International Journal of Chemical and Biochemical Sciences
(ISSN 2226-9614)

Journal Home page: www.iscientific.org/Journal.html

Chronic disease modeling - MATLAB environment

Irha Basit[1], Haroon Ur Rashid Kayani[2], Fatma Hussain[3]*

[1 3]Department of Biochemistry, Faculty of Sciences, University of Agriculture, Faisalabad, Pakistan and [2]Department of Software Engineering, Lahore Garrison University, Lahore, Pakistan

Abstract

Chronic diseases are prevailing at an alarming rate leading to morbidity and mortality. Effective diagnostic and therapeutic strategies must be implemented to curb such diseases worldwide. MATLAB (matrix laboratory), a multi-pattern numerical computational program is solving large complex problems in many fields including biological sciences, chemistry, mathematics, and engineering. It is used as an important tool for the detection and prevention of diseases. On this platform, several tools and models are used which gives information after data analysis. Some other tools are also used such as WEKA (Waikato Environment for Knowledge Analysis), artificial neural network, J48 algorithm and Java implementation LVQ (Learning vector quantization). There are many updated versions of MATLAB that are now-a-days being used. Each updated version is better than the previous one. All these softwares are indulged in data analysis from the previous history and predicting the onset of the disease.

Keywords: Artificial neural network, fuzzy logic, machine learning, fuzzy expert system, algorithms

Full length article **Corresponding Author*, e-mail: fatmauaf@yahoo.com

1. Introduction

MATLAB (matrix laboratory), a multi-pattern numerical computational program is solving large complex problems in many fields including biological sciences, chemistry, mathematics, and engineering [1]. It is used as an important tool for the detection and prevention of disease. On this platform, several tools and models are used which gives information after data analysis. It's graphical features facilitate in depiction of conclusions [2].

1.1 Computer-Based Disease Diagnosis

CAD (computer-aided diagnosis) system is designed to follow the mathematical functions provided by Mathworks Inc. [3]. It is used for the detection of neurodegenerative diseases (Alzheimer disease, Parkinson disease) and diabetes mellitus. MATLAB also includes an artificial neural network model with mathematical and computational strategies. It is a biological neural network. It works based on different attributes. For the prediction of heart disease, there are 15 attributes. At first, 13 attributes were used but after research, two additional factors obesity and smoking were included for efficient disease prediction [4]. In MATLAB, fuzzy expert system is now being used for providing a good evaluation. Fuzzy logic and tables are based on knowledge in healthcare for the diagnosis of disease. The Mamdani model of fuzzy system and UCI machine learning repository have excellent results and these can predict the disease with only 6 attributes [5].

1.2 Heart Disease

Ephzibah and Sundarapandian [6] proposed a system to predict heart diseases. Input system has all the selected features which are required for the prediction of disease and output includes the information which diagnose disease. Date is converted into fuzzy linguistics variables, its terms and data mining functions. After this step, defuzzification steps are performed.

1.3 J48 Classification Algorithm–Diabetes mellitus

The decision tree algorithm provides information about the relationship between attribute-vectors. Based on practices, new illustrations can be found out [6]. This kind of algorithm creates the rules and predicts the target variable. J48, an updated ID3 has few extra characteristics that fill the absent values, decision trees pruning, commencement of rules. It has WEKA (Waikato Environment for Knowledge Analysis) mining tool. J48 uses the JAVA implementation of the C4.5 algorithm. Tree

pruning is also associated with it. For more accurate results over-fitting and pruning is done. In some algorithms, sorting is done to deliver maximum pure leaf. It creates the trees to analyze the data and results are given [7]. Researchers have given a new approach for predicting diabetes from the patient's history. For this purpose, Indian diabetic patients were recruited along with non-diabetic subjects. The updated version of J48 gives more accurate results. To generate J-48 modifies, WEKA, is used as API (Application Program Interface). This was a big achievement as it gave 99.87% accurate results [8, 9]. Accuracy and errors of different algorithms are shown below (Fig.1)

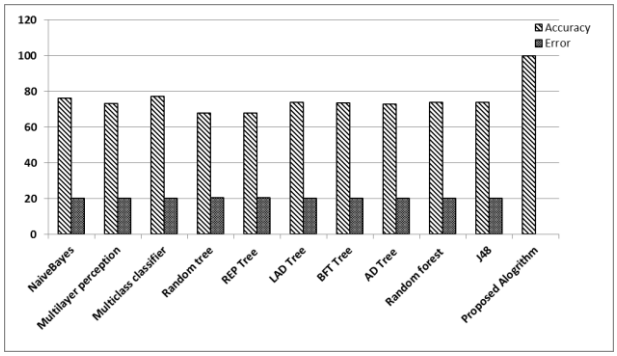

Fig.1 Accuracy and error of different programs

1.4 Parkinson Disease (PD)

Neuropathology and histopathology are the two criteria based on which PD is diagnosed. The sensitivity and specificity of the clinical features are the factors on which the PD classification is done. PD cannot be predicted by checking just one factor because it is a complicated disease that interconnects many perspectives [10]. Two basic classifiers are KStar and IBK. Diagnostic and predictive value of various medical biometrics determines presence of PD. Some other software used are, WEKA 3 and Mat lab v7 [11].

1.5 Kidney Disease

Learning vector quantization (LVQ) algorithm and radial basis function are used for the diagnosis of kidney stone disease. Different models are used to excel in accuracy. WEKA tool and multilayer perception is the best method [12].

1.6 Liver Disease

This disease can be predicted by support vector machine (SVM) and Naïve Bayes algorithms. SVM has high accuracy while Naive Bayes takes less time for a prediction [13]. Some intelligent techniques were also used to identify liver patients including WEKA mining tools, J48, SVM, Bayesian network and random forest. In these tools, the percentage of accuracy was observed and the Random forest provided the highest value of accuracy and it was 71.35 percent [14]. Mining techniques such as FT tree, Naïve Bayes and K star were used to analyze liver disease based on seven attributes. The highest accuracy rate was given by FT tree [15].

1.7 Dengue

Dengue under certain conditions can become critical enough to cause death. Therefore, best simulation model should be made into consideration. For dengue, scientists used a Decision tree, ANNs, and Rough set theory (RS). RS provided excellent results with 99.72 percent accuracy [16]. For the prediction of Arbovirus-Dengue, support vector machine tool was used and accuracy was 0.9042 [17]. Artificial neural network for dengue hemorrhagic fever (DHF) gave 90 percent correctness [18].

1.8 Hepatitis

Ft tree, K star, J48, LMT, Naïve Bayes, Naïve Bayes updatable and ANNs are the data mining algorithms that were used for the prediction of hepatitis. A comparative analysis revealed that Naïve Bayes have best accuracy of 96.52 percent in 0 seconds [19]. Comparative analysis between Naïve Bayes and back propagation classifiers was made to diagnose Hepatitis. For this, the data set was made with 20 attributes and 155 instances. Out of two, Naïve Bayes gave 97 percent accuracy [20]. Table 1 highlights the contribution of different computational models in healthcare.

Table 1 Literature Review

Reference	Contribution
Pan and Tompkins, 1985 [21]	Real-Time signals for QRS complexes of ECG signals
Speckt, 1990 [22]	Probabilistic neural networks. It is also known as Bayesian classifiers
El-Solh *et al*., 1999 [23]	Diagnostic tool for chest diseases
Ashizawa *et al*., 1999 [24]	Artificial neural networks - lung diseases
Aliferis *et al*., 2002 [25]	Machine learning models for lung cancer classification
Shanthi *et al*., 2009 [26]	Prediction of thromboembolic stroke IJBB
Hanif *et al*., 2009 [27]	Artificial neural network to check the severity of asthma and its prevention
Temurtas, 2009 [28]	Diagnosis of thyroid disease by neural networks
Er *et al*., 2010 [29]	Diagnosis of tuberculosis by neural networks
Anbarasi *et al*., 2010 [30]	Prediction of heart disease with genetic algorithms
Yeh *et al*., 2011 [31]	Detection of cerebrovascular disease using data mining
Udak and Mfon, 2013 [32]	Fuzzy support system for the detection of Cholera
Hashmi and Khan, 2015 [33]	Diagnostic blood test for hepatic diseases using fuzzy logic
Badnjevic *et al*., 2015 [34]	Asthma and chronic obstructive pulmonary disease (COPD) categorization
Lauraitis *et al*., 2018 [35]	For the detection of Huntington disease

2. Conclusion

The incidence of chronic diseases at an exponential rate and progression has led to the development of numerous softwares. MATLAB (matrix laboratory), a multi-pattern numerical computational program is solving large complex problems in many fields including biological sciences, chemistry, mathematics and engineering. MATLAB programs deliver early diagnosis and treatment options. These include WEKA, Artificial neural network, J48 algorithm, Java implementation LVQ and many more. Each having its pros and cons has variable benefit for healthcare sector.

References

[1] J. Pan and W. Tompkins. (1985). Real time algorithm detection for QRS. IEEE Trans. Eng Biomed Eng, 32(3): 230-236.

[2] H.W. Hethcote. (2000). The mathematics of infectious diseases. SIAM review, 42(4): 599-653.

[3] R. Mikut., M. Reischl., O. Burmeister and T. Loose. (2006). Data Mining in medical time series. Biomedizinische Technik. 51: 288-293.

[4] C. S. Dangare and S. S. Apte. (2012). Improved study of heart disease prediction system using data mining classification techniques. Int J Computer Appl. 47(10): 44-48.

[5] K. Noh., H. G. Lee., H.S. Shon., B.J. Lee and K. H. Ryu. (2006). Associative classification approach for diagnosing cardiovascular disease. In Intelligent computing in signal processing and pattern recognition, Springer, 345: 721- 727.

[6] E.P. Ephzibah and D.V. Sundarapandian. (2012). Framing fuzzy rules using support sets for effective heart disease diagnosis. Int J Fuzzy Logic Sys, 2.

[7] T.S. Korting. (2006). C4. 5 algorithm and multivariate decision trees. Image Processing Division, National Institute for Space Research–INPE Sao Jose dos Campos–SP, Brazil.

[8] A. Nadali, E.N. Kakhky and H.E. Nosratabadi. (2011). Evaluating the success level of data mining projects based on CRISP-DM methodology by a Fuzzy expert system," Electronics Computer Technology (ICECT), 2011 3rd International Conference on Electronics Computer Technology. IEEE, 6: 161,165.

[9] J. Quinlan. (1996). Learning decision tree classifiers. ACM Computing Surveys (CSUR), 28(1):71-72.

[10] D. Aarsland., K. Andersen., J.P. Larsen and A. Lolk. (2003). Prevalence and characteristics of dementia in Parkinson disease: an 8-year prospective study. Arch Neurol. 60(3): 387.

[11] F. Aström, and R. Koker. (2011). A parallel neural network approach to prediction of Parkinson's Disease." Expert Sys Applications. 38(10): 12470-12474.

[12] K. Kumar and B. Abhishek. (2012). Artificial neural networks for diagnosis of kidney stones disease. Int J Inform Technol Comp Sci. 20-2

[13] S. Vijayarani and S. Dhayanand. (2015). Liver disease prediction using SVM and Naïve Bayes algorithms. Int J Sci Eng Technol Res. 4(4): 816-820.

[14] A. Gulia., R. Vohra and P. Rani. (2014). Liver patient classification using intelligent techniques. Int J Comp Sci Inform Technol. 5(4): 5110-5115.

[15] P. Rajeswari and G.S. Reena. (2010). Analysis of liver disorder using data mining algorithm. Global J Comp Sci Technol.

[16] N.D.A Tarmizi., F. Jamaluddin., A. Abu Bakar., Z.A. Othman., S. Zainudin and A.R.Hamdan. (2013). Malaysia dengue outbreak detection using data mining models. J Next Generation Information Technology (JNIT). 4(6): 96-107.

[17] A. Fathima and D. Manimegalai. (2012). Predictive analysis for the arbovirus-dengue using svm classification. Int J Eng Technol. 2(3):521-527.

[18] F. Ibrahim., M.N. Taib., W.A.B.W.Abas., C.C. Guan and S. Sulaiman. (2005). A novel dengue fever (DF) and dengue hemorrhagic fever (DHF) analysis using artificial neural network (ANN). Comp Methods Prog Biomed.79(3): 273-281.

[19] F.M. Ba-Alwi and H.M. Hintaya. (2013). Comparative study for analysis the prognostic in hepatitis data: data mining approach. Spinal Cord.11: 12.

[20] B. Karlik. (2012). Hepatitis disease diagnosis using backpropagation and the naive bayes classifiers. IBU J Sci Technol.1(1).

[21] J. Pan and W. Tompkins. (1985). Real Time Algorithm detection for QRS. IEEE Trans. Eng. Biomed Eng. 32(3): 230-236.

[22] D.F. Speckt. (1990). Probabilistic neural networks. Neural Networks. 3(1):109-118.

[23] A.A. El-Solh., C.-B. Hsiao., S. Goodnough., J. Serghani and B.J.B. Grant. (1999). Predicting active pulmonary tuberculosis using an artificial neural network. Chest. 116: 968-973.

[24] K. Ashizawa., T. Ishida., H. MacMahon., C.J. Vyborny., S. Katsuragawa and K. Doi. (1999). Artificial neural networks in chest radiography: Application to the differential diagnosis of interstitial lung disease. Acad Radiol. 6(1): 2-9.

[25] C.F. Aliferis., D. Hardin and P.P. Massion. (2002). Machine learning models for lung cancer classification using array comparative genomic hybridization. Proceedings of the AMIA symposium on biomedical informatics, Vanderbilt University, Nashville, TN, USA. pp. 7-11

[26] D. Shanthi., G.Sahoo and N.Saravanan. (2009). Designing an artificial neural network model for the prediction of thrombo-embolic stroke. Int J Biometric Bioinform. 3(1). 10-18

[27] N.H.H.M. Hanif., W.H. Lan., H.B. Daud and J. Ahmad. (2009). Classification of control measures for asthma using artificial neural networks. Computer. 1(2):3.

[28] F. Temurtas. (2009). A comparative study on thyroid disease diagnosis using neural networks Expert Sys Applica. 36: 944-949.

[29] O. Er., F. Temurtas and A.C. Tanrikulu. (2010). Tuberculosis disease diagnosis using artificial neural networks. J Med Sys. 34 (3): 299-302.

[30] M. Anbarasi., E. Anupriya and N. C. S. N. Iyengar. (2010). Enhanced prediction of heart disease with feature subset selection using genetic algorithm. Int J Eng Sci Technol. 2(10), 5370-5376.

[31] D.Y Yeh., C. H. Cheng and Y.W. Chen. (2011). A predictive model for cerebrovascular disease using data mining. Expert Sys Applications. 38(7): 8970-8977.

[32] U. A. Uduak and M.M. Mfon. (2013). Proposed fuzzy framework for cholera diagnosis and monitoring. Int J Comput Appl. 82 (17): 1-10.

[33] A. Hashmi and M.S. Khan. (2015). Diagnosis blood test for liver disease using fuzzy logic. Int J Sci: Basic Appl Res. 20(1):151-183.

[34] A. Badnjevic., M. Cifrek., D. Koruga and D. Osmankovic. (2015). Neuro-Fuzzy classification of asthma and chronic obstructive pulmonary disease. S1:1-9 BMC Med Inf Decision Making. 15(S3): S1.

[35] A. Lauraitis, R. Maskeliunas and R. Damasevicius. (2018). ANN and fuzzy logic based models to evaluate Huntington disease symptoms. J Healthcare Eng (2018).