International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Advancing Preeclampsia Prediction with Machine Learning: A Comprehensive Systematic Literature Review

R. Topan-Aditya Rahman [1], Muhammad-Modi Lakulu *[2], Ismail-Yusuf Panessai [3]

Submitted: 19/04/2023 **Revised**: 21/06/2023 **Accepted**: 06/07/2023

Abstract: Preeclampsia is one of the leading causes of maternal mortality, which is a serious problem during pregnancy, which is further complicated by issues related to pathophysiology and etiology. The focus of this research is on the early detection of preeclampsia by using machine learning with multiple algorithms. Specifically, the aim of this study is to identify the causes of preeclampsia. A total of 21 articles were obtained from four scientific databases, namely ScienceDirect, Scopus, IEEE, and PubMed, which were published between 2018 and 2022, using several keywords such as "Artificial Intelligence", "Machine Learning", "Prediction", and "Preeclampsia". The method of the review adhered to the principles outlined in a guideline published by Preferred Reporting Items for Systematic Review and Meta-analysis (PRISMA). The systematic review of these articles was focused on the accuracy of prediction of preeclampsia using machine learning. The results showed machine learning was the most popular method, garnering 40.4% of mentions, followed by deep learning (11.5%), hybrid learning (2%), and other methods (19%), while other types of features, such as cell DNA, cohort, and resonance imaging received sizable mention (46.1%). Stochastic Gradient Boosting (SGBoost), which had an accuracy of 97.3%, was the most accurate algorithm. Machine learning is, therefore, deemed to be the best method for predicting pregnancy outcomes in light of these findings. Clearly, further research is needed to determine the best algorithm for developing prenatal diagnosis models, particularly for the early detection of preeclampsia.

Keywords: *Artificial Intelligence, Machine Learning, Pregnancy, Preeclampsia, Prediction.*

1. Introduction

According to the World Health Organization (WHO), over 800 women die every day from preventable causes related to pregnancy around the world [1]. One of the leading causes of maternal and fetal morbidity and mortality is preeclampsia (PE) [2]. This condition poses inherent dangers that can result in serious medical complications for both the mother and the baby. High blood pressure (hypertension), fluid retention (edema), and an increased protein rate in the urine are all symptoms of preeclampsia (proteinuria) [3]. Because such indications are not visible at the outset of pregnancy, preeclampsia may be difficult to detect. Preeclampsia is multi-organ hypertension complicating 3% to 8% of pregnancies [4]. It affects 2%–8% of all pregnancies and is a leading cause of maternal and perinatal morbidity and mortality [5]. In its severe form, PE can cause stroke, renal failure, intracranial hemorrhage, coagulopathy, pulmonary edema, bleeding, and even death [6]. Preeclampsia is the leading cause of maternal death in the United States [7]. Although the etiology and pathogenesis of preeclampsia are unknown, it is acknowledged that childbirth is the only cure; consequently, a false-positive test predicting preeclampsia may result in unduly early delivery [8]. Premature and low birth weight newborns result from this predicament, leading to more inefficient use of the neonatal Intensive Care Unit (ICU). To make matters worse, predictive models of preeclampsia developed in prior studies have four flaws: (1) no significant predictor exists for all preeclampsia subtypes; (2) biased predictive performance; (3) low precision or positive predictive value; and (4) the need for high resource settings to develop predictive models [9]. Many studies are used to define a woman as a woman at high risk of developing preeclampsia. In this study, the variables as risk factors include maternal age, education, parity, pregnancy interval, hypertension, hemoglobin, body mass index, history of preeclampsia and gestational diabetes mellitus [10], [11], [12], [13].

Over the past ten years, artificial intelligence technology (AI) has significantly been used in the industrial and health sectors. Most recently, AI-based technologies have been used to diagnose pregnancy [14]. AI-based technologies have been hailed as a potent tool for assessing disparate data sources. In this regard, over 75% of studies of AI-based technology indicate that the use of such technologies can help improve the detection of pregnancy or pregnancy disorders [15]. Moreover, they have the potential to assist in

[1] *Graduate Program Doctoral Student, Sultan Idris Education University, Malaysia; Sari Mulia University, Indonesia*
ORCID ID: 0000-0003-1123-2934
[2] *Faculty of Computing and Meta-Technology, Sultan Idris Education University, Malaysia.*
SCOPUS ID: 35810606800
[3] *Faculty of Computing and Meta-Technology, Sultan Idris Education University, Malaysia.*
ORCID ID: 0000-0003-1211-7137
* *Corresponding Author Email: topanaditya85@gmail.com; modi@meta.upsi.edu.my*

decision-making and improve medical care. Specifically, AI can be applied in health care for modeling, diagnosis, early detection, and monitoring [16]. However, its application in this sector is complicated by the noisy class of outcomes caused by the disease's variable gene expression. On the plus side, such problems can be solved utilizing machine learning (ML) technologies [17]

For example, machine learning is well-suited for predictive modeling of pregnancy outcomes [18]. Essentially, ML, also known as supervised, semi-supervised, unsupervised, or reinforcement learning, is a subset of artificial intelligence that involves the use of algorithms and computer models to achieve a certain goal [19]. ML techniques are often used in decision-making scenarios to produce higher predicted accuracy than standard generalized linear models [20]. As time and technology progress, the application of ML algorithms grows and becomes more advanced [21]. According to various research findings, using machine learning algorithms to detect preeclampsia has a greater level of accuracy [11]. Deep learning, in particular, is a machine learning technique that uses neural networks, which are comparable to neurons in the human brain, to extract many levels of data representation from a given input in order to solve a problem [22]. Preeclampsia is currently detected early through antenatal care examinations, thus there is a need for artificial technology support that can diagnose the incidence of preeclampsia [23]. Despite the fact that there have been numerous AI-based studies in the health sector, literacy support for the detection of preeclampsia events based on machine learning is still relatively low, therefore the discussion and debate about AI in relation to preeclampsia need to gain traction accordingly.

2. Method

2.1. Information Sources

Relevant publications were obtained from a number of reputable online scientific databases, including Scopus, ScienceDirect, IEEE, and PubMed. The articles were chosen based on a number of inclusion criteria related to the topic under investigation. The currency of articles (articles published between 2018 and 2022), the type of publications (journal articles and conference proceedings), the publication language (English), the scope of publication (full texts), and the type of fields (medical healthcare covering pregnancy, women's health, preeclampsia, artificial intelligence, and machine learning) were used as inclusion criteria.

2.2. Selection Study

This review followed the principles outlined in the Preferred Reporting Items for Systematic Review and Meta-analysis (PRISMA) guidance. The PRISMA SLR Protocol for Preeclampsia is depicted in Figure 1. The search procedure was applied to all approved databases by entering relevant keywords as search strings, such as (“Prediction” OR “Detection”) AND (“Preeclampsia” OR “Eclampsia”) AND (“Artificial Intelligence” OR “Machine Learning”). Also, advanced search options were used in all database engines to access conference journals and papers.

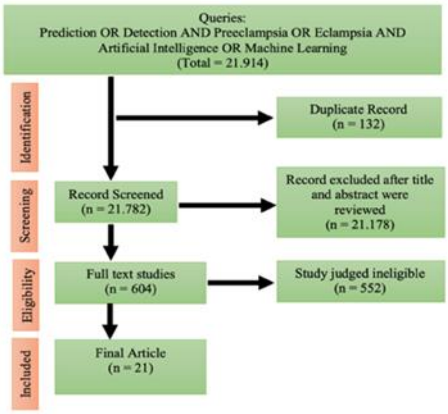

Fig 1. PRISMA SLR Protocol of Preeclampsia

2.3. Eligibility Criteria

The procedure for selecting articles in the selected database consisted of four stages. The first stage was to identify relevant articles that match the research topic by verifying their titles and abstracts, and duplicate records. At this stage, the screening of the articles was limited to full-text articles, scientific journals, and proceedings written in English and published between 2018 and 2022. The second stage entailed the screening of articles using a record screener to record the titles and abstracts of articles that had been reviewed. The third stage involved determining the eligibility of full-text publications reported in previous research. Finally, in the fourth stage, a qualitative synthesis of the final 21 articles was performed.

2.4. Country of Author

Seven of the 21 articles chosen were from the United States, three were from the United Kingdom, two each from China and Indonesia, and one each from France, Ecuador, India, Iran, Poland, Sweden, and Romania. Figure 2 shows the country of origin of the authors.

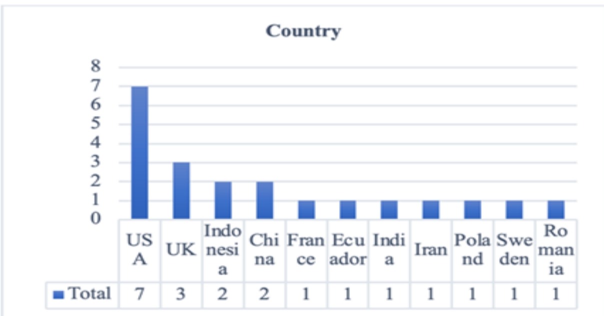

Fig 2. Country of Author

2.5. Research Methodology

The literature review utilized a total of 21 preeclampsia-related articles, and the results indicate that 86% of the studies largely focused on machine learning, with the remaining studies concentrating on hybrid learning (9%) and deep learning (5%).

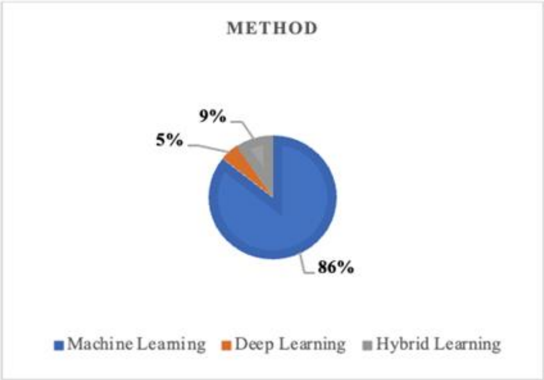

Fig 3. Methodology used in the article

2.6. Type of Algorithm discussed the article

As reviewed, the machine learning technique is more frequently employed than the deep learning approach in applications linked to pregnancy. Given its simplicity and accuracy, Random Forest, Logistic Regression, and Support Vector Machine have all been extensively employed in this context [24].

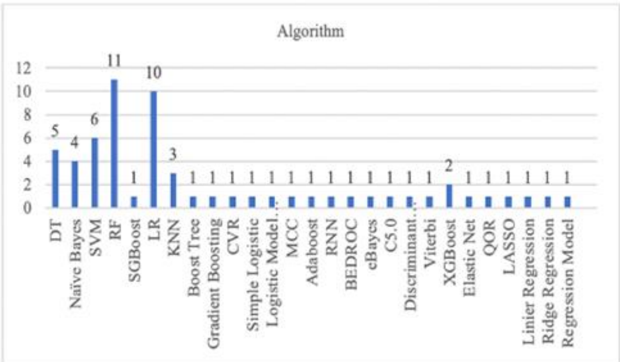

Fig 4. Algorithm used in the articles

Figure 4 shows the type algorithms discussed in the selected articles. As shown, Random Forest (RF) received the most mentions with 11 publications, followed by Logistic Regression (LR) with 10. The additional algorithms described in the chosen articles include the following: SVM (6), Decision Tree (5), Naive Bayes (4), K-Nearest Neighbor (3), and XGBoost (2). Only one mention was given to the remaining algorithms, which included Stochastic Gradient Boosting, Boost Tree, Gradient Boosting, Crossover, simple logistic, Logistic model, Matthew Correlation Coefficient, Adaboost, Recurrent neural network, Boltzmann-Enhanced Discrimination, eBayes, C5.0, Discriminant Analysis, Viterbi, Elastic Net, Quantile Ordinal Regression, LASSO, Linier Regression, Ridge Regression, and Regression Model.

3. Result

3.1. Main Features of selected articles

In this study, a classification of features was developed to help identify features used by researchers. Crucially, the selection of appropriate features is very important to achieve optimal algorithm performance. As highlighted, there are several factors that have contributed to the increasing number of incidences of preeclampsia among expecting women, which may lead to seizures, unconsciousness, and coma in a worst case scenario. As such, these women should have regular pregnancy checks, and the healthcare workers should treat their needs promptly. The classification of features of the 21 selected articles helped revealed the number (percentage) of articles with the following category of features: maternal demographics and characteristics (53%), maternal history (14%), medical condition (12%), laboratory data (15%), nutrition (3%), and protein extraction (3%), as shown in Figure 5. The sub-features of each feature is highlighted in Table 2.

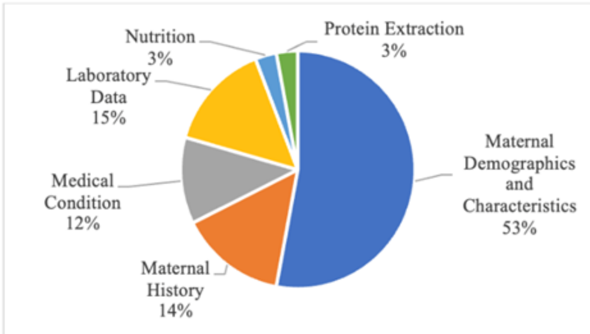

Fig 5. Main features of selected articles

Table 1. Main Characteristics of Selected Articles Type

Type of study	Temporality	Year of publication
Case Control (45,6%)	Retrospective (52,6%)	2018 (4,8%)
Cohort (33,3%)		2019 (19,1%)
Cross Sectional (4,8%)	Prospective (47,4%)	2020 (23,8%)
Explorative (19,1%)		2021 (52,3%)

Table 1 shows the main characteristics of the types of selected articles. Case control studies account for 45.6% of all studies, cohort studies account for 33.3%, cross-sectional studies account for 4.8%, and exploratory studies account for 19.1%. In terms of temporality, 52.6% employed a retrospective study strategy, whereas 47.4% used a prospective study approach. The year of article publication used in 2018 was 4.8%, 19.1% in 2019, 23.8% in 2020, and 52.3% in 2021. Table 2 highlights the sub-features of the main features.

Table 2. Sub main features

features	Sub features	Ref
Maternal demographics and characteristic	Maternal Age, Parity, Education, Job, BMI, Race, ethnicity, Dietary habits, blood group, region of birth, number of pregnancies, Source of drinking water, Health insurance provider, fetal gender, Interval Pregnancy, Smoking, Alcohol, Fetal Sex.	Jhee, et al (2019), Liu., et al (2021), Venkatesh, et al (2021), Sufriyana, et al (2020), He, et al (2021), Sufriyana,, et al (2020), Yoffe, et al (2018), Gao, et al (2019), Maung, et al (2019), Moufarrej, et al (2021), Saroj, et al (2021), Manoochehri,,et al (2021), Lewandowska, et al (2020), Wang, et al (2021), Sandstrom (2019), Rocha, et al (2021), Li, et al (2021), Pietsch, et al (2021)
Maternal History	Previous miscarriage (abortion), PE in a previous pregnancy, Prior Cesarean Delivery, History of seizure disorder, Family History of hypertension, family history of diabetes.	Li, et al (2021), Sandstrom (2019), Lewandowska, et al (2020), Jhee, et al (2019), Yoffe, et al (2018)
Medical Condition	Diabetes Mellitus Chronic, Preeclampsia, Chronic Hypertension, Hemoglobin, cardiac disease, obesity.	Venkatesh, et al (2021), He, et al (2021), Lewandowska, et al (2020), Wang, et al (2021)
Laboratory Data	WBC, 103/uL, Hemoglobin, Platelet counts, BUN, Creatinine, Total biliburin, Potassium, TCO2, Calcium, Magnesium, UPCR, Blood are ribosomal, Globulin RNAs, Uterine artery (UtA) Doppler measure, Mean arterial pressure, Glucose, Creatinine.	Sufriyana, et al (2020), Gao, et al (2019), Wang, et al (2021), Jhee, et al (2019), Rocha, et al (2021)
Nutrition	Alpha-linolenic acid. Salt, Caffeine, Organic vegetables, Fruits, Fibres, Fish, Dairy product calcium, Iron, Water, Olive oil, vitamins	Marin, et al (2020)
Protein extraction	ACE, ERBB2, BCHE, F2, CES1, AGTR1, REN, MMP9, MMP2 and ABCG2	Tejera, et al (2021)

As shown, the main feature of maternal demographics and characteristics encompasses several sub-features such as maternal age, parity, education, job, BMI, race, ethnicity, dietary habits, blood group, region of birth, number of pregnancies, source of drinking water, health insurance provider, fetal gender, interval pregnancy, smoking, alcohol, and fetal, which appear in 18 articles. The main feature of maternal history covers a number of sub-features including previous miscarriage (abortion), PE in a previous pregnancy, prior caesarean delivery, history of seizure disorder, family history of hypertension, and family history of diabetes, which appear in 5 articles. For the main feature of medical conditions, several sub-features such as chronic diabetes mellitus, preeclampsia, chronic hypertension, hemoglobin, and cardiac disease obesity are discussed in 4 articles. For the main feature of laboratory data, there are 5 articles discussing its sub-features such as WBC, 103/uL, Hemoglobin, Platelet counts, BUN, Creatinine, Total biliburin, Potassium, TCO2, Calcium, Magnesium, UPCR, Blood are ribosomal, Globulin RNAs, Uterine artery (UtA) Doppler measure, Mean arterial pressure, and Glucose, Creatinine. For main feature of nutrition, there is only 1 article that discusses its sub-features such as alpha-linolenic acid. Salt, Caffeine, Organic vegetables, Fruits, Fibres, Fish, Dairy product calcium, Iron, Water, Olive oil, and vitamins. Likewise, there is only 1 article that touches on several sub-features of the main feature of protein extraction, namely ACE, ERBB2, BCHE, F2, CES1, AGTR1, REN, MMP9, MMP2 and ABCG2.

Interestingly, several studies have focused on the use of various parameters in predicting the incidence of preeclampsia, such as alcohol consumption, ethnicity, drug abuse among pregnant woman, Crown rump length, Uterine Pulpability Index, and the pregnant woman's medical condition, with the results providing several benefits for clinical decision-making [25]. In addition to the aspects listed above that cause preeclampsia, studies have been conducted to predict the occurrence of preeclampsia utilizing plasma or DNA examinations [26].

3.2. Relates to the success of machine learning to predict preeclampsia

Preeclampsia (PE) is a hypertensive pregnancy disease that affects both maternal and fetal health and affects 3% to 8% of pregnancies [27]. PE can induce stroke, renal failure, cerebral hemorrhage, coagulopathy, pulmonary oedema, bleeding, and even mortality in its severe form [4], [6]. Preeclampsia is currently diagnosed as a pregnancy syndrome in pregnant women characterized by hypertension (blood pressure ≥ 140 mmHg and/or diastolic blood pressure ≥ 90 mmHg two times with an interval of at least 4 hours) and proteinuria (≥ 0.3 g/24 hours or 2+ on a dipstick testing) that occurs after 20 weeks of gestation [28].

Preeclampsia affects only humans and some primates, and despite several investigations, understanding its etiology has been difficult [7], [28]. Preeclampsia is defined as hypertension that develops 20 weeks after pregnancy with symptoms of proteinuria or other terminal organ damage, and it is associated with increased maternal and neonatal morbidity and mortality, particularly if it occurs early [29].

In the management of preeclampsia complications, serious clinical management can be a useful preventive measure [30]. Neonatal morbidity, which can cause neonatal premature, can be avoided by early preeclampsia incidence prediction [9] [31]. As a basis for identifying causes or factors for maternal and child healthcare, artificial intelligence and machine learning offer a novel way for predictive modeling, diagnosis, early detection, and monitoring in perinatal that aids medical professionals in giving pregnant women the care they need [32], [33].

In order to ensure that pregnant women receive the right care and resource management is more efficient, numerous studies have created algorithms related to forecasting the prevalence of preeclampsia helped by machine learning [34]. To identify risk factors for preeclampsia incidence, such as gestational age, parity, chronic hypertension, gestational hypertension, and a history of preeclampsia, is an example of constructing a predictive model using machine learning [35].

Preeclampsia incidence has been predicted using a variety of factors, including alcohol consumption, ethnicity, drugs abuse, length of the Crown rump, uterine pulsatility index, and the overall health of pregnant women [25]. The results of these studies are useful for clinical decision-making. Studies have been conducted to predict the incidence of preeclampsia by plasma or DNA testing in addition to using the factors of preeclampsia mentioned above. [13], [26].

Many researchers have also developed numerous algorithm models, such as AI models, machine learning, and deep learning, through the construction of algorithm models to establish the optimal algorithm for predicting the incidence of preeclampsia [17], [36]. In addition to the development of machine learning algorithm models, many literature reviews have been published that discuss the role of machine learning and artificial intelligence in clinical practice and provide an overview of how the developed model can predict pregnancy cases as the basis for optimal care and treatment therapy [34].

3.3. Performance of machine learning method

Based on the findings of numerous selected publications, it is reasonable to conclude that machine learning can be used to predict the incidence of preeclampsia. the application of algorithm in research is based on the most accurate modelling. The result of the integration of machine learning relating to the prediction incidence of preeclampsia in all selected articles are base on the Accuracy, AUC, Sensitivity, and Specificity. To this end, several researchers have developed a number of algorithm models to determine the most optimal algorithm in predicting the incidence of preeclampsia. There are 21 articles in this study that focus on identifying or detecting such as ailment. The performance result of algorithm are displayed in Table 3.

Table 3. Performance of machine learning method

Author	Sample Size	Algorithm	Best Model	Performance Metrics			
				AUC	Acc	Sens	Spec
Jhee, et al (2019)	11,006	Decision Tree Naïve Bayes Support Vector Machine Random Forest Stochastic Gradient Boosting Logistic Regression	Stochastic Gradient Boosting (SGBoost)	-	0,973	0,603	-

Liu,F., et al (2021)	1,648	Random Forest SVM Decision Tree K-Nearest Neighbor Naïve Bayes Boost Tree	Boost Tree	0,701	0,951	-	
Venkatesh, et al (2021)	10,100	Random Forest Gradient Boosting Tree Logistic Regression	Logistic Regression (LR)	-	0,940	-	-
Sufriyana, et al (2020)	95	Random Forest CVR Naïve Bayes Simple Logistic Logistic Model Tree Multi Class Classifier Logistic Regression	Random Forest (RF)	0,97	0,926	0,907	0,930
He, et al (2021)	64	Random Forest	Random Forest (RF)	0,81	0,88	-	-
Sufriyana,, et al (2020)	23,301	Logistic Regression, Decision Tree, Artificial Neural Network Random Forest SVM Ensemble algorithm	Random Forest (RF)	0,88	0,82	0,52	0,91
Yoffe, et al (2018)	100	Logistic Regression	Logistic Regression (LR)	0,86	0,76	0,87	0,72
Munchel, S. Et al (2020)	43	Adaboost	Adaboost	-	0,850	-	-
Gao, et al (2019)	25,689	Recurrent Neural Network	Recurrent Neural Network (RNN)	0,827	-	0,965	0,330
Max t. Aung, et al (2019)	173	Random Forest	Random Forest (RF)	0,84	-	-	-
Tejera, et al (2021)	155	Boltzmann-Enhanced Discrimination of RO (BEDROC)	BEDROC	0,831	0,830	-	-
Moufarrej, et al (2021)	404	Empirical Bayes (eBayes)	eBayes	0,710	-	0,750	0,560
Saroj, et al (2021)	90	Logistic Regression Decision Tree	Logistic Regression (LR)	0,81	0,78	1,00	0,79
Manoochehri,,et al (2021)	1,452	Logistic Regression K-Nearest Neighbor C5.0 Discriminant analysis Random Forest	Support Vector Machine (SVM)	-	0,791	0,800	0,780

		Support Vector Machine					
Marin, et al (2020)	110	Viterbi	Viterbi	-	0,790	0,950	0,700
Lewandowska, et al (2020)	912	Logistic Regression	Logistic Regression (LR)	0,600	0,716		
Wang, et al (2021)	907	Logistic Regression Random Forest Support Vector Machine Naïve Bayes eXtreme Gradient Boosting	Random Forest (RF)	0,711	0,817	0,815	0,984
Sandstrom (2019)	62,562	Random Forest	Random Forest (RF)	0,680	-	0,293	-
Rocha, et al (2021)	3.876.666	eXtreme Gradient Boosting Elastic Net Quantile Ordinal Regression LASSO Linear Regression Ridge Regression Decision Tree	eXtreme Gradient Boosting (XGBoost)	RMSE 2.09 (2.090–2.097)	-	-	-
Li, et al (2021)	3.759	Logistic regression Random Forest Support Vector Machine eXtreme Gradient Boosting	eXtreme Gradient Boosting (XGBoost)	0,955	0,920	-	0,447
Pietsch, et al (2021)	110	Regression Model	Regression Model	0.580	-	-	0,680

As shown in Table 2, it can be seen that the SGBoost algorithm method is the best method in predicting the incidence of preeclampsia with an accuracy rate of 97.3% and a sensitivity value of 0.603. These results are similar to those of Sufriyana, et al. (2020), who used 95 samples, a considerably lower number, and had a very high accuracy rate of 92.6%. Similar results were found in He, et al.'s research from 2021, which had a sample size of 64 and an accuracy rate of 81%. In comparison, Sandstrom's research (2019) has a fairly low accuracy rate of 68% while having a large number of samples. Liu, F., et al. (2021) assert that the Boost Tree Algorithm, which can achieve an accuracy rate of 95.1%, offers a high level of accuracy with smaller sample sizes.

Clearly, the use of a small and big number of samples must be compared in order to demonstrate which method provides the best modeling based on the level of accuracy. Additionally, even though many algorithm models have already been created by earlier researchers, many still use retrospective study methodologies (table 1). Because the sample employed is often pregnant women who have preeclampsia, it is believed that the use of the retrospective study method is insufficient to serve as a basis for the machine learning method to be applied in forecasting the occurrence of preeclampsia. Therefore, it is important to conduct studies using the cohort study method to ensure the research results are more credible.

In managing preeclampsia complications, serious clinical management can be used as an effective preventive measure [30]. Neonatal morbidity, which can cause neonatal premature, can be avoided by early preeclampsia incidence prediction [31]. Artificial intelligence and machine learning offer a new approach for predictive modeling, diagnosis, early detection, and monitoring in perinatal, which can help practitioners in determining the causes of such ailments such that they can provide the best possible care for pregnant women [22] [33]. In this regard, machine learning can be leveraged to help diagnose various diseases and genetic disorders, including preeclampsia [37].

4. Discussion

Preeclampsia is one of the most dangerous causes of complications in pregnancy and deaths of women, which need to be addressed urgently [38]. Preeclampsia early diagnosis has been a clinical problem to date [39]. The development of models using technology, such as machine learning models and the examination of plasma and cell DNA, has been the subject of extensive research and investigation, but the outcomes have not been able to effectively predict, treat, or prevent preeclampsia. To improve prenatal care services, it may be concluded that AI-based technology is advised for predicting the prevalence of preeclampsia [40].

Preeclampsia is a type of new-onset hypertension that can harm organs after 20 weeks of pregnancy. Many organ systems are involved in this complex pathophysiological process. The mechanism underlying the pathophysiology of preeclampsia remains unknown [5]. Histopathological investigation of the placenta in late pregnancy is frequently used to investigate pathophysiology. However, this simply leads to more pronounced structural defects, particularly villi [41]. Several biomarkers, including cardiovascular, placental, and urine biomarkers, are used in pathophysiology [42].

To date, pregnancy screening continues to be performed using the conventional method, with antenatal checks. However, this may not be fully effective as such a method can give rise to misleading results, given that the number of incidences of preeclampsia continues to grow unabatedly [23]. According to an expert, who is also an obstetrics and gynecology specialist, most practitioners dealing with pregnancy issues are primarily concerned with treatment rather than prevention, which inevitably leads to higher cost [43]. It is hardly surprising that early detection of preeclampsia has remained as an unresolved issue [44].

In this context, machine learning plays a vital role, providing solutions with numerous applications [45], such as disease detection and diagnosis [46]. This comprehensive review of 21 articles helped identify a number of ML techniques for diagnosing preeclampsia. With artificial intelligence, it is not a far-fetched notion for a machine to think like a human when dealing with challenges. Although not all of the techniques highlighted above can be employed in all disciplines, all existing algorithms can be further developed to serve certain specific application domains, such as medical healthcare. Recent trends show that machine learning a lot developed in the field of medicine [47].

As highlighted, preeclampsia prediction models based on machine learning technologies are able to predict the onset of a variety of diseases. As such, they can be used by clinicians to identify pregnant women who are at high risk of developing PE. Given the high accuracy of numerous algorithms, particularly SGBoost and Random Forest with Boost Tree, future research should investigate their utility in predicting or detecting preeclampsia in pregnant women, particularly those in early pregnancy with a gestational period of fewer than 20 weeks. The findings of such studies will certainly enhance existing techniques for providing critical care to these women.

5. Conclusion

Unquestionably, pregnancy is a crucial stage in a woman's life that can have an impact on both her physical and psychological wellbeing. A pregnancy has not only an element of happiness, but also an aspect of the mother's anxiety about physiological changes throughout pregnancy that might affect her emotions [15]. Pregnancy is a complicated process in which each stage of decidualization, placentation, implantation, organogenesis, fetal development, and fetal growth is dependent on the success of the previous phases. A healthy pregnancy is dependent on a complex network of interconnected biological changes involving the maternal immune system, hormone homeostasis, and placentation [48].

Pregnant women suffering from preeclampsia are particularly vulnerable to placental oxidative stress, which can cause ischemia injury [49]. Depending on the mother's health, early disease symptoms frequently interfere with pregnancy or deliveries. Preeclampsia is currently the leading cause of fetal and mother death during pregnancy. As a result, early detection of this condition is necessary, and it can be reinforced with the use of new medications. In this field of research, machine learning is well suited for modeling and predicting pregnancy outcomes. Thus far, preeclampsia has not been successfully predicted, treated, or prevented despite extensive research and investigations using technology-assisted model generation techniques, including machine learning and plasma and DNA cell analysis. In order to identify the most accurate pregnancy diagnosis model, particularly for the early detection of preeclampsia, more studies are needed.

In light of the above revelations, future studies should focus on developing more robust prediction models using several algorithms, especially SGBoost and Random Forest, to help predict and detect any signs of symptoms of preeclampsia among pregnant women in early pregnancy with a gestational period of fewer than 20 weeks. The findings of such studies can certainly help improve the understanding of this ailment, which can lead to better and improved healthcare of pregnant women. This will not only save lives of these women but also can significantly save cost as expensive treatments can be avoided through early detection of preeclampsia [50], [51].

Acknowledgements

This manuscript is part of the doctoral dissertation of the first author who was a doctoral student of the Sultan Idris Education University, Malaysia

Author contributions

R. Topan Aditya Rahman, Muhammad-Modi Lakulu, and Ismail-Yusuf Panessai contributed to the design and implementation of the research, the analysis of results, and manuscript writing.

Conflicts of interest

The authors declare that they do not have any conflicts of interest in pursuing this research.

References

[1] A. Bertini, R. Salas, S. Chabert, L. Sobrevia, and F. Pardo, "Using Machine Learning to Predict Complications in Pregnancy: A Systematic Review," *Front Bioeng Biotechnol*, vol. 9, no. January, pp. 1–16, 2022, doi: 10.3389/fbioe.2021.780389.

[2] I. C. Udenze, A. P. Arikawe, C. C. Makwe, and O. F. Olowoselu, "A prospective cohort study on the clinical utility of second trimester mean arterial blood pressure in the prediction of late-onset preeclampsia among Nigerian women," *Niger J Clin Pract*, vol. 20, no. 6, pp. 741–745, 2017, doi: 10.4103/1119-3077.208948.

[3] I. Bellos, V. Pergialiotis, D. Loutradis, A. Papapanagiotou, and G. Daskalakis, "The role of hemoglobin degradation pathway in preeclampsia: A systematic review and meta-analysis," *Placenta*, vol. 92, pp. 9–16, 2020, doi: 10.1016/J.PLACENTA.2020.01.014.

[4] D. D. Smith and M. M. Costantine, "The role of statins in the prevention of preeclampsia," *Am J Obstet Gynecol*, vol. 226, no. 2, pp. S1171–S1181, Feb. 2022, doi: 10.1016/j.ajog.2020.08.040.

[5] C. W. Ives, R. Sinkey, I. Rajapreyar, A. T. N. Tita, and S. Oparil, "Preeclampsia—Pathophysiology and Clinical Presentations: JACC State-of-the-Art Review," *J Am Coll Cardiol*, vol. 76, no. 14, pp. 1690–1702, 2020, doi: 10.1016/j.jacc.2020.08.014.

[6] V. R. Kay, N. Wedel, and G. N. Smith, "Family History of Hypertension, Cardiovascular Disease, or Diabetes and Risk of Developing Preeclampsia: A Systematic Review," *Journal of Obstetrics and Gynaecology Canada*, vol. 43, no. 2. J Obstet Gynaecol Can, pp. 227-236.e19, Feb. 01, 2021. doi: 10.1016/j.jogc.2020.08.010.

[7] E. Tejera, Y. Pérez-Castillo, A. Chamorro, A. Cabrera-Andrade, and M. E. Sanchez, "A multi-objective approach for drug repurposing in preeclampsia," *Molecules*, vol. 26, no. 4, Feb. 2021, doi: 10.3390/MOLECULES26040777.

[8] E. Jung *et al.*, "The etiology of preeclampsia," *Am J Obstet Gynecol*, vol. 226, no. 2, pp. S844–S866, 2022, doi: 10.1016/j.ajog.2021.11.1356.

[9] H. Sufriyana, Y. W. Wu, and E. C. Y. Su, "Prediction of preeclampsia and intrauterine growth restriction: Development of machine learning models on a prospective cohort," *JMIR Med Inform*, vol. 8, no. 5, 2020, doi: 10.2196/15411.

[10] E. Bartsch, K. E. Medcalf, A. L. Park, and J. G. Ray, "Clinical risk factors for pre-eclampsia determined in early pregnancy: systematic review and meta-analysis of large cohort studies," *BMJ*, vol. 353, p. i1753, Apr. 2016, doi: 10.1136/bmj.i1753.

[11] R. Chu *et al.*, "Predicting the Risk of Adverse Events in Pregnant Women With Congenital Heart Disease," *J Am Heart Assoc*, vol. 9, no. 14, Jul. 2020, doi: 10.1161/JAHA.120.016371.

[12] M. Lewandowska, "The association of familial hypertension and risk of gestational hypertension and preeclampsia," *Int J Environ Res Public Health*, vol. 18, no. 13, 2021, doi: 10.3390/ijerph18137045.

[13] M. Liu *et al.*, "Development of a prediction model on preeclampsia using machine learning-based method: a retrospective cohort study in China," *Front Physiol*, vol. 13, no. August, pp. 1–9, Aug. 2022, doi: 10.3389/fphys.2022.896969.

[14] I. *Abuelezz et al.*, "Contribution of Artificial Intelligence in Pregnancy: A Scoping Review," *Stud Health Technol Inform*, vol. 289, pp. 333–336, 2022, doi: 10.3233/SHTI210927.

[15] A. M. Oprescu, G. Miró-Amarante, L. García-Díaz, L. M. *Beltrán*, V. E. Rey, and M. Romero-Ternero, "Artificial intelligence in pregnancy: A scoping review," *IEEE Access*, vol. 8, pp. 181450–181484, 2020, doi: 10.1109/ACCESS.2020.3028333.

[16] C. D. Valle and R. Almazan, "Towards an understanding of artificial intelligence in government," in *Proceedings of the 19th Annual International Conference on Digital Government Research:*

Governance in the Data Age, New York, NY, USA: ACM, May 2018, pp. 1–2. doi: 10.1145/3209281.3209397.

[17] H. Sufriyana, Y. W. Wu, and E. C. Y. Su, "Artificial intelligence-assisted prediction of preeclampsia: Development and external validation of a nationwide health insurance dataset of the BPJS Kesehatan in Indonesia," *EBioMedicine*, vol. 54, 2020, doi: 10.1016/j.ebiom.2020.102710.

[18] C. Espinosa *et al.*, "Data-Driven Modeling of Pregnancy-Related Complications," *Trends Mol Med*, vol. 27, no. 8, pp. 762–776, 2021, doi: 10.1016/j.molmed.2021.01.007.

[19] E. Nsugbe, "A cybernetic framework for predicting preterm and enhancing care strategies: A review," *Biomedical Engineering Advances*, vol. 2, no. November, p. 100024, 2021, doi: 10.1016/j.bea.2021.100024.

[20] F. Liu *et al.*, "C1431T variant of PPARγ is associated with preeclampsia in pregnant women," *Life*, vol. 11, no. 10, 2021, doi: 10.3390/life11101052.

[21] W. Zhang, C. Wu, H. Zhong, Y. Li, and L. Wang, "Prediction of undrained shear strength using extreme gradient boosting and random forest based on Bayesian optimization," *Geoscience Frontiers*, vol. 12, no. 1, pp. 469–477, Jan. 2021, doi: 10.1016/j.gsf.2020.03.007.

[22] R. Ramakrishnan, S. Rao, and J. R. He, "Perinatal health predictors using artificial intelligence: A review," *Women's Health*, vol. 17, 2021, doi: 10.1177/17455065211046132.

[23] F. T. D. S. Lusyana. Purwadi. Taufik, "Implementasi Sistem Pakar Untuk Mendiagnosa Penyakit Dalam Pada Manusia Menggunakan Metode Dempster Shafer," *Jurnal Sarjana Teknik Informatika*, vol. 1, no. 3, pp. 118–127, 2022.

[24] R. Susetyoko, W. Yuwono, E. Purwantini, and N. Ramadijanti, "Perbandingan Metode Random Forest, Regresi Logistik, Naïve Bayes, dan Multilayer Perceptron Pada Klasifikasi Uang Kuliah Tunggal (UKT)," *Jurnal Infomedia: Teknik Informatika, Multimedia & Jaringan*, vol. 7, no. 1, pp. 8–16, 2022.

[25] R. K. Saroj and M. Anand, "Environmental factors prediction in preterm birth using comparison between logistic regression and decision tree methods: An exploratory analysis," *Social Sciences & Humanities Open*, vol. 4, no. 1, p. 100216, 2021, doi: 10.1016/j.ssaho.2021.100216.

[26] M. Pietsch *et al.*, "APPLAUSE: Automatic Prediction of PLAcental health via U-net Segmentation and statistical Evaluation," *Med Image Anal*, vol. 72, p. 102145, 2021, doi: 10.1016/j.media.2021.102145.

[27] M. Amin-Beidokhti, H. Sadeghi, R. Pirjani, L. Gachkar, M. Gholami, and R. Mirfakhraie, "Differential expression of *Hsa-miR-517a/b* in placental tissue may contribute to the pathogenesis of preeclampsia," *Journal of the Turkish-German Gynecological Association*, vol. 22, no. 4, pp. 273–278, Dec. 2021, doi: 10.4274/jtgga.galenos.2021.2021.0062.

[28] S. et al Munchel, "A molecular signal f or p re eclampsia," vol. 16, no. September, p. 2020, 2020.

[29] D. L. Rolnik, K. H. Nicolaides, and L. C. Poon, "Prevention of preeclampsia with aspirin," *Am J Obstet Gynecol*, vol. 226, no. 2, pp. S1108–S1119, 2022, doi: 10.1016/j.ajog.2020.08.045.

[30] Y. Li *et al.*, "Novel electronic health records applied for prediction of pre-eclampsia: Machine-learning algorithms," *Pregnancy Hypertens*, vol. 26, no. August, pp. 102–109, Dec. 2021, doi: 10.1016/j.preghy.2021.10.006.

[31] J. H. Jhee *et al.*, "Prediction model development of late-onset preeclampsia using machine learning-based methods," *PLoS One*, vol. 14, no. 8, p. e0221202, Aug. 2019, doi: 10.1371/journal.pone.0221202.

[32] R. Ramakrishnan, S. Rao, and J. R. He, "Perinatal health predictors using artificial intelligence: A review," *Women's Health*, vol. 17, 2021, doi: 10.1177/17455065211046132.

[33] L. Myatt, "The prediction of preeclampsia: the way forward," *Am J Obstet Gynecol*, vol. 226, no. 2, pp. S1102-S1107.e8, 2022, doi: 10.1016/j.ajog.2020.10.047.

[34] C. Gao, S. Osmundson, D. R. Velez Edwards, G. P. Jackson, B. A. Malin, and Y. Chen, "Deep learning predicts extreme preterm birth from electronic health records," *J Biomed Inform*, vol. 100, no. January, p. 103334, 2019, doi: 10.1016/j.jbi.2019.103334.

[35] K. K. Venkatesh, J. Beckham, A. M. Friedman, K. Boggess, E. Jelovsek, and D. M. Stamilio, "A prediction model for postpartum readmission for hypertension or preeclampsia," *Am J Obstet Gynecol*, vol. 224, no. 2, pp. S42–S43, Feb. 2021, doi: 10.1016/j.ajog.2020.12.058.

[36] Q. Wang, D. Liu, and G. Liu, "Value of Ultrasonic Image Features in Diagnosis of Perinatal Outcomes of Severe Preeclampsia on account of Deep Learning Algorithm," *Comput Math Methods Med*, vol. 2022, 2022, doi: 10.1155/2022/4010339.

[37] P. Mobadersany, L. A. D. Cooper, and J. A. Goldstein, "GestAltNet: aggregation and attention to improve

deep learning of gestational age from placental whole-slide images," *Laboratory Investigation*, vol. 101, no. 7, pp. 942–951, 2021, doi: 10.1038/s41374-021-00579-5.

[38] P. Nisha and J. Gonzales, "乳鼠心肌提取 HHS Public Access," *Physiol Behav*, vol. 177, no. 7, pp. 917–918, 2017, doi: 10.1001/jamainternmed.2017.1422.Preeclampsia.

[39] S. Rana, E. Lemoine, J. P. Granger, and S. A. Karumanchi, "Preeclampsia," *Circ Res*, vol. 124, no. 7, pp. 1094–1112, Mar. 2019, doi: 10.1161/CIRCRESAHA.118.313276.

[40] T. Tomimatsu, K. Mimura, M. Endo, K. Kumasawa, and T. Kimura, "Pathophysiology of preeclampsia: An angiogenic imbalance and long-lasting systemic vascular dysfunction," *Hypertension Research*, vol. 40, no. 4, pp. 305–310, Apr. 2017, doi: 10.1038/HR.2016.152.

[41] C. Bertholdt, M. Dap, M. Beaumont, J. Duan, and O. Morel, "New insights into human functional ultrasound imaging," *Placenta*, vol. 117, no. August 2021, pp. 5–12, 2022, doi: 10.1016/j.placenta.2021.10.005.

[42] T. M. MacDonald, S. P. Walker, N. J. Hannan, S. Tong, and T. J. Kaitu'u-Lino, "Clinical tools and biomarkers to predict preeclampsia," *EBioMedicine*, vol. 75, p. 103780, 2022, doi: 10.1016/j.ebiom.2021.103780.

[43] D. R. (Fakultas K. U. G. M. Hadiati, "Terapi Preeklamsia Perlu Dikembangkan untuk Menekan Angka Kematian Ibu," *Article text*, no. 07 September 2017, 2017.

[44] X. Han *et al.*, "Differential dynamics of the maternal immune system in healthy pregnancy and preeclampsia," *Front Immunol*, vol. 10, no. JUN, pp. 1–13, 2019, doi: 10.3389/fimmu.2019.01305.

[45] A. Martínez-Velasco, L. Martínez-Villaseñor, and L. Miralles-Pechuán, "Machine learning approach for pre-eclampsia risk factors association," *ACM International Conference Proceeding Series*, no. November, pp. 232–237, 2018, doi: 10.1145/3284869.3284912.

[46] R. Kuciene and V. Dulskiene, "Associations of maternal gestational hypertension with high blood pressure and overweight/obesity in their adolescent offspring: a retrospective cohort study," *Sci Rep*, vol. 12, no. 1, pp. 1–9, 2022, doi: 10.1038/s41598-022-07903-z.

[47] A. Roihan, P. A. Sunarya, and A. S. Rafika, "Pemanfaatan Machine Learning dalam Berbagai Bidang: Review paper," *IJCIT (Indonesian Journal on Computer and Information Technology)*, vol. 5, no. 1, pp. 75–82, May 2020, doi: 10.31294/ijcit.v5i1.7951.

[48] L. Davidson and M. R. Boland, "Towards deep phenotyping pregnancy: A systematic review on artificial intelligence and machine learning methods to improve pregnancy outcomes," *Brief Bioinform*, vol. 22, no. 5, pp. 1–29, 2021, doi: 10.1093/bib/bbaa369.

[49] A. Ho *et al.*, "Placental magnetic resonance imaging in chronic hypertension: A case-control study," *Placenta*, vol. 104, no. November 2020, pp. 138–145, 2021, doi: 10.1016/j.placenta.2020.12.006.

[50] M. N. Moufarrej *et al.*, "Early prediction of preeclampsia in pregnancy with cell-free RNA," *Nature*, vol. 602, no. 7898, pp. 689–694, 2022, doi: 10.1038/s41586-022-04410-z.

[51] L. M. Amaral, K. Wallace, M. Owens, and B. LaMarca, "Pathophysiology and Current Clinical Management of Preeclampsia," *Curr Hypertens Rep*, vol. 19, no. 8, p. 61, Aug. 2017, doi: 10.1007/s11906-017-0757-7

[52] Patil, A. ., & Govindaraj, S. K. . (2023). ADL-BSDF: A Deep Learning Framework for Brain Stroke Detection from MRI Scans towards an Automated Clinical Decision Support System. International Journal on Recent and Innovation Trends in Computing and Communication, 11(3), 11–23. https://doi.org/10.17762/ijritcc.v11i3.6195

[53] Verma, D. N.. (2022). Access Control-Based Cloud Storage Using Role-Fully Homomorphic Encryption Scheme. Research Journal of Computer Systems and Engineering, 3(1), 78–83. Retrieved from https://technicaljournals.org/RJCSE/index.php/journal/article/view/46