PERFORMANCE COMPARISON OF DECISION TREE, RANDOM FOREST, AND XGBOOST MODELS; AND ITS INTERPRETABILITY USING SHAP FOR RECOGNIZING THE NECESSITY OF CAESAREANS SECTION OF CHILDBIRTH

ASWAN SUPRIYADI SUNGE[1], EDI ABDURACHMAN[2], YAYA HERYADI[3], IMAN H. KARTOWISATRO[4]

[1,3]Computer Science Departement, BINUS Graduate Program-Doctor of Computer Science, Bina Nusantara University, Jakarta, Indonesia

[2]Trisakti Institute of Transportation and Logistic, Trisakti University, Jakarta, Indonesia

[4]Computer Science Departement-BINUS Graduate Program, Doctor of Computer Science,

Computer Engineering Departement, Faculty of Engineering

Bina Nusantara University, Jakarta, Indonesia

E-mail: [1]aswan.sunge@binus.ac.id, [2]edia@itltrisakti.ac.id, [3]yaya.heryadi@binus.edu, [4]ihkartowisastro@binus.ac.id

ABSTRACT

Birth is one of the processes that every woman will go through after pregnancy, many risks are faced from birth failure to death of the baby and mother. One of the methods for reducing this risk is by Caesarean section, but before doing this, medical personnel or obstetricians need medical records from the patient before making the decision, from this analysis Machine Learning (ML) is needed to analyze whether Caesar is needed or not for the candidate. mother to give birth. The data used in analyzing the prediction of Caesarean birth with external data which consists of 32 attributes totaling 3602 are then tested with the Decision Tree, Random Forest, and XGBoost models then compared which one has the highest performance of each model with the overall data and selected data then tested with the model SHAP to see which features are highest in Caesar's predictions. The test results using all the features show that Random Forest (0.86) achieves the highest accuracy. However, using selected features such as Age, Duration until the next pregnancy in days, and Obesity has the highest accuracy, namely XGBoost (0.85). Finally, in the SHAP model, the highest feature of the 32 features is ICD10O82 (Contractions but without any indication) and ICD10O80 (Infection of the female reproductive organs). From the results of testing with the ML model and it is also seen that the very domain features are used as a guideline for cesarean prediction decision-makers.

Keywords: *Prediction, Caesar, Classification, Comparison, SHAP.*

1. INTRODUCTION

Childbirth is part of the final cycle of a woman's life after the pregnancy process, moreover, it is a happy moment for expectant mothers and their partners to welcome the birth of a baby [1]. In the process of pregnancy, many things must be prepared because it has an impact on the safety of the mother and baby, this becomes the fear of the mother before giving birth such as miscarriage, baby defects to death [2-5] especially depression after giving birth which has physical and mental consequences for the mother [6]. In addition, the biggest fear is the mother's first birth from insomnia, obesity, and general problems after childbirth [7], whether from a vaginal birth or Caesarean [8-9]. In the delivery process without any health problems for the mother and child through a normal delivery process or the vaginal route, but if there is interference with the safety of the mother or the baby, then a Caesarean section is performed immediately [10]. However, in determining a Caesarean section, an analysis of the patient's history of medical record data and an analysis of predictions determined by experienced health

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

workers and doctors to see whether a Caesarean section is necessary.

Therefore, Machine Learning (ML) is needed to analyze whether a cesarean section is necessary. ML is very suitable for analyzing data, especially predicting making decisions such as marketing, and industry in analyzing big data, especially medicine, which is one of the fields, namely whether a cesarean birth is necessary [11-14]. Moreover, ML is currently widely used, especially in the health sector in classification, regression, clustering, and especially prediction [15-19].

However, the use of ML depends on the selection of an algorithm model, which is expected to predict accuracy in making decisions. Among the algorithms that are widely used is the Decision Tree, because it is easy to implement, and interpret and can be used in various fields, but this model has several drawbacks, namely it cannot apply in regression or clustering, there are many duplicates in subtrees and limited in output per attribute [20], the advantages of the Random Forest model can be used for tens and hundreds of features or the amount of data, especially the low correlation between models, also better performance than decision trees and faster processing [21] but the disadvantages of this model require large enough data and only labeled data or class [22], and finally the XGBoost model which is an ensemble technique based on gradient boosting which is faster and more efficient, especially in computing [23], but the drawback is that it takes quite a long time to make a model if the data is small [24].

From the test results above, another accurate test is also needed because the ML model is considered difficult to understand and explain in interpreting predictions, therefore SHapley Additive exPlanations (SHAP) are used to see the contribution of features that are very dominant or highest in Caesar's predictions and obtain Duration_till_next_pregnancy which is very influential. So, in this study, the aim is to compare the three algorithms to see which is the best in terms of accuracy, precision, recall, and F1 in Caesar's predictions and then see which features are most influential and highest with the SHAP model. It is hoped that it will be able to develop a machine learning model that has not existed before, and its benefits can prevent death for mothers and prospective babies in labor.

2. REVIEW OF RELATED WORK

Several studies on the prediction of Caesarean births or related to medical have been done before. First, Lakshmi et al. [25] predicting Caesarean section in pregnant women having complications during pregnancy with Algorithm C.45, data used in Bangalore district, Karnataka state for this study, and three medical centers from Bangalore district were identified for data collection. From this research, the C.45 algorithm gets very good accuracy and performance. Akbulut et al.[26], Predict when giving birth or undergoing pregnancy by creating a system so that the patient knows what to do with the Averaged Perceptron model, Boosted Decision Tree, Bayes Point Machine, Decision Forest, Decision Jungle, Locally-Deep Support Vector Machine, Logistic Regression, Neural Network, Support Vector Machine and data used from pregnant women received at the radio diagnostic center (Radyo Emar, Bakirkoy, Istanbul). From the test results, it was found that Decision Forest with high accuracy was 89.5% compared to other models. Beksac et al. [27] Making a computerized system is obtained when the patient enters the delivery room at the beginning of the first stage of labor after a pelvic examination and examination/evaluation of the mother and fetus in predicting normal delivery and cesarean section using the Supervised Artificial Neural Network (ANN) model. The data used is from the Department of Obstetrics and Gynecology, Division of Perinatology, Hacettepe University Medical Faculty, Ankara, Turkey. With the ANN test, the specificity and sensitivity of the system are 97.5% and 60.9% respectively, which means that it is very good in testing with this model. Khan et al. [28] Prediction of Caesarean Delivery using the Ensemble Machine Learning Method with the XGBoost, AdaBoost, and CatBoost models with the dataset used by Campillo Artero. The results of this study were XGBoost with accuracy-88.91%, then AdaBoost showed 88.69% accuracy, and Catboost had 87.66% accuracy. Sujatha [29] Seccar's Initial Prediction for Treatment Decision Making, with Bayes, NN, MP Classifier, and data used from external UCI data, the research results obtained from the three models tested, the Naive Bayes model has a prediction accuracy of 87.5%. Lastly, Wang et al. [30] Prediction of heart disease by medical records for the past 3 years using SHAP provides a very good interpretation of the contributing factors to the disease.

3. MATERIAL AND DATA SET

There have been several studies on analyzing and predicting pregnancy and birth in the field of computer science and the health sector that are interrelated, this section discusses several techniques used by previous researchers in using several models, namely the Decision Tree, which is a very reliable, effective, decision-making technique. easy to understand and provides high classification accuracy with simple interpretability is also used in various fields, especially medical, one of which is in birth [31-33], Random Forest, a technique that is widely used in predictions in the medical field because it performs very well and efficiently, especially in bagging models and boosting but in better performance than Neural Networks which require more data [34-36]. XGBoost is an ensemble technique of machine learning that is reliable and efficient in speed in calculating training data and is also more complex than other decision tree algorithms [37-38]. In the SHAP method, it is used as a model, one of which is in the medical field because it looks at the dominant features and sees features that greatly affect other features and considers each specific feature that influences each prediction made [39-42].

2.1 Data Set

The data used in this research is secondary data from https://datadryad.org/stash/dataset/doi:10.5061/dryad.g7t04 with a total of 3602 data with 32 attributes and two classes, namely Caesar (1) or Not Caesar (0).

Table 1: Atributes, Description and Value Data Set.

Attributes	Description	Value
Age	Year of patient	Number
Duration till next pregnancy in days	Next day to next pregnancy	Number
Obesity	Weight	0 or 1
Ruptured membranes	Premature rupture of membranes	0 or 1
Placenta Disorders	Complicated disorders in which part or all the placenta separates from the uterine wall before the baby is born	0 or 1
Condition of the placenta		0 or 1
Premature Separation	A condition in which the placenta separates from the uterine lining	0 or 1
Bleeding		0 or 1
Fake Childbirth	Sporadic uterine contractions	0 or 1
Pregnancy Over 40 Weeks		0 or 1
Premature Labor	Deliveries that occurred before 37 weeks of gestation. Fetuses born usually have less weight	0 or 1
Primary contractions are not strong		0 or 1
Medical diagnosis	Describe the presence of labor abnormalities in the 1st stage	0 or 1
Perineal laceration	Damage to the female genital organs that usually occurs during childbirth	0 or 1
Other obstetric trauma	Fear after Childbirth	0 or 1
Postpartum haemorrhage	Bleeding from the birth canal immediately after delivery	0 or 1
Placenta retained without bleeding	A condition when the placenta or placenta does not come out on its own or is stuck in the uterus after delivery	0 or 1
Complications of anesthesia during labor	Effects after anesthesia	0 or 1
Other complications of labor and delivery	Childbirth complications	0 or 1
Genital and pelvic tract infections after abortion	Infection of the female reproductive organs	0 or 1
Single birth with forceps and vacuum extractor	First childbirth with vacuum extractor)	0 or 1
Meets fault without indication	Mistakes diagnosis	0 or 1
Chronic postrheumatic arthropathy	The disease is chronic and occurs for life and usually affects both men and women	0 or 1
Pre-existing essential hypertension complicating pregnancy, third trimester	Hypertension that is not caused by certain conditions, diseases, or drugs but pregnancy	0 or 1
Posttraumatic urethral stricture, male, unspecified	The tube that carries urine from the bladder out of the body	0 or 1
Anterior dislocation of left humerus	A condition in which the bony hump of the upper arm detaches from the shoulder joint.	0 or 1
Acute suppurative otitis media with spontaneous	Inflammation of the middle ear, is a condition that occurs when a virus or bacteria causes the area behind the eardrum to become inflamed	0 or 1
Eye and vision examination after vision screening failed without abnormal findings	Evaluate your vision condition and check for eye diseases	0 or 1
Goose neck deformity on left finger	A disorder that makes the shape of the fingers look like the neck of a swan	0 or 1
Twin pregnancy	Identical twins who share both a placenta and an amniotic sac	0 or 1
Latex allergy status	The immune system's reaction to the proteins in natural rubber latex, a product made from the liquid of the rubber tree	0 or 1

Journal of Theoretical and Applied Information Technology
15th May 2023. Vol.101. No 9

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

2.2 Data Splitting

After the data is obtained, labeling, and searching for empty data are then divided which are numerical and categorical data based on the data obtained. If it has been divided into two, namely the training dataset which is used to display the results of the model that is appropriate to use while the data testing is to test and see the performance of the model obtained. The data set is divided into 70% training data and 30% testing data.

4. RESEARCH METHOD

The following describes how to get Caesar's prediction model.

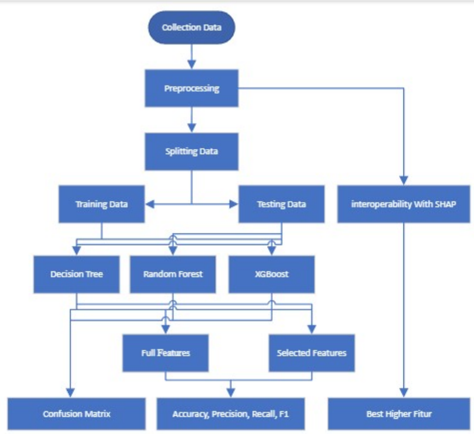

Figure 1: Research Framework

First, the data obtained amounted to 3602 with class Caesar and not Caesar, then the data was cleaned to find missing data and then the data was divided into 2 training data 70% of the total data and testing data 30% of the amount of data to reduce errors and see which models are suitable for use [43]. Second, testing with 3 models, namely Decision Tree, Random Forest, and XGBoost, to see the confusion matrix to measure the performance of the classification model being tested [44]. After that look at the accuracy of the number of truths of each amount of data then see precision as the level of accuracy between the information requested and the answers given by the system then recall as a presentation of the relevant amount of data from the selected item and F1 which is known as F-Measure as the process of taking precision and recall to calculate the performance of an algorithm [45] to be able to compare and measure the performance resulting from each model by using the full data set of all the features used and only using 3 features namely age, duration of pregnancy and obesity, to compare which is best from full data set and selected features. Finally, the entire data set is tested using the SHAP method to see the specific features that are very influential or the highest for each prediction made.

5. RESULT AND DISCUSSION

In this discussion, the differences with previous research are comparisons of models that are often used in the medical field, and then the best confusion matrix is tested based on positive and negative values. A very striking difference in testing this research is using data on all the attributes and selected attributes, then using SHAP by displaying the best feature sequence.

The data used in Caesar prediction is 3602 with 32 attributes then pre-processing the data so that the process is in a form that is easier to understand then the data is divided into 70% as training data and 30% as testing data. After that, it was tested with 3 models of classification, namely Decision Tree, Random Forest, and XGBoost, then the accuracy of each model will be seen based on the overall data and based on certain attributes used. By using the overall data in table 1, the highest confusion matrix is Random Forest, which means that the value is a positive prediction of Caesar, but positive Caesar is 650 and a negative prediction or not Caesar but is negative or not Caesar is 590.

Table 2: Comparison Confusion Matrix

Predicted	Actually		Model
	Positive (1)	Negative (0)	
Positive (1)	601	133	Decision Tree
Negative (0)	112	595	
Positive (1)	650	84	Random Forest
Negative (0)	117	590	
Positive (1)	630	104	XGBoost
Negative (0)	112	595	

With training data using all the attributes in table 1, the highest accuracy and precision are obtained, namely from Random Forest, while recall, namely XGBoost and F1, namely Random Forest and XGBoost, can be seen in table 3. what is displayed is more positive or Caesar predictions than negative or non-Caesar, more over Precision does not depend on Accuracy.

Table 3: Full Features

Model	Accuracy	Precision	Recall	F1
Decision Tree	0.84	0..86	0.80	0.83
Random Forest	0.86	0.88	0.83	0.85
XGBoost	0.85	0.85	0.85	0.85

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

With training data using certain attributes with selected attributes, namely Age, Duration till next pregnancy in days and Obesity, the highest accuracy is obtained, namely XGBoost, while the highest precision is Decision Tree and Random Forest, while Recall and F1 are the highest, namely XGBoost as shown in table 4. If it is seen that Recall is higher with Accuracy, Precision and F1 because it prefers to choose False Positives which are better than False Negatives, for example there are patients who are predicted not to be Caesar, but the condition is Caesar.

Table 4: Using Selected Features

Model	Accuracy	Precision	Recall	F1
Decision Tree	0.84	0.78	0.91	0.84
Random Forest	0.84	0.78	0.91	0.82
XGBoost	0.85	0.74	0.95	0.86

After testing by testing the model above, then interpreting the ML prediction model correctly and presenting the predicted results visually, which is expected to be understood by health workers or obstetricians. Therefore, apply the SHAP model to XGBoost to achieve the best prediction and interpretation results. The SHAP value evaluates the importance of an output containing all features and provides a locally consistent and accurate attribute value for each feature in the prediction model. With data training using certain attributes with the SHAP model, it is shown below.

Figure 2: The interpretation of model prediction results with the two samples. (The values of each variable are normalized values)

From Figure 2 is an illustration selected in Caesar's prediction based on the SHAP model. The arrow direction means the influence of each factor on the prediction. So, it is predicted to have a value of 1.66 but the base value is -0.02656, which means that the feature value has increased or has an effect in the form of pink visual color, but the feature value has no increase or effect in blue. The most influential feature is ICD10O80 but the ICD10O82 feature has a significant effect, which means a de-crease in prediction.

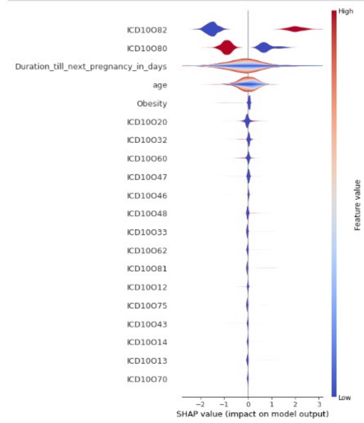

Figure 2: The importance ranking of the top 20 risk factors with stability and interpretation using the optimal model.

The figure above shows the 20 features in Caesar's prediction in the model. The feature rating (Y-Axis) indicates the importance of the predictive model results. The SHAP value (X-Axis) is a unified index that represents the influence of certain fea-tures in the model. In each row of salient features, the attribution of all patients or expectant mothers to outcomes is drawn with different color dots, where the red dot represents a high-risk score, and the blue dot represents a low-risk score. Looking at the features above, ICD10O82 has a high relationship with a positive value but not so much weight, however, the largest value is the negative value and the ICD10O80 feature has a positive value but not too high a level of influence on the predicted results. Furthermore, Duration_till_next_pregnancy_in_days and age have balanced positive and negative values.

From the results of the three models used from the Confusion Matrix, Accuracy, Precision, Recall, F1, and SHAP, it can be used as an option in Caesar's predictions. therefore, the role of the health worker or doctor can be used as the decision-making Caesar.

6. CONCLUSIONS

The results of all tests, from the Confusion Matrix, it is obtained from the three algorithm models, namely Random Forest, then based on the overall data test the highest Accuracy and Precision, namely Random Forest and Recall and FI, namely XGBoost, but if seen from the data selected from certain attributes, namely Accuracy, Recall, and F1 namely XGBoost. Finally, in the test to see the interpretation with a very do-main SHAP, namely Duration_till_next_pregnancy and the second age has a bal-ance value of each feature.

However, in this study, there are still deficiencies, namely the data used today may be different from the data used in the future with different and more diverse attributes, which in turn results in predictions, and the dominant feature will differ from each feature of the data used. It is also hoped that the data used can be used as a clustering model and using other algorithms such as K-NN, SVM, CNN, RNN, or other algorithms. For Interpretable, maybe use another model as a comparison and see how far is the easiest to understand in the display. It is hoped that it can develop and design a system or application in Caesarean predictions using a programming language and can be combined with an integrated IoT system so that it can help in the medical world for the safety of mothers and prospective babies.

REFERENCES:

[1] M. Berg and K. Dahlberg, "A phenomenological study of women's experiences of compli-cated childbirth," Midwifery. Vol. 14, Issue 1, March 1998, Pages 23-29.

[2] M. Aksoy, A. N. Aksoy, A. Dostbil, M. G. Celik, and I. Ince, "The Relationship between Fear of Childbirth and Women's Knowledge about Painless Childbirth," Hindawi Publish-ing Corporation Obstetrics and Gynecology International Volume 2014, Article ID 274303, 7 pages.

[3] T. H. T. Lai, S. T. Kwok, W. Wang, M. T. Y. Seto, and K. W. Cheung, "Fear of childbirth: Validation study of the Chinese version of Wijma delivery expectancy experience question-naire version B", Midwifery, Vol. 108, May 2022, 103296.

[4] P. Khosrav, S. P. Beiranvand, B. Beiranvand and Z. B. Khalesi, "Relationship between Primigravid women's awareness, attitude, fear of childbirth, and mode of delivery prefer-ence," European Journal of Obstetrics & Gynecology and Reproductive Biology: X Volume 14, April 2022, 100143.

[5] B. Areskog, B. Kjessler and N. Uddenberg, "Identification of women with significant fear of childbirth during late pregnancy," Gynecol Obstet Inves. 1982; 13:98–107.

[6] J. Söderquist, B. Wijma, G. Thorbert and K. Wijma, "Risk factors in pregnancy for post-traumatic stress and depression after childbirth," BJOG: An International Journal of Obstet-rics & Gynaecology, 116(5), 672–680.

[7] L. Nilsson, T. Thorsell, E. H. Wahn, and A. Ekstrom, "Factors Influencing Positive Birth Experiences of First-Time Mothers," Hindawi Publishing Corporation Nursing Research and Practice Vol. 2013, Article ID 349124, 6 pages.

[8] S. Räisänen, S. M. Lehto, H. S. Nielsen, M. Gissler, M. R. Kramer and S. Heinonen, "Fear of childbirth predicts postpartum depression: a population-based analysis of 511 422 single-ton births in Finland," BMJ: British Medical Journal on JSTOR. 2013;3: e004047.

[9] L. D. Giglio, S. Federici, S. Ruggieri, G. Borriello, MA. D'Errico, C. D. Angelis and C. Pozzilli, "Cesarean section in women with MS: A choice or a need? Multiple Sclerosis and Related Disorders," Vol. 38, February 2020, 101867.

[10] A. M. Darcy, A. K. Louie, and L. W. Roberts, "Machine Learning and the Profession of Medicine," Journal of the American Medical Association. February 2016. 315(6):551.

[11] B. Nistal-Nuño, "Developing machine learning models for prediction of mortality in the medical intensive care unit," Computer Methods and Programs in Biomedicine Vol. 216, April 2022, 106663.

[12] H. Erik, "Editorial Commentary: Big Data and Machine Learning in Medicine. Arthroscopy: The Journal of Arthroscopic & Related Surgery," Vol. 38, Issue 3, March 2022, Pages 848-849.

[13] V. K. Verma and S. Verma, "Machine learning applications in healthcare sector: An over-view," Materialtoday; Proceedings. Vol 57, Part 5, 2022, Pages 2144-2147.

[14] A. Alanazi, "Using machine learning for healthcare challenges and opportunities," Informat-ics in Medicine Unlocked. Vol. 30, 2022, 100924.

[15] B. Ozaydin, E. S, Berner and J. J. Cimino, "Appropriate use of machine learning in healthcare". Intelligence-Based Medicine. Vol/ 5, 2021, 100041.

[16] H. Pallathadka, M. Mustafa, D. T. Sanchez, G. S. Sajja, S. Gour and M. Naved, "Impact of Machine Learning on Management, Healthcare and Agriculture," Materialtoday; Proceed-ings. Available online 22 July 2021.

[17] B. Milovic and M. Milovic, "Prediction and Decision Making in Health Care using Data Mining," International Journal of Public Health Science (IJPHS) Vol. 1, No. 2, December 2012, pp. 69~78 ISSN: 2252-8806.

[18] A. Bais, K. Deshmukh, and M. Shrivastava, "Implementation of Decision Tree," Interna-

tional Journal of Engineering and Advanced Technology (IJEAT) ISSN: 2249 – 8958, Volume-2, Issue-2, December 2012.

[19] Breiman, L, "Random Forests. Machine Learning," 45, 5–32 (2001). https://doi.org/10.1023/A:101093340432

[20] Kulkarni, V. Y. & Sinha, P, "Random Forest Classifiers: A Survey and Future Research Directions," International Journal of Advanced Computing, 36(1):1144–1153. 2013

[21] Bentéjac, C., Csörgő, A. & Martínez-Muñoz, G, "A comparative analysis of gradient boost-ing algorithms," Artif Intell Rev 54, 1937–1967 (2021). https://doi.org/10.1007/s10462-020-09896-5

[22] Ramraj S, Nishant Uzir, Sunil R, and Shatadeep Banerjeeb, "Experimenting XGBoost Algorithm for Prediction and Classification of Different Datasets," International Journal of Con-trol Theory and Applications ISSN: 0974–5572. International Science Press. Vol. 9, Num. 40. 2016

[23] A. S. Sunge, Y. Heryadi, E. Abdurachman, and I. H. Kartowisatro, "Machine Learning Methods for Predicting the Necessity of Caesareans Section of Childbirth," IEEE 5th Inter-national Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE), 2021, pp. 100-104.

[24] C. J. Sims, L. Meyn, R. Caruana, R. B. Rao, T. Mitchell, and M. Krohn, "Predicting cesarea delivery with decision tree models," American Journal of Obstetrics and Gynecology Vol. 183, Issue 5, November 2000, Pages 1198-1206.

[25] Lakshmi, B.N, Indumathi, T.T, and Ravi, Nandini, "A Study on C.5 Decision Tree Classification Algorithm for Risk Predictions During Pregnancy" International Conference on Emerging Trends in Engineering, Science and Technology (ICETEST - 2015). Procedia Technology 24 (2016) 1542 – 1549.

[26] Akhan Akbulut, Egemen Ertugul, Varol Topcu, "Fetal health status prediction based on maternal clinical history using machine learning techniques" Comput Methods Programs Biomed. 2018 Sep; 163:87-100. doi: 10.1016/j.cmpb.2018.06.010. Epub 2018 Jun 14. PMID: 30119860.

[27] Beksac MS, Tanacan A, Bacak HO, Leblebicioglu K. Computerized prediction system for the route of delivery (vaginal birth versus cesarean section). J Perinat Med. 2018 Oct 25;46(8):881-884. doi: 10.1515/jpm-2018-0022. PMID: 29570455.

[28] K han, Nafiz Imtiaz and Mahmud, Tahasin and Islam, Muhammad Nazrul and Mustafina, Sumaiya Nuha, "Prediction of Cesarean Childbirth using Ensemble Machine Learning Methods", iWAS '20: Proceedings of the 22nd International Conference on Information Integration and Web-based Applications & ServicesNovember 2020 Pages 331–339https://doi.org/10.1145/3428757.3429138.

[29] Kamepalli Sujatha. (1). Early Prediction and Analysis of Caesarian Section Delivery for Better Healthcare Operational Decisions. International Journal of Advanced Science and Technology, 29 (3), 10789 - 10799. Retrieved from http://sersc.org/journals/index.php/IJAST/article/view/27886.

[30] K. Wang, J. Tian, C. Zheng, H. Yang, J. Ren, Y. Liu, Q. Han, and Y. Zhang, "Interpretable prediction of 3-year all-cause mortality in patients with heart failure caused by coronary heart disease based on machine learning and SHAP," Computers in Biology and Medicine, Vol. 137, 2021, 104813, ISSN 0010-4825, https://doi.org/10.1016/j.compbiomed.2021.104813

[31] M. Lipschuetz, J. Guedalia, A. Rottenstreich, M. N. Persky, S. M. Cohen, D. Kabiri, G. Levin, S. Yagel, R. Unger, and Y. Sompolinsky, "Prediction of vaginal birth after cesarean deliveries using machine learning," American Journal of Obstetrics and Gynecology Vol. 222, Issue 6, June 2020, Pages 613.e1-613.e12.

[32] M. N. Islam, S. N. Mustafina, T. Mahmud, N. I. Khan, "Machine learning to predict preg-nancy outcomes: a systematic review, synthesizing framework, and future research agenda," BMC Pregnancy Childbirth 22, 348 (2022). https://doi.org/10.1186/s12884-022-04594-2.

[33] R. R. S. Clark, and J. Hou, "Three machine learning algorithms and their utility in exploring risk factors associated with primary cesarean section in low-risk women," A Methods Paper. Res Nurs Health. 2021 Jun;44(3):559-570. doi: 10.1002/nur.22122. Epub 2021 Mar 2. PMID: 33651381; PMCID: PMC8068617.

[34] S. A. Abbas, R. Riaz, S. Z. H. Kazmi, S. S. Rizvi, and S. J. Kwon, "Cause Analysis of Caesarian Sections and Application of Machine Learning Methods for Classification of Birth Data," in IEEE Access, vol. 6, pp. 67555-67561, 2018, doi: 10.1109/ACCESS.2018.2879115.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[35] M. A. Deif, R. E. Hammam, A. A. A Solyman, "Gradient Boosting Machine Based on PSO for prediction of Leukemia after a Breast Cancer Diagnosis," International Advanced Science Engineering Information Technology Vol.11.

[36] A. F. M. Batista, C. S. G, Diniz, E. A. Bonilha, I. Kawachi and A. D. P. C. Filho, "Neonatal mortality prediction with routinely collected data: a machine learning approach," BMC Pedi-atrics (2021) 21:322.

[37] Yuchun Liu, Zhihui Liu, Xue Luo, and Hongjingtian Zhao, "Diagnosis of Parkinson's dis-ease based on SHAP value feature selection," Biocybernetics and Biomedical Engineering, Vol. 42, Issue 3, 2022, Pages 856-869, ISSN 0208-5216, https://doi.org/10.1016/j.bbe.2022.06.007.

[38] A. Guryanov, "Efficient Computation of SHAP Values for Piecewise-Linear Decision Trees," International Conference on Information Technology and Nanotechnology (ITNT), 2021, pp. 1-4, doi: 10.1109/ITNT52450.2021.9649051.

[39] S. M. Lundberg, and S. Lee, "A Unified Approach to Interpreting Model Predictions," Ad-vances in Neural Information Processing Systems 30 (NIPS 2017).

[40] Stone. M, "Cross-Validatory Choice and Assessment of Statistical Predictions," Journal of the Royal Statistical Society. Series B (Methodological), vol. 36, no. 2, 1974, pp. 111–47. JSTOR, http://www.jstor.org/stable/2984809. Accessed 8 Sep. 2022.

[41] Gholamy. Afshin, Kreinovich. Vladik, and Kosheleva. Olga, "Relation Between Training and Testing Sets," A Pedagogical Explanation. A Pedagogical Explanation" (2018). Depart-mental Technical Reports (CS). 1209. https://scholarworks.utep.edu/cs_techrep/1209.

[42] C. Sammut and G. I. Webb, "Encyclopedia of Machine Learning," Springer Science & Business Media, 2011.

[43] M. Hossin and M. N. Sulaiman, "A Review on Evaluation Metrics for Data Classification Evaluations," International Journal of Data Mining and Knowledge Management Process. vol.5, no.2, 2015.

[44] D. M. Powers, "Evaluation: From Precision, Recall and F-Measure to ROC, Informedness, Markedness and Correlation," Journal of machine learning research, vol. 2, no. 1, 2011, pp. 37–63.

[45] Y. Sasaki, "The truth of the F-measure," Teach Tutor mater. vol. 1, no. 5, 2007, pp. 1-5