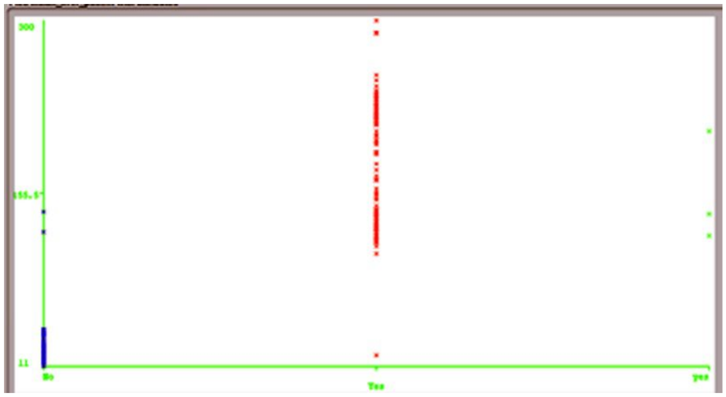

Figure 1: Showing the accuracy value for Artificial Neural Network

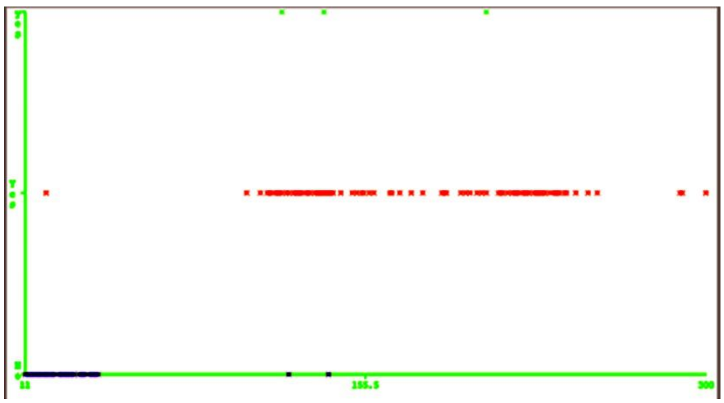

Figure 2: Showing the accuracy value for Radical Basis Function

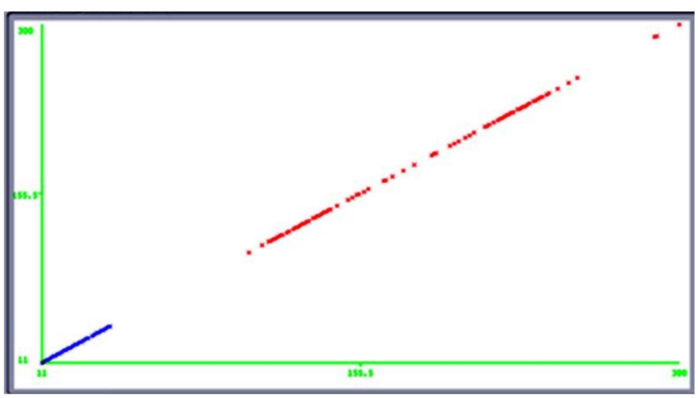

Figure 3: Showing the accuracy value for ARTMAP

Table: 1Showing the comparative analysis of Artificial Neural Network, Radical Basis Function, and ARTMAP with respect to various parameters.

Sr. No.	Parameters	Artificial Neural Network	Radical Basis Function	ARTMAP
1	Kappa statistics	0.979	0.9452	0
2.	Mean Absolute error	0.0116	0.0184	0.328
3.	Root mean squared error	0.0803	0.13535	0.4046
4.	Relative absolute error	3.5371%	5.6213%	100%

ISSN 2348 – 8034
Impact Factor- 5.070

5.	Root-relative absolute error	19.8605%	33.434%	100%
6.	TP Rate	1.000	1.000	0.000
7.	FP Rate	0.000	0.000	0.000
8.	Precision	1.000	1.000	0.000
9.	Recall	1.000	1.000	0.000
10.	F-Measure	1.000	1.000	0.000
11.	MCCC	1.000	1.000	0.000
12.	ROC Area	1.000	1.000	0.497

IV. DISCUSSIONS

The results of the present study indicated that among Artificial Neural Network, Radial Basis function, and ARTMAP, The Artificial Neural Network yielded best results. On comparing ARTIFICIAL NEURAL NETWORKs, RADIAL BASIS FUNCTION and ARTMAP indicated that ARTIFICIAL NEURAL NETWORKs are more superior and beneficial than its counterpart. ARTIFICIAL NEURAL NETWORKs function in a sequential and logical order[10]. Firstly, they adopt the sound theory and then proceed towards implementation and experimentation whereas, RADIAL BASIS FUNCTION and ARTMAP adopt "a hands-on" approach where application and experiments are given more priority and theory is involved later[10]. ARTIFICIAL NEURAL NETWORKs have several other benefits over RADIAL BASIS FUNCTION and ARTMAP such as ARTIFICIAL NEURAL NETWORKs involve simple geometric interpretation and gives a solution which is viable, while RADIAL BASIS FUNCTION and ARTMAP suffer from multiple complexities and its solutions are limited to a local level only[17]. Unlike RADIAL BASIS FUNCTION and ARTMAP, the computational complexity of ARTIFICIAL NEURAL NETWORKs does not dependent upon the input space[18]. On one hand, ARTIFICIAL NEURAL NETWORKs useempiricalrisk minimization while on the other hand, ARTMAP and RADIAL BASIS FUNCTION use structural risk management[19]. The main reason for ARTIFICIAL NEURAL NETWORKs to be more popular and preferred is their capacity to overcome the biggest problem experienced with ARTMAP and RADIAL BASIS FUNCTION, i.e overfitting[20]. In the nutshell, the ARTIFICIAL NEURAL NETWORK classifier is generally considered as the best algorithm because of its highest classification accuracy[10]and the results of the present study are also incoherence with the same.

Our model trained many instances and predicted the success rate of patients after LT successfully. Out of 100 % patients, 70 % of patients were alive after LT without any difficulty. The post-transplantation outcome of each patient depends upon the pre-transplantations tate of the patient, the graft quality and the complication of surgery[10]. Sometimes the complications occur instantly after surgery or in the long run. When complications occur, stay in the intensive care unit is extended and the chance of mortality increases[21].

Currently, collection and sharing of liver organs are the most important aspects of LT[22]. The scarcity of donors is the main problem faced by patients and every organ allocation has to be accurate in such a scenario[10]. LT has advanced from an experimental therapy to a mainstream treatment option for a wide range of acute and chronic liver diseases[22].Now a day's LT is one of the challenging areas in the field of organ transplantation[22]. The liver gets damaged due to not only alcohol or liver disease but also it gets affected by improper food usage as well as genetic disorders[23]. Clinical studies showed that in the next decade, more than 90% of people will be affected by liver problems[22-23]. The prediction by medical experts is based on MELD score, but not all the time the MELD score will give the exact outcome. The components of MELD score includes Bilirubin, Creatinine, and INR, out of which the creatinine value will be changing according to the body weight of the patient[24]. With the same dataset, the graft survival rate is 79.11% and graft failure rate is 20.89% using MELD score. In addition to the MELD score, various machine learning techniques are introduced for the forecasting of increased survival after LT[10]. ARTIFICIAL NEURAL NETWORK is a new biologically inspired computing approach which is a very powerful advancement in the field of computers and medicine. In order to perform machine learning operations in engineering, medicine, mathematics, economics, science, geology and many others, the role of artificial neural networks has been very successful[25].

V. CONCLUSION

It has been observed that classification of liver diseases is more accurate in Artificial Neural Network data mining than Radial Basis Function and ARTMAP modality but at the same time is costlier. Artificial Neural Network prediction of the liver provides a good basis for analyzing the texture of the liver, whereasRadial Basis Function and ARTMAP impose some difficulties to analyze the structure of the liver.Further analyzing texture is a challenge but Radial Basis Function and ARTMAP are cost effective. Researchers have been made in all the three modalities to diagnose the diseases related to the liver. Artificial Neural Network techniques provide good accurate results according to our study. A lot can be improved in the accuracy of diagnosing the liver diseases based on texture analysis. The accuracy completely depends on classifiers applied. The results may improve by applying a combination of classifiers

VI. ETHICAL STATEMENT

The research presented in the study is true and best to my knowledge. None of the data has been manipulated or tampered with in any respect.

VII. ACKNOWLEDGMENTS

The authors are sincerely thankful to the Department of Computer Sciences,D.A.V.Institute of Engineering and Technology, Jalandhar (144001), Punjab, Indiafor providing the necessary infrastructure facilities during the tenure of this study.

REFERENCES

1. *Brooks, Rodney A. "Intelligence without representation." Artificial intelligence (1991); 47(1-3) : 139-159.*
2. *Kaur, Harleen, and Siri Krishan Wasan. "Empirical study on applications of data mining techniques in healthcare." Journal of Computer science (2006); 2(2): 194-200.*
3. *Vijayarani, S., and S. Dhayanand. "Liver disease prediction using SVM and Naïve Bayes algorithms." International Journal of Science, Engineering and Technology Research (2015); 4(4): 816-820.*
4. *Raghupathi, Wullianallur, and Viju Raghupathi. "Big data analytics in healthcare: promise and potential." Health information science and systems (2014): 2(1); 3.*
5. *Ramana, Bendi Venkata, M. Surendra Prasad Babu, and N. B. Venkateswarlu. "A critical study of selected classification algorithms for liver disease diagnosis." International Journal of Database Management Systems (2011); 3(2): 101-114.*
6. *Marino, Ignazio Roberto, et al. "Effect of donor age and sex on the outcome of liver transplantation." Hepatology (1995); 22(6): 1754-1762.*
7. *Chuang, Chun-Ling. "Case-based reasoning support for liver disease diagnosis." Artificial Intelligence in Medicine (2011); 53(1) : 15-23.*
8. *Murray, Christopher JL, Alan D. Lopez, and World Health Organization. "The global burden of disease: a comprehensive assessment of mortality and disability from diseases, injuries, and risk factors in 1990 and projected to 2020: summary." (1996).*
9. *Rana, A1, et al. "Survival outcomes following liver transplantation (SOFT) score: a novel method to predict patient survival following liver transplantation." American journal of transplantation (2008); 8(12): 2537-2546.*
10. *Raji, C. G., and SS Vinod Chandra. "Graft survival prediction in liver transplantation using artificial neural network models." Journal of Computational Science (2016);16: 72-78.*
11. *Ben-David, Arie. "Comparison of classification accuracy using Cohen's Weighted Kappa." Expert Systems with Applications(2008); 34(2): 825-832.*

12. *Willmott, Cort J., and Kenji Matsuura. "Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance." Climate research(2005); 30(1): 79-82.*
13. *Maiorov, Vladimir N., and Gordon M. Crippen. "Significance of root-mean-square deviation in comparing three-dimensional structures of globular proteins." Journal of molecular biology(1994); 235(2): 625-634.*
14. *Frey, Brendan J., and Delbert Dueck. "Clustering by passing messages between data points." science (2007); 315(5814): 972-976.*
15. *Powers, David Martin. "Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation." (2011).*
16. *Gorodkin, Jan. "Comparing two K-category assignments by a K-category correlation coefficient." Computational biology and chemistry(2004);28(5-6): 367-374.*
17. *Venkatesh, S., and S. Gopal. "Orthogonal least square center selection technique–a robust scheme for multiple source partial discharge pattern recognition using radial basis probabilistic neural network." Expert Systems with Applications(2011); 38(7): 8978-8989.*
18. *Jain, Anil K., Jianchang Mao, and K. Moidin Mohiuddin. "Artificial neural networks: A tutorial." Computer (1996); 29(3): 31-44.*
19. *Bottou, Léon, and Vladimir Vapnik. "Local learning algorithms." Neural computation (1992);4(6): 888-900.*
20. *Wu, Shelly Xiaonan, and Wolfgang Banzhaf. "The use of computational intelligence in intrusion detection systems: A review." Applied soft computing(2010); 10(1): 1-35.*
21. *Khuri, Shukri F., et al. "Determinants of long-term survival after major surgery and the adverse effect of postoperative complications." Annals of surgery(2005);242(3): 326.*
22. *Adam, René, et al. "Evolution of liver transplantation in Europe: report of the European Liver Transplant Registry." Liver transplantation(2003); 9(12): 1231-1243.*
23. *Bernal, William, et al. "Acute liver failure." The Lancet(2010); 376(9736): 190-201.*
24. *Fallon, Michael B., et al. "Model for end-stage liver disease (MELD) exception for hepatopulmonary syndrome." Liver transplantation (2006); 12(S3).*
25. *Hertz, John A. Introduction to the theory of neural computation. CRC Press, 2018*