ISSN: 0974-8571 Vol.13 No. 2 December , 2021

International Journal of Computational Intelligence in Control

Optimized Machine Learning Classifier for Early Prediction of Fetal Abnormalities

R.Chinnaiyan[1], Dr.Stalin Alex[2]

[4]Research Scholar , Department of Computer Science and Engineering, Jain Deemed to be University, Bangalore, India
Professor & Head , Department MCA , AMC Engineering College , Bangalore, Karnataka, India
vijayachinns@gmail.com
[5] Associate Professor, Department of CST (DS) , Jain Deemed to be University , Bangalore, Karnataka, India
drstalinalex@gmail.com

Date of Submission: 21st October 2021 Revised: 11th November 2021 Accepted: 9th December 2021

How to Cite: ***R.Chinnaiyan, Dr.Stalin Alex (2021). Optimized Machine Learning Classifier for Early Prediction of Fetal Abnormalities . International Journal of Computational Intelligence in Control 13(2)***

Abstract

This research paper investigates the machine learning classifiers Random Forest, Naïve Bayes and Support Vector Machine for the better analysis and early prediction of fetal abnormalities with the datasets of cardiotocographs. Cardiotocography is an important process in pregnancy such as monitoring the baby. It looks at whether a child's heartbeat is healthy or not. This can also determine whether a baby's movement in the womb is normal or not. This research work uses 21 abdominal details from CTG datasets imported from Machine Learning Repository. The proposed models will be used as a guide for early analysis and prediction the condition of the fetus whether it is under normal, suspicious or pathological conditions

.

Keywords: Carditocography, Fetal, Abnormalities, Machine Learning, Prediction

INTRODUCTION

Machine learning approaches uses the capability of computers for training and learning without being explicitly programmed. Machine learning approaches supports the doctors in making precise medical diagnoses, analysis and prediction which helps in rational decision making process. The existing machine learning approaches are dedicated on segmenting fetal brain pictures to identify abnormalities. The fetal abnormalities in a pregnant women appears due to deviation in the normal growth of fetus. Fetal abnormalities are deliberated irreparable and the reasons due to abnormalities is unnoticed. The classification of fetal abnormalities in prenatal women is triggered due to many factors such us ., Primary abnormalities , Secondary abnormalities , Deformationis , Dysplasiais , Agenesia , Sequence: and Syndromeis . Various fetal abnormalities are depicted in the following figure 1.

The main objective this proposed research is to introduce optimized ensemble based machine learning classification approaches for early analysis and prediction of fetal abnormalities.

2. LITERATURE REVIEW

Yu Lu., et.al.,(2020), Ensemble Approach An ensemble machine learning model has been proposed based on the genetic algorithm with parallel optimization of multiple parameters to predict the fetal weight at varying gestational age with an accuracy of 64.3%.

The limitation of their work is Estimation of fetal birth weight among twins. Md Rafiul Hassan et.al.,(2020)., Hill Climbing Approach., An automated tool for predicting IVF pregnancy success based on emerging machine learning classifiers, namely MLP, SVM, C4.5, CART and random forest. With an accuracy of 97.88%, 98.38%, 94.28%, 95.42%, 98.25%.

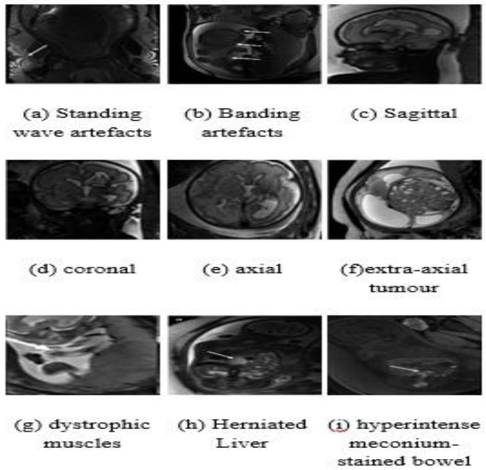

Figure 1. Diverse fetal abnormalities

Future study will consider multi collinearity among features to further improve feature selection strategy and thereby pregnancy prediction outcome. H. N. Xie et.al.,(2020)., Deep-learning algorithms, Deep-learning algorithms can be trained for segmentation and classification of normal and abnormal fetal brain ultrasound images with an accuracy 96.2 % .

Further research on the differential diagnosis of fetal intracranial abnormalities. J.Jayashre et.al.,(2020) ., MRMR feature selection algorithms., Paper presents MRMR feature selection algorithms with four classification for Fetal risk prediction using python with an accuracy of 98%. The predictive potential of these approaches remains controversial and still unreliable.

3. PROPOSED METHODOLOGY

Steps

Step 1: *Begin*

Step 2: *Import the Fetal CTG dataset from Machine Learning Repository*

Step 3: *Perform Data Pre-Processing*

Step 4: *Generate Feature Subset FS*

Step 5: *Prepare test and train data*

Step 6: *Build the model RF, NB ,SVM*

Step 7: *Test and Train the Models using RF, NB and SVM*

Step 8: *Evaluate the performance of all the models with test data*

Step 9: *Model Validation*

Step 10: *Predict the performanace of ML Classifiers RF , NB and SVM*

Step 11: *End*

4. IMPLEMENTATION

This proposed research work compares the machine learning classifiers random forest, the Naïve Bayes with SVM of machine learning. The confusion matrix with RMSE is estimated for each method and are depicted in the figures.

Table 1: Attributes of Fetal Health Datasets (AFHD)

Attribute No	Attribute Name	Attribute No.	Attribute Name
AFHD1	LB-FHR	AFHD 13	Min–minimum of FHR histogram
AFHD 2	AC-# of accelerations per second	AFHD 14	Max–Maximum of FHR histogram
AFHD 3	FM-# of fetal movements per second	AFHD 15	Nmax-# of histogram peaks
AFHD 4	UC-# of uterine contractions per second	AFHD 16	Nzeros-# of histogram zeros
AFHD 5	DL-# of light decelerations per second	AFHD 17	Mode–histogram mode
AFHD 6	DS-# of severe decelerations per second	AFHD 18	Mean–histogram mean
AFHD 7	DP-# of prolonged decelerations per second	AFHD 19	Median–histogram median
AFHD 8	ASTV–percentage of time with abnormal short term variability	AFHD 20	Variance–histogram variance
AFHD 9	MSTV–mean value of short term variability	AFHD 21	.Tendency–histogram tendency
AFHD 10	ALTV–percentage of time with abnormal long term variability	AFHD 22	CLASS–FHR pattern class code (1 to 10)
AFHD 11	MLTV–mean value of long term variability	AFHD 23	NSP–fetal state class code (N = Normal; S = Suspect; P = Pathologic)
AFHD 12	Width–width of FHR histogram		

Table 2: Fetal State Class Code

S.No	State	Class Code	Total	
1	N – Normal	1	1655	
2	S- Suspect	2	295	
3	P- Pathologic	3	176	
Total			2126	

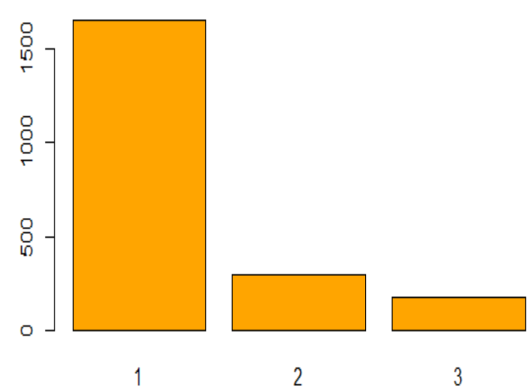

Figure 2. Fetal State

Table 3: Summary of Fetal Health Datasets

S. No	Attribute Number	Min	Max	Mean	Median
1	AFHD1	106.0	160.0	133.3	133.0
2	AFHD 2	0.000	0.019	0.003	0.002
3	AFHD 3	0.000	0.481	0.009	0.000
4	AFHD 4	0.000	0.015	0.004	0.004
5	AFHD 5	0.000	0.015	0.002	0.000
6	AFHD 6	0.000	1.000	3.293	0.000
7	AFHD 7	0.000	0.005	0.000	0.000
8	AFHD 8	12.00	87.00	46.00	49.00
9	AFHD 9	0.200	7.000	1.333	1.200
10	AFHD 10	0.000	91.00	9.847	0.000
11	AFHD 11	0.000	50.70	8.188	7.400
12	AFHD 12	3.000	180.0	70.45	67.50
13	AFHD 13	50.00	159.0	93.58	93.00
14	AFHD 14	122.0	238.0	162.0	164.0
15	AFHD 15	0.000	18.00	4.068	3.000
16	AFHD 16	0.000	10.00	0.324	0.000
17	AFHD 17	60.00	187.0	137.0	139.0
18	AFHD 18	73.00	182.0	134.0	136.0
19	AFHD 19	77.00	186.0	138.0	139.0
20	AFHD 20	0.000	259.0	18.81	7.000
21	AFHD 21	-1.000	1.000	0.320	0.000

5. RESULTS

The effectiveness of these three methods is shown in Table 2. According to the accuracy of the RMSE, the random forest was significantly more efficient with 99.93 % of accuracy compared to the other two methods NB with 83.78% of accuracy and SVM with 93.37% of accuracy. Results shows that the Random Forest provides the better prediction when compared to NB and SVM Models.

On the other hand, the inexperienced Naïve Bayes had only 83.78% accuracy. Outcome analysis based on the matrix of confusion is shown in Tables 3, 4, and 5.

5.1 Accuracy and RMSE

Table 4.

Classifier	Accuracy	*Root Mean Square Error*
RF	99.93	0.1506
Naïve Bayes	83.78	0.33
SVM	93.37	0.21

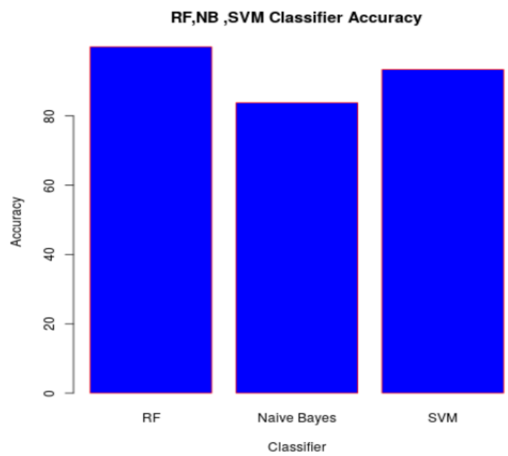

Figure 3. Accuracy

5.2 Confusion Matrix for Random Forest

Table 5.Confusion Matrix of RF

	Normal	Suspect	Pathologic
Normal	1177	17	1
Suspect	49	148	4
Pathologic	3	5	6

5.3 Confusion Matrix for Naïve Bayes

Table 6. Confusion Matrix of NB

	Normal	Suspect	Pathologic
Normal	1041	38	5
Suspect	110	151	38
Pathologic	44	12	84

5.4 Confusion Matrix for SVM

Table 7. Confusion Matrix of SVM

	Normal	Suspect	Pathologic
Normal	1170	50	4
Suspect	24	151	22
Pathologic	3	1	0

5.5 Performance Measure of RF, NB and SVM

The performance is estimated with the below mentioned equations. .

$$Accuracy = \frac{\mathrm{TP} + \mathrm{TN}}{TP + TN + FP + FN}$$

$$Precision = \frac{\mathrm{TP}}{TP + FP}$$

$$Recall = \frac{\mathrm{TP}}{TP + FN}$$

$$F1 - score = \frac{2.Precision.Recall}{Precision + Recall}$$

Table 8. Sensitivity , Specificity and Accuracy

Classifier	Sensitivity	Specificity	Accuracy
Random Forest	98.70	99.70	99.93
Naïve Bayes	88.48	91.61	83.78
SVM	97.91	83.54	93.37

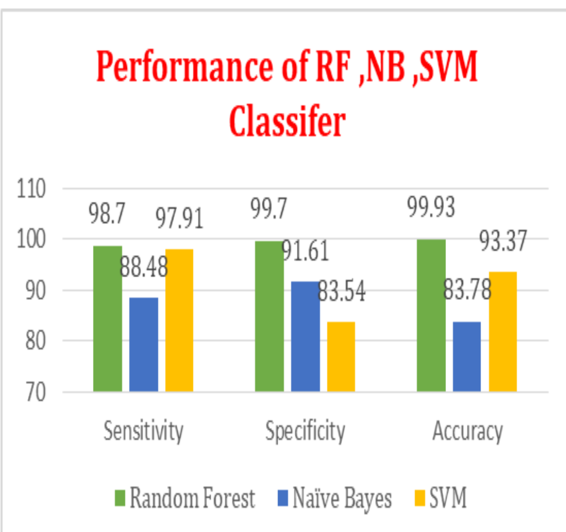

Figure 4. Performance of RF, NB and SVM Classifier

6. CONCLUSION

This research work compared the machine learning classifiers Random Forest, Naïve Bayes and Support Vector Machine in classifying embryonic conditions based on CTG Data. Tests based on ten-fold verification. Performance was measured using RMSE accuracy. Table 2 shows that the random forest was significantly more efficient compared to the other two methods

References

1. Abdulhamit Subasia, Bayader Kadasaa, Emir Kremic, “Classification of the Cardiotocogram Data for Anticipation of Fetal Risks using Bagging Ensemble Classifier”, Procedia Computer Science 168 (2020) 34–39
2. Alessio Petrozziello, Ivan Jordanov, Aris T.Papageorghiou, Christopher W.G. Redman, and Antoniya Georgieva,” Deep Learning for ContinuousElectronic Fetal Monitoring in Labor”, Preprint, Researchgate
3. Attallah O, Sharkas MA, Gadelkarim H. Fetal Brain Abnormality Classification from MRI Images of Different Gestational Age. *Brain Sciences*. 2019; 9(9):231.
4. Balachandar S., Chinnaiyan R. (2019) Centralized Reliability and Security Management of Data in Internet of Things (IoT) with Rule Builder. In: Smys S., Bestak R., Chen JZ., Kotuliak I. (eds) International Conference on Computer Networks and Communication Technologies. Lecture Notes on Data Engineering and Communications Technologies, vol 15. Springer, Singapore
5. Balachandar S., Chinnaiyan R. (2019) Reliable Digital Twin for Connected Footballer. In: Smys S., Bestak R., Chen JZ., Kotuliak I. (eds) International Conference on Computer Networks and Communication Technologies. Lecture Notes on Data Engineering and Communications Technologies, vol 15. Springer, Singapore
6. Comert Z., Kocamaz A. F., Subha V. (2018). Prognostic model based on image-based time-frequency features and genetic algorithm for fetal hypoxia assessment. *Comput. Biol. Med*. 99 85–97.
7. Daniel LaFreniere, Farhana Zulkernine, David Barber, Ken Martin. “Using Machine Learning to Predict Hypertension
8. R.Vani, “Weighted Deep Neural Network BasedClinical Decision Support System for the Determination of Fetal Health”, International Journal of Recent Technology and Engineering (IJRTE)ISSN: 2277-3878, Volume-8 Issue-4, November 2019,8564-8569.
9. Ragab DA, Sharkas M, Attallah O. Breast Cancer Diagnosis Using an Efficient CAD System Based on Multiple Classifiers. *Diagnostics*. 2019; 9(4):165.
10. S.Balachandar , R.Chinnaiyan (2018), Centralized Reliability and Security Management of Data in Internet of Things (IoT) with Rule Builder, Lecture Notes on Data Engineering and Communications Technologies 15, 193-201.
11. S.Balachandar , R.Chinnaiyan (2018), Reliable Digital Twin for Connected Footballer, Lecture Notes on Data

Engineering and Communications Technologies 15, 185-191.

12. S.Balachandar , R.Chinnaiyan (2018), A Reliable Troubleshooting Model for IoT Devices with Sensors and Voice Based Chatbot Application, International Journal for Research in Applied Science & Engineering Technology,Vol.6,Iss.2, 1406-1409.
13. M. Swarnamugi ; R. Chinnaiyan, "IoT Hybrid Computing Model for Intelligent Transportation System (ITS)", IEEE Second International Conference on Computing Methodologies and Communication (ICCMC), 15-16 Feb. 2018.
14. M. Swarnamugi; R. Chinnaiyan, "Cloud and Fog Computing Models for Internet of Things", International Journal for Research in Applied Science & Engineering Technology, December 2017.
15. G Sabarmathi, R Chinnaiyan (2019), Envisagation and Analysis of Mosquito Borne Fevers: A Health Monitoring System by Envisagative Computing Using Big Data Analytics, Lecture Notes on Data Engineering and Communications Technologies book series (LNDECT, volume 31), 630-636. Springer, Cham
16. S. Balachandar, R. Chinnaiyan (2019), Internet of Things Based Reliable Real-Time Disease Monitoring of Poultry Farming Imagery Analytics, Lecture Notes on Data Engineering and Communications Technologies book series (LNDECT, volume 31), 615- 620. Springer, Cham
17. M Swarnamugi, R Chinnaiyan (2019), IoT Hybrid Computing Model for Intelligent Transportation System (ITS), Proceedings of the Second International Conference on Computing Methodologies and Communication (ICCMC 2018), 802-806.
18. G. Sabarmathi, R. Chinnaiyan (2016) , Big Data Analytics Research Opportunities and Challenges - A Review, International Journal of Advanced Research in Computer Science and Software Engineering, Vol.6 , Issue.10, 227-231
19. G. Sabarmathi, R. Chinnaiyan, Investigations on big data features research challenges and applications, IEEE Xplore Digital LibraryInternational Conference on Intelligent Computing and Control Systems (ICICCS), 782 – 786.