Classifying Cardiotocography Data based on Rough Neural Network

Belal Amin[1]
Department of Technological Development
Tanta Branch
Worker University
Cairo, Egypt

A. A. Salama[3], Khaled Mahfouz[5]
Department of Mathematics and Computer Science
Faculty of Sciences
Port Said University
Port Said, Egypt

Mona Gamal[2]
Machine Learning and Information Retrieval Department
Faculty of Artificial Intelligence
Kafrelsheikh University
Kafrelsheikh, Egypt

I.M. El-Henawy[4]
Computer Science Department,
Faculty of Computers and Information
Zagazig University
Zagazig, Egypt

***Abstract*—Cardiotocography is a medical device that monitors fetal heart rate and the uterine contraction during the period of pregnancy. It is used to diagnose and classify a fetus state by doctors who have challenges of uncertainty in data. The Rough Neural Network is one of the most common data mining techniques to classify medical data, as it is a good solution for the uncertainty challenge. This paper provides a simulation of Rough Neural Network in classifying cardiotocography dataset. The paper measures the accuracy rate and consumed time during the classification process. WEKA tool is used to analyse cardiotocography data with different algorithms (neural network, decision table, bagging, the nearest neighbour, decision stump and least square support vector machine algorithm). The comparison shows that the accuracy rates and time consumption of the proposed model are feasible and efficient.**

Keywords—Accuracy rate; cardiotocography; data mining; rough neural network; WEKA tool

I. Introduction

Dealing with uncertain and inconsistency data in diagnosing diseases is a very challenging problem in medical field. Cardiotcography (CTG) [1, 2, 3 and 4] is one of the most common diagnostic devices in the last few decades representing features of fetus Heart Rate (FHR) and Uterine Contraction (UC) during pregnancy. The features are organized in a dataset with 21 input attributes and 3 classes of fetus state classified into Normal, Suspicious and Pathologic.

CTG is probably the most widely used technique in all obstetrics. It was introduced by Orvan Hess and Ed Hon at Yale University in 1957 [5]. Before that, the only device used was a stethoscope to determine fetal status and maternal uterine contractions. Therefore, the birth process was as a black box. Before 2008, fetal heart rate was classified as either reassuring or non-reassuring. The NICHD Workgroup [6] proposed a terminology of a three-tiered system to replace the older; they were normal, indeterminate and abnormal. In 2015 FIGO [6] updated the terminology of CTG monitoring device into normal, pathological and suspicious states. The device has several benefits for patients [7, 6, 5]. For example, it helps doctors monitoring more than one patient at the same time, predicting whether the mother needs a cesarean section or not, detecting low and high risk for patients in labour to make decisions quickly. Hence, Fetal Heart Rate (FHR) monitoring remains a widely used method for detecting changes in fetal oxygenation that can occur during labor.

Data mining provides various classification techniques with a suitable accuracy rate and time for such medical data to make decisions or discover patterns in datasets.

In the last decades, the researcher provided many papers on data mining techniques supporting bioinformatics to classify and process data with efficient performance. Dr C Sunder [1] used supervised artificial neural network and support vector machine to classify the CTG dataset depending on training data. But the model didn’t have a good performance to classify a suspicious state as the other two states normal and pathological. Dr Ahmed Abou El-Fetouh [8] used hybrid rough neural network model to analyze the performance of breast cancer classification using different sizes of training data. The paper used WEKA [9, 10] tool to measure accuracy rate of Neural Networks and compare the results. But it didn’t estimate the consumption time of the RNN [8, 11] model and didn’t use more algorithms to compare with the proposed model. Dr Suman [3] used WEKA [9, 10] mining tool to analyze classification techniques like (neural network, Bayesian classification and decision tree) to provide which technique has the best and efficient performance. Comparison among different algorithms determines that each algorithm performs the best result according to its parameters, but he did not determine which one was the best in general to use as classification technique [9], Dr Divya Bhatnagar [10]. Provided analyses of CTG data set and generated classification rules to identify normal, suspicious and pathological cases using WEKA classifiers. He didn’t apply simulation for his results or provide hybrid models for improving classification accuracy rate. Z. Cömert [12] presented the comparative metrics of five machine learning techniques such as Artificial Neural Network (ANN) [1, 2, and 13], support vector machine [14], extreme learning machine [15], radial basis function network [16] and random forest [17]. He found that ANN technique is the most efficient in the sensitivity and specificity measures. Dr Mona Gamal[18] used hybrid model of fuzzy rough feature selection and rough neural networks to classify dataset of breast cancer and measure accuracy rate and consumed time of processing data.

The importance of the proposed model is to present a solution of limitations in the previous researches. Providing a good performance to classify all states of fetus heart rate. Also, comparing its results with various algorithms to prove that it satisfies a good accuracy rate in suitable time consumption. And providing analysis of CTG attributes using a WEKA application to visualize it.

The proposed model depends on Rough Neural Network (RNN) [8, 11] which is built on a neural network structure [1, 2 and 13] and rough sets theory [8, 11]. RNN is characterized by various advantages such as the ability to deal with fault tolerance, simplicity and relief of structure, parallel processing of dataset and self-adapted. In addition, RNN advantages of rough set in performing sustainable amount of uncertain data and reduction attributes without losing information. RNN is composed of multilayers input, hidden and output layers. The simulated model measures accuracy rate and time consumption on (CTG) dataset.

The paper is organized as follows; Section 2 presents information about CTG device and its role in diagnosing fetal status. Section 3 provides the proposed model, algorithm and its benefits. Section 4 presents an experimental result of the model and analysis of other classification techniques in comparison. At the end, conclusion and future work are documented in section 5.

II. Cardiotocography

Cardiotocography [2] is common medical devices; many re-searches analyze datasets to achieve improved accuracy in diagnosing the state of fetal heart rate under uncertain situations. The device produces a simultaneous recording and traces patterns of the FHR and the UC during pregnancy period and before delivery. Now the Cardiotocography readings are organized and stored for medical researches.

The CTG dataset consists of measurements of FHR and UC for the fetus, the important features of Cardiotocograms classified by an obstetricians' expert, and the data set is available publicly at the data mining repository of University of California Irvine (UCI) [4]. (Last accessed April 2019). Data set was split into training data and testing data with percentages 70% and 30% respectively.

The data set has 21 attributes and classified according to the FHR pattern or fetal state class code [3, 4]. In this study, fetal state class code is used as the target attribute instead of FHR pattern class code and classification into one of three groups Normal, Suspicious or Pathological (NSP) classes. The dataset includes a total of 2126 samples. Attributes description is given in Table I.

TABLE I. CTG Data Set Attributes Description

CTG data set attributes description	
Attribute	**Description**
LB	Fetal Heart Rate baseline (beats per minute)
AC	number of accelerations per second
FM	number of fetal movements per second
UC	number of uterine contractions per second
DL	number of light decelerations per second
DS	number of severe decelerations per second
DP	number of prolonged decelerations per second
ASTV	percentage of time with abnormal short term variability
MSTV	mean value of short term variability
ALTV	percentage of time with abnormal long term variability
MLTV	mean value of long term variability
Width	width of FHR histogram
Min	minimum of FHR histogram
Max	maximum of FHR histogram
Nmax	number of histogram peaks
Nzeros	number of histogram zeros
Mode	histogram mode
Mean	histogram mean
Median	histogram median
Variance	histogram variance
Tendency	histogram tendency
CLASS	FHR pattern class code (1 to 10)
NSP	fetal state class code (N = normal; S = suspicious ; P = pathologic)

III. Proposed Model

The proposed model uses RNN [8, 11], which depends on combining Neural Network (NN) [1, 2, and 13] and rough set theory [8, 11]. The proposed model applies the supervised learning model of the RNN and formed from three consecutive phases which are preprocessing, training and testing phases as in the following:

1) Preprocessing phase: where medical dataset is normalized to avoid anomaly values of features and improve the efficiency of medical data in implementation stage.

2) Training phase: where the RNN is trained to reach best weights helps in discovering patterns of data and reduce absolute error by using a feed forward algorithm, and back propagation algorithm to update upper and lower weights to reach a better classification of CTG data set.

3) Testing phase: where the trained RNN is measured against new instances of data to calculate the accuracy rates using the relation: Accuracy Rate = 1 – absolute error. Moreover, the time consumption is determined to prove the performance of the proposed model.

The RNN structure replaces the traditional neuron by two neurons (lower neuron, upper neuron) to represent lower and upper approximations of each attribute in the CTG data set, its structure formed from 4 layers input, 2 hidden and output layers. The hidden layers have rough neurons, which overlap and exchange information between each other, While the input and output layers consists of traditional neurons as in Fig. 1.

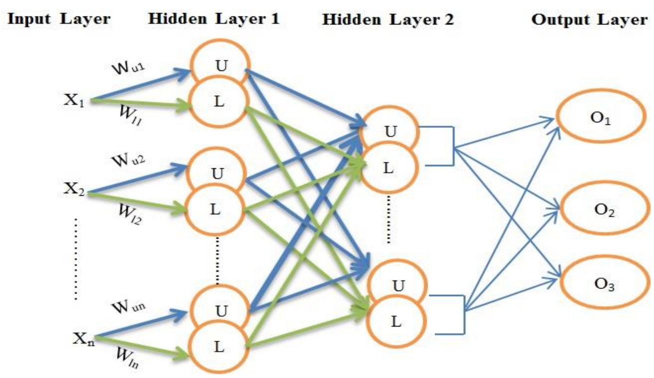

Fig. 1. Rough Neural Network (RNN) Structure.

Input layer is composed of neuron for each data attribute. The output layer represents the three FHR classes, the hidden layers rough neurons are determined by the Baum-Haussler rule [19].

$$N_{hn} = \frac{N_{ts} * Te}{N_i + M_o} \quad (1)$$

Where N_{hn} is the number of hidden neurons, N_{ts} is the number of training samples, Te is the tolerance error, N_I is the number of inputs (attributes or features), and N_O is the number of the output.

During training process, the normalized input data is multiplied by its weight and computed in sigmoid activation function.

$$f(x) = \frac{1}{1+e^{-\lambda x}} \quad (2)$$

Steps of the proposed model architecture:

Step I: preprocessing phase

1. Read features of each objects in dataset
2. Normalize all values of data by equation

$$Nor = \frac{X - \min}{\max - \min} \quad (3)$$

Step II: Training phase

3. Initialize random (upper, lower) weights of network
4. Feed forward of attribute values and multiply in both direction (U_w, L_w)
5. Compute (I_U,I_L) of hidden layers by relations:

$$I_{Ln} = \sum_{j=1}^{n} W_{Lnj} O_{nj} \quad (4)$$

$$I_{Un} = \sum_{j=1}^{n} W_{Unj} O_{nj} \quad (5)$$

6. Compute (O_U , O_L) of hidden layers by relations:

$$O_{Ln} = \text{Min}\,(f\,(I_{Ln}),\ f(I_{Un})) \quad (6)$$

$$O_{Un} = \text{Max}\,(f\,(I_{Ln}),\ f(I_{Un})) \quad (7)$$

7. Check fetus according to comparing between actual output (T) and output value (O), where output represent by

$$O = O_{Ln} + O_{Un} \quad (8)$$

8. If output is error, then use back propagation algorithm, and compute error.

$$\Delta = T - O \quad (8.1)$$

9. Update (upper, lower) weights of network by derivation of activation function: new weight = old weight + (Δ * η *derivative* activation of(input)) (10) where η is learning rate of model
10. Repeat 5, 6, 7, 8 and 8.1 until reduction error as possible as.

Step III: Testing phase

Classify new sample of objects and determine the accuracy rate of the model by using relation Accuracy = 1–absolute error, also calculate time consumption in model processing.

The flowchart of the hybrid proposed model is shown in the following Fig. 2.

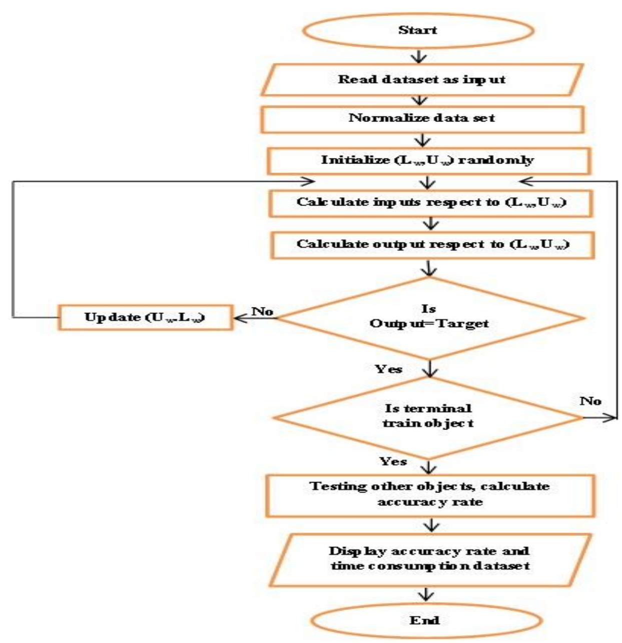

Fig. 2. Flowchart of Rough Neural Network (RNN) Algorithm Steps.

IV. Experimental and Results

CTG is an important medical device that has 21 features to determine the state of fetus heart rate and uterine contraction at the same time. Obstetricians could classify the state of fetal as normal, pathologic or suspicious state according to values of its features. So it's vital to visualize [20, 21] of cardiotcography device features by using WEKA version 3.7.2 [9, 10] application as in Fig. 3. The attribute is drawn to illustrate a visual qualitative understanding of the distribution.

A boxplot is a statistical way to summarize large amounts of data of each feature and display each of minimum, maximum, range, median and distribution of data. Also, it shows the symmetry of data, the upper and lower quartiles, which represent the numbers above and below the high and lower quarters of the data. The CTG data set boxplot is presented in Fig. 4.

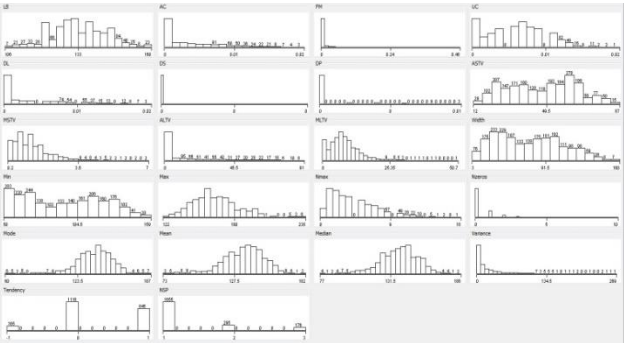

Fig. 3. Distribution of CTG Features.

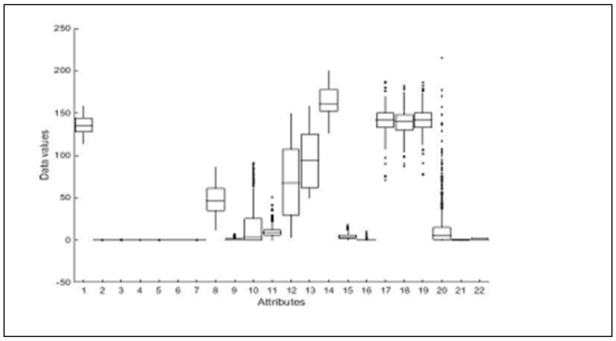

Fig. 4. Boxplot of CTG Features.

The proposed model is implemented in the structure of the rough neural network (RNN) as it's one of the best models in processing uncertain medical data. It is composed of four layers input, two hidden and output layers. The input layer is formed from conventional neuron, while hidden layers are formed from pairs of neurons called upper and lower neurons, which overlap to exchange information and using the interval of values. The output layers formed of conventional neurons represent classes of CTG data set which could be normal, suspicious and pathological states. The proposed model is implemented by C# language. In Windows 7 by specification device, processor Intel ®core™ i5, Ram 4 GB, 64- bit operating. The processing is in three steps which are the preprocessing phase that normalizes features values to avoid anomalies, the training phase, which achieves the RNN learning by using back propagation algorithm to update the weights on networks. The third phase is testing which measures the accuracy rate of the classifier and time consumption during processing the model, The CTG data set is divided into training and testing datasets with a percentage of 70% as training data to learn machine and 30% as testing data to compute accuracy rate of the model.

The WEKA [9,10] tool to analysis CTG dataset using different data mining classifiers such as Nearest Neighbor [22,23], Neural Network[1,2,13], Bagging [24], Decision Table [25,26,27], Decision Stump [28,29] and Least Square Support Vector Machine algorithm [3,30] and compute accuracy rate as in the following Table II.

As shown in the table our model achieves the best accuracy rate compared to other classifiers and it satisfies more efficiency performance. Fig. 5 represents a chart of them as the following.

The time consumption in second of the model is computed and compared to other classifiers as in the following Table III and Fig. 6 observed the RNN model has an acceptable consumption time.

TABLE II. COMPARISON BETWEEN DIFFERENT DATA MINING ALGORITHMS IN ACCURACY RATE

Algorithm	**Accuracy Rate**
Rough Neural Networks	92.95 %
Nearest Neighbor	84.99 %
Neural Network	83.12 %
Bagging	85.15 %
Decision Table	77.85 %
Decision Stump	66.05 %
Support Vector Machine for regression	74.92 %

TABLE III. COMPARISON BETWEEN DIFFERENT DATA MINING ALGORITHMS IN TIME CONSUMPTION

Algorithm	**Accuracy Rate**
Rough Neural Networks	14.25
Nearest Neighbor	0.03
Neural Network	10.01
Bagging	0.39
Decision Table	0.35
Decision Stump	0.04
Support Vector Machine for regression	18.37

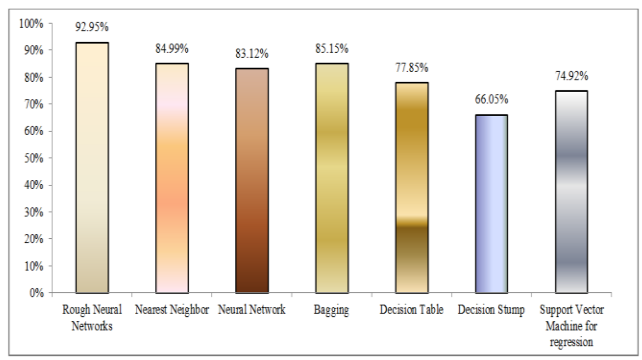

Fig. 5. Comparison between different Data mining Algorithms in Accuracy Rate.

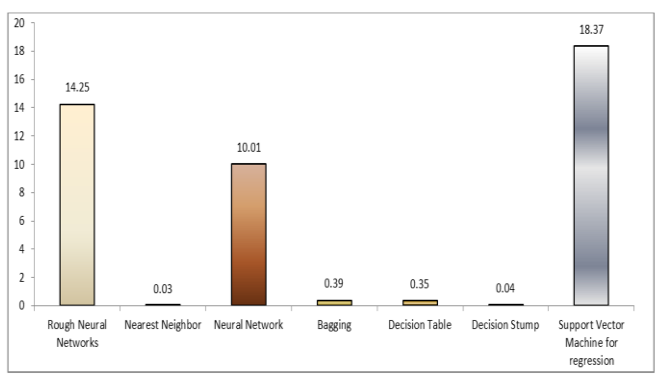

Fig. 6. Comparison between different Data mining Algorithms in Time Consumption.

V. Conclusion and Future Work

Features of fetus Heart Rate (FHR) and Uterine Contraction (UC) during pregnancy are very important in monitoring fetus and mother's health. Data mining provides important techniques for dealing with uncertain medical data. Rough Neural Network classifier is based on neural network and rough set theory. RNN is a well-tested algorithm which satisfies efficiency and provides a good diagnosing of diseases rapidly with high accuracy. RNN structure is composed of rough neurons that manage the upper and lower boundaries in the input and hidden layer instead of traditional neuron with full connection between upper neurons, and lower neurons. During training, RNN learns its weights basing on the back propagation algorithm to updates upper and lower weighs boundaries of input and hidden layers. Through the testing phase, the system measures the accuracy of data and time consumption of the processing model.

The paper presents a distributed and boxplot visualization of CTG features by WEKA tool. Also, the paper provides an implementation of the proposed model, computes the accuracy rate of CTG data set based on absolute error and time consumption. Comparisons between the proposed model and different data mining algorithms such as Nearest Neighbor, Neural Network, Bagging, Decision Table, Decision Stump and Least Square Support Vector Machine algorithm proved the feasibility of the RNN in classifying the CTG data basing on accuracy rate in suitable time.

The future work, several improvements should be made on the accuracy rate of the proposed model technique and apply other data mining techniques. Also, feature selection algorithms would be applied to remove irrelevant features.

References

[1] C. Sunder, M. Chitradevi, and G. Geetharamani, "Incapable of identifying suspicious records in CTG data using ANN based machine learning techniques", Journal of Scientific& Industrial Research,vol 73, , Augest 2014, pp. 510-516.

[2] C. Sunder, M. Chitradevi, and G. Geetharamani," Classification of Cardiotocogram Data using Neural Network based Machine Learning Technique", International Journal of Computer Applications, vol 47, June 2012.

[3] Jagannathan D., "Cardiotocography, "A Comparative Study between Support Vector Machine and Decision Tree Algorithms", International Journal of Trend in Research and Development, vol 4, February 2017.

[4] http://archive.ics.uci.edu/ml/datasets/cardiotocography.(Accessed February 2019).

[5] https://www.sciencedirect.com/topics/medicine-and-entistry/ cardiotocography (Accessed February 2019).

[6] https://en.wikipedia.org/wiki/Cardiotocography.(accessed February 2019).

[7] R. J. Pardeshi, D. V. Gopalghare, M. Sase, and P. P. Panigrahi, "Carditocograghy In Early Labour- A Screning Test For Prediction Of Fetal Outcome", International Journal of Innovative Medicine and Health Science, vol 4, pp. 81-85,2015.

[8] A. A. El-Fetouh, M. Gamal, "An Intelligent Model in Bioinformatics based on Rough-Neural Computing", International Journal of Computer Applications, vol 64, Febuary 2013.

[9] Suman, and P. Mittal, "A Comparative Performance Analysis of Classification Algorithms Using Weka Tool Of Data Mining Techniques", International Journal of Computer Science and Information Technologies, Vol. 5, 2014.

[10] D. Bhatnagar., P. Maheshwari., "Classification of Cardiotocography Data with WEKA", International Journal of Computer Science and Network, vol 5, April 2016.

[11] S. Ding, J.Chen, X Xu, and J. Li," Rough Neural Networks: A Review", Journal of Computational Information Systems,vol 7, 2011.

[12] Z. Cömert, and A.F. Kocamaz, "Comparison of Machine Learning Techniques for Fetal Heart Rate Classification", 3rd International Conference on Computational and Experimental Science and Engineering (ICCESEN 2016), vol 132, 2017.

[13] David K, "A Brief Introduction to Neural Network", http://www.dkriesel.com/en/science/neural_networks, 2005.

[14] W. S. Noble, "What is a support vector machine?", Nature Biotechnology, vol 24 , Dec 2006.

[15] S. Ding, X. Xu, and R. Nie, "Extreme learning machine and its Applications", Journal of Neural Computing and Applications, Vol 25, Pages 549-556,Sep 2014.

[16] Q. Que, and M. Belkin," Back to the Future: Radial Basis Function Networks Revisited", 19th International Conference on Artificial Intelligence and Statistics (AISTATS), vol 51, 2016.

[17] G. Biau, "Analysis of a Random Forests Model", Journal of Machine Learning Research, vol 13, pp 1063-1095, 2012.

[18] M. Gamal, "Medical Diagnostic System Basing Fuzzy Rough Neural-Computing for Breast Cancer", International Conference on Advanced Intelligent Systems and Informatics, Springer, pp. 467-479, 2016.

[19] E.Baum, D.Huassler," what size net gives valid generalization", the Neural Information Processing Systems Conference., 1988.

[20] R. Dawson," How Significant Is A Boxplot Outlier?", Journal of Statistics Education, vol 19, 2011.

[21] H. Wickham and L. Stryjewski, "40 years of boxplots", Technical report, http://vita.had.co.nz/papers/boxplots.pdf, 2011.

[22] Y. N. Law, and C. Zaniolo., "An Adaptive Nearest Neighbor Classification Algorithm for Data Streams", Springer-Verlag Berlin Heidelberg, pp. 108-120, 2005.

[23] S. B. Imandoust, and M. Bolandraftar," Application of K-Nearest Neighbor (KNN) Approach for Predicting Economic Events: Theoretical Background", International .Journal of Engineering Research and Applications, vol 3, Sep-Oct 2013, pp 605-610.

[24] P. Shrivastava, and M. Shukla," Uses the Bagging Algorithm of Classification Method with WEKA Tool for Prediction Technique", International Journal of Advanced Computational Engineering and Networking, vol 2, December 2014.

[25] H. Lu, and H. Liu," Decision Tables: Scalable Classification Exploring RDBMS Capabilities", Proceedings of the 26th International Conference on Very Large Databases, Cairo, Egypt, 2000.

[26] J. Kaur, and Seema Baghla, "Modified Decision Table Classifier by Using Decision Support and Confidence in Online Shopping Dataset", International Journal of Computer Engineering & Technology (IJCET), vol 8, 2017, pp. 83–88.

[27] P. Sewaiwar, and K. K. Verma., "Comparative Study of Various Decision Tree Classification Algorithm Using WEKA", International Journal of Emerging Research in Management &Technology, Oct 2015.

[28] Ayinde A.Q, Adetunji A.B, Bello M and Odeniyi O.A., "Performance Evaluation of Naive Bayes and Decision Stump Algorithms in Mining Students' Educational Data", IJCSI International Journal of Computer Science Issues, vol 10, July 2013.

[29] T. Kannapiran, and M. Wadhawa," Analysis and Comparison Study of Data Mining Algorithms Using RAPID Miner", International Journal of Computer Science, Engineering and Applications (IJCSEA), vol 6, Feb 2016.

[30] E. YJlmaz, and Ç. KJlJkçJer," Determination of Fetal State from Cardiotocogram Using LS-SVM with Particle Swarm Optimization and Binary Decision Tree", Computational and Mathematical Methods in Medicine, 2013.