Check for updates

A PCA-Enhanced Ensemble Framework for Maternal Health Risk Prediction with SMOTE-ENN Balancing

Tamim Mahmud[1]* , Kulsom Akter Sinthia[1] , Prosenjit Chandra Biswas[1],
Md Sanower Hossain[1] , Sourov Kumar[1] , Sanzida Afrin Anamika[1]

[1]Department of Computer Science and Engineering, Daffodil International University,

Savar, Dhaka-1216, Bangladesh.

Email: {mahmud23105101249@diu.edu.bd, sinthia23105101410@diu.edu.bd, biswas15-14568@diu.edu.bd, hossain2305101726@diu.edu.bd, kumar23105101296@diu.edu.bd, anamika23105101044@diu.edu.bd

Abstract. Maternal health complications are a major issue of concern in the world especially in developing countries. Early detection of the high-risk pregnancies is very important in minimizing the morbidity and mortality of the mother. The traditional clinical tests often rely on a few features and manual examination, which can neglect complex nonlinear associative links between the physiological predictors. In this report, an ensemble learning model incorporating Principal Component Analysis (PCA) to reduce the number of dimensions and SMOTE-ENN to balance the classes is proposed to predict the risk of maternal health with high precision. A stacked model with K-Nearest Neighbor (KNN), Random Forest (RF), and XGBoost is used to increase the predictive accuracy. Experimental results on a maternal health dataset indicate that the suggested stacking ensemble is more effective than a single model. It has a precision of 98%, an accuracy of 97.59%, a recall of 98% and an F1-score of 98%. It does well particularly in segmenting the minority; the Mid Risk segment. The methodology provides a reliable and robust framework assist in identifying high-risk pregnancies in the early stages, which will help make timely clinical interventions.

Keywords: Maternal Health Risk, Ensemble Learning, PCA, SMOTE-ENN, Stacking Ensemble, Machine Learning,Pregnancy Prediction

1. Introduction

Maternal health remains a large issue in virtually all parts of the world as challenges during pregnancy and delivery have continued to result in maternal morbidity and mortality, particularly in the underdeveloped regions [1]. The early identification of the risk factors can significantly improve the therapeutic decision-making and reduce the risk of adverse outcomes. On the other hand the traditional methods used in assessment would often rely on manual ratings and on a limited amount of clinical information. They may not be in a position to detect any complex and nonlinear associations between the physiological indicators. It is currently a sound technique to better and more trustworthy predict the health concerns of pregnant individuals by the usage of data-driven and machine learning (ML) technologies.

However, recent advancements to supervised learning algorithms have enabled the fast and precise modeling of healthcare information to be used in solutions such as diagnosis and classification. Ensemble-based methods such as the Random Forest (RF) and Extreme Gradient Boosting (XGBoost) have performed well using different types of medical data to make predictions [2], [3]. The biggest issues with maternal health datasets are there are not enough cases that fall in the category of Mid Risk, as opposed to the cases that fall into the category of Low Risk and the ones that fall into the category of High Risk. Due to such imbalance, models are biased towards the majority of the classes, it implies that they cannot recollect or practice what they learn in the minority classes. Classical ML models also cannot

Even though there has been some progress in using machine learning to predict maternal health, there are still some big problems that haven't been solved in the current research:

- Most studies use only one machine learning model and don't use dimensionality reduction or advanced class-balancing methods [7].
- Severe class imbalance, especially the lack of Mid Risk samples, keeps making the model's performance worse and makes it hard to remember minority-class samples [8].
- Only a few studies look at using PCA to reduce features, and almost none look at using PCA with a hybrid SMOTE-ENN resampling strategy in a stacking ensemble to predict maternal health risks.
- A lot of the methods that are already out there are tested on small datasets that aren't validated enough, which makes them less useful and reliable in real-world situations.
- These gaps show that we need a single ML pipeline that can work with maternal health data that has low-dimensional features, class distributions that are very uneven, and nonlinear patterns.

© The Author(s) 2026
M. S. Arefin et al. (eds.), *Proceedings of the International Conference on Intelligent Data Analysis and Applications (IDAA 2025)*, Advances in Intelligent Systems Research 206,
https://doi.org/10.2991/978-94-6239-664-7_96

traditional methods used in assessment would often rely on manual ratings and on a limited amount of clinical information. They may not be in a position to detect any complex and nonlinear associations between the physiological indicators. It is currently a sound technique to better and more trustworthy predict the health concerns of pregnant individuals by the usage of data-driven and machine learning (ML) technologies.

However, recent advancements to supervised learning algorithms have enabled the fast and precise modeling of healthcare information to be used in solutions such as diagnosis and classification. Ensemble-based methods such as the Random Forest (RF) and Extreme Gradient Boosting (XGBoost) have performed well using different types of medical data to make predictions [2,3]. The biggest issues with maternal health datasets are there are not enough cases that fall in the category of Mid Risk, as opposed to the cases that fall into the category of Low Risk and the ones that fall into the category of High Risk. Due to such imbalance, models are biased towards the majority of the classes, it implies that they cannot recollect or practice what they learn in the minority classes.

Even though there has been some progress in using machine learning to predict maternal health, there are still some big problems that haven't been solved in the current research:

- Most studies use only one machine learning model and don't use dimensionality reduction or advanced class-balancing methods [7].
- Severe class imbalance, especially the lack of Mid Risk samples, keeps making the model's performance worse and makes it hard to remember minority-class samples [8].
- Only a few studies look at using PCA to reduce features, and almost none look at using PCA with a hybrid SMOTE-ENN resampling strategy in a stacking ensemble to predict maternal health risks.
- A lot of the methods that are already out there are tested on small datasets that aren't validated enough, which makes them less useful and reliable in real-world situations.
- These gaps show that we need a single ML pipeline that can work with maternal health data that has low-dimensional features, class distributions that are very uneven, and nonlinear patterns.

This paper makes the following important contributions to solve these problems:

- Proposes a preprocessing module based on PCA that cuts down on redundancy while keeping over 95% of the variance in the physiological features.
- Presents a combined SMOTE-ENN resampling method that both oversamples minority classes and gets rid of noisy majority samples at the same time.
- Creates a stacking ensemble that uses KNN, Random Forest, and XGBoost to make the model more stable and reliable at making predictions.
- Uses stratified 10-fold cross-validation to do a full evaluation with metrics like accuracy, precision, recall, F1-score, and multi-class ROC-AUC.

– Shows a big improvement in finding Mid Risk cases, which is one of the hardest and most important parts of predicting maternal health risk.

This paper suggests a PCA-XGBoost pipeline that uses SMOTE-ENN (Synthetic Minority Oversampling and Edited Nearest Neighbors) in order to overcome these problems. One could reduce noise and redundancy and retain over 95 percent of the variance on the PCA (Principal Component Analysis) step which is more efficient [4]. Hybrid SMOTE- ENN technique is effective to equalize the distribution of the classes by creating fake samples belonging to the minor classes and removing the noisy ones belonging to the majority ones [5,6]. Furthermore, one of the stacking ensemble models incorporating K-Nearest Neighbors (KNN), Random Forest, and XGBoost are developed to render predictions more solid [7].

In short, this paper suggests a single ML pipeline that includes PCA-based dimensionality reduction, SMOTE-ENN hybrid resampling, and a stacking ensemble (KNN, RF, XGBoost) that was tested with stratified 10-fold cross-validation. This design is made to deal with the severe class imbalance and limited feature space of the maternal health risk dataset. It gets 97.59% accuracy and 98% recall for the under-represented Mid Risk class.

To the best of our knowledge, this is the first work on the Maternal Health Risk dataset that jointly integrates PCA-based dimensionality reduction, SMOTE-ENN hybrid resampling and a stacking ensemble, and evaluates the framework with stratified 10-fold cross-validation.

2 Related Work

The maternal health is also a well-explored topic in medical and computational perspectives and focused on the detection and estimation of maternal risk factors based on various analytical and machine learning (ML) methods. The authors of this paper highlighted the importance of implementation science and innovation to achieve maternal health-related Sustainable Development Goals (SDGs), and it is important to note that data-informed strategies are essential to improve the health outcomes [1]. Ojong et al. [2] also discussed the factor that maternal well-being is composed of numerous factors which are categorized as physical, psychological, and social. Lee [3] added another facet to the health agenda by integrating oral health to maternal health into the global health structures implying the importance of holistic care to a mother.

Drife et al. [4] and Deardorff et al. [5] argued that improvement in maternal outcomes depends on use of evidence-based practice in healthcare delivery and education. Sajedinejad et al. [6] conducted a cross-sectional world health research that highlighted the inequality of the regional maternal mortality and revealed that maternal outcomes are greatly influenced by accessibility to quality healthcare. With the advent of computational approaches, a large number of researchers began to work on ML-based risk prediction strategies. Mu et al. [7] applied physiological measures such as blood pressure, blood glucose and heart rate to predict maternal health risk achieving admirable classification accuracy.

Similarly, Raza et al. [8] proposed a feature engineering-based approach to ensemble learning which improved the predictive performances by incorporating the predictive techniques of the gradient boosting as well as the feature selection. Yang and Wu [9] and Sandsaeter et al. [10] did some research on gestational diabetes and preeclampsia that it correlates with future cardiovascular complications as a consequence, thus the importance of early prediction should not be ignored.

The article by Khanlou et al. [11] and Phillimore [12]also discussed the health issues of mothers who relocate to a new country or belong to another race, which demonstrated that the ease of access to health care and its effectiveness varies.

Predictive modeling has emerged as an option with the use of artificial intelligence (AI) and big data analytics in the healthcare industry. Arshad et al. [13] studied the obstacles and opportunities related to using big data intelligence to transform healthcare systems, whereas Basile et al. [14] presented the implementation of the business intelligence systems in the clinical decision support systems. Zitnik et al. [15] provided an extensive analysis of machine learning application to the fusion of biological and medical information, outlining foundations and future opportunities in improving predictive analysis in healthcare.

While current research has examined numerous machine learning methodologies for predicting maternal health, notable limitations persist. Most previous studies utilize singular models and fail to integrate dimensionality reduction methods, such as PCA, to mitigate redundancy in physiological data. Moreover, the pronounced imbalance in the Maternal Health Risk dataset—particularly the deficiency of Mid Risk samples—has not been adequately addressed, leading to biased models with suboptimal minority-class recall. Using only SMOTE may add fake noise, and using only undersampling may remove useful samples, which leaves a gap in strong hybrid balancing methods.

Existing studies employ individual classifiers, basic oversampling, or constrained evaluation frameworks; however, none integrate PCA-based variance-preserving feature reduction with a hybrid SMOTE-ENN resampling strategy within a stacking ensemble, nor do they validate this approach using stratified 10-fold cross-validation on the Maternal Health Risk dataset. In contrast, our work combines redundancy in physiological features, severe Mid Risk underrepresentation, and model instability into a single pipeline. This makes the predictions more stable and sensitive to minority groups.

3 Methodology

The overall workflow of the proposed framework for prediction of maternal health risk is shown in Fig. 1. , it ranges through the whole process from data preprocessing to feature selection to class balancing and updatedmodel training. Our pipeline consists of PCA to reduce the dimensionality, SMOTE-ENN to resolve the class imbalance, and a stacking ensemble for enhanced predictionstability. This integrated approach guarantees that the model works with clean, balanced andinformative data leading to good classification performance of maternal risk levels.

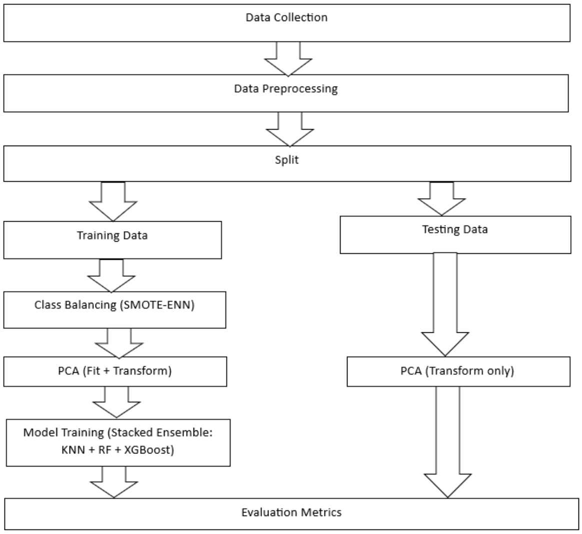

Fig. 1. System architecture of the proposed model

To deal with such issues, the current research uses a machine learning (ML)-based approach of combining data preprocessing, the balancing of classes, and dimensionality reduction, as well as, the use of ensembles of learning. The pipeline used in the methodology will be hybrid and will include SMOTE-ENN that will deal with imbalanced classes, Principal Component Analysis (PCA) that will be used to extract features, and a stacked way which is a combination of K-Nearest Neighbors (KNN), Random Forest (RF), and XGBoost to boost accuracy and robustness [5,7].

3.1 Dataset Description

The research is based on the publicly available Maternal Health Risk Dataset on Kaggle [8]. There are 1,200 samples in the dataset that have 6 physiological variables: age, systolic blood pressure (SBP), diastolic blood pressure (DBP), blood sugar (BS), body temperature, and heart rate. RiskLevel is a categorical target variable in that it shows a pregnancy to be Low Risk (0), Mid Risk (1), or High Risk (2).

The class distribution was not executed equally as the Central Risk category was significantly underrepresented, as well as the Extreme Risk category. This imbalance is the reason why techniques of balancing the classes should be applied to guarantee sound model training [9].

The dataset contains only 1,200 samples from a single source, which makes the results less diverse and less applicable to other situations. The small sample size, especially the fact that there aren't enough people in the Mid Risk class, could make the trained models less reliable. Future research should utilize larger or multicenter datasets to enhance the reliability and clinical applicability of predictions.

Table 1. Dataset's attributes description

Description	Type	Possible Values	Units of Measurement
Age of a person	Numeric	13–20	years
Systolic blood pressure	Numeric	70–150	mmHg
Diastolic blood pressure	Numeric	13–130	mmHg
Risk level classification	Categorical	high risk, mid risk, low risk	-
Heart rate	Numeric	2–98	BPM
Glucose level	Numeric	6–18	mmol/L

3.2 Data Preprocessing

Severalpre-processing steps were carried out to refine the dataset for the robust development of this model. First, any records with missing or invalidvalues were removed to avoid biased learning and maintain quality of the data. Then all the numerical features were standardized by StandardScaler that scales each variable to have a zeromean and unit variance. This normalization process is important for KNN, PCA and ensemblelearning algorithms as it prevents large numeric ranges of some features from dominating the learning algorithm.

Lastbut not least, the categorical target variable RiskLevel was converted into numerical annotation to make it machine learning-friendly. These pre-processing steps were necessary in order to getthe database cleaned and scaled homogenously for an effective downstream modelling [10].

3.3 Class Balancing Using SMOTE-ENN

The unwanted effect of poor performance of the models is the class imbalance, especially the underrepresented Mid Risk class. This issue is addressed by the SMOTE-ENN approach which oversamples minority categories (SMOTE) and removes the noisy samples of the majority (ENN) [11,12]. The class distribution as seen before SMOTE-ENN is presented in Fig. 2. and 3. It can be used, followed by the balanced distribution, as illustrated in Fig. 3. and this is more precise when making good model training.

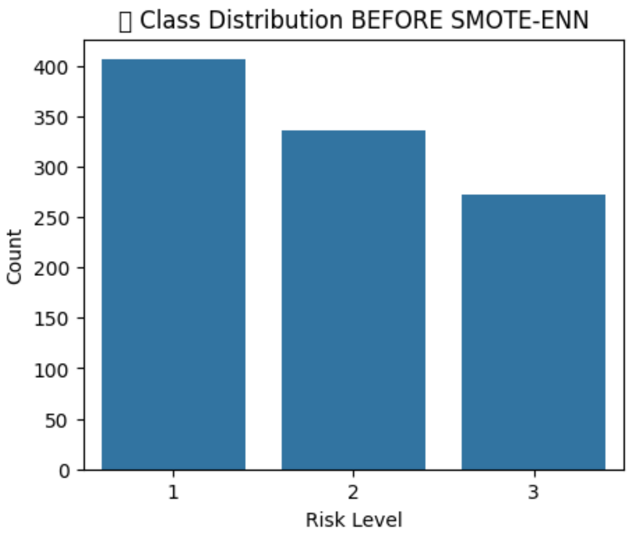

Fig. 2. Class distribution before SMOTE-ENN

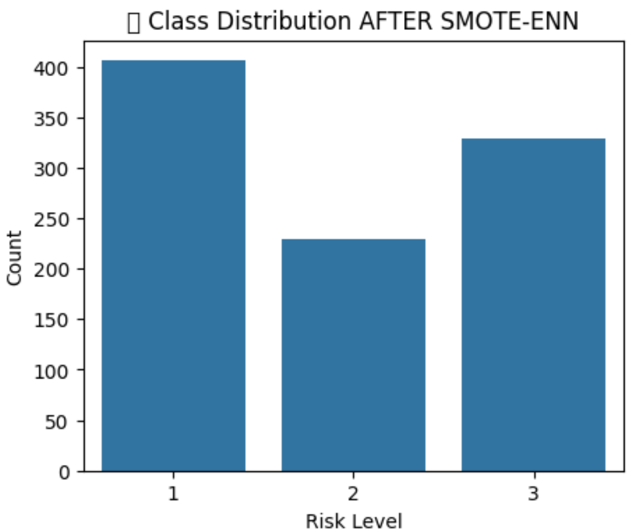

Fig. 3. Class distribution after SMOTE-ENN

3.4 Dimensionality Reduction Using PCA

Cutting down on redundancy of features was done through Principal Component Analysis (PCA) that enabled the calculations to be more efficiency [13,14]. The initial six features were conformed into five key components that retained over 95 percent of the explained variance. PCA aids in the removal of noise, faster model training, as well as preventing overfitting in small datasets.

3.5 Model Architecture

The proposed stacking ensemble architecture is a platform by which the predictions of three base learners are integrated, which include K-Nearest Neighbors (KNN), Random Forest (RF), and XGBoost. The ultimate decision is then determined by a meta-learner. The dataset passes through preprocessing before it can proceed to any further step, this preprocessing involves scaling the features, dimensional reduction done through PCA, and balance of the classes by use of SMOTE-ENN. Then each base learner is trained with the processed dataset to identify distinctive trends on the data. The probabilistic outputs of these base models are then combined by the meta-learner that is typically a Logistic Regression classifier to create the ones that are more general and less variable. Such an ensemble design through layers embraces the best of other algorithms and thus makes maternal health risk classification more strong and precise [7].

3.6 Evaluation Metrics

We used a number of different metrics to check how well the model worked [15].

- **Accuracy:** The percentage of correct classifications overall.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

- **Precision:** This evaluates how resistant the model is to false positives.

$$\text{Precision} = \frac{TP}{TP + FP} \tag{2}$$

- **Recall:** This evaluates how resistant it is to false negatives.

$$\text{Recall} = \frac{TP}{TP + FN} \tag{3}$$

- **F1-Score:** The F1-Score is the harmonic mean of recall and precision.

$$F1 = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \tag{4}$$

4 Discussion and results

4.1 Experimental Setup

The experiments used the Maternal Health Risk Dataset containing 1,200 samples with six physiological features. Data were split into 70% training and 30% testing using stratified sampling. Preprocessing included handling missing values, normalization, and target encoding. SMOTE-ENN addressed class imbalance, while PCA reduced dimensions while preserving over 95% variance [5,6,13].

4.2 Data Split Strategy and Validation Process

A 70–30 stratified train–test split was used to divide the dataset. This way, the original proportion of Low Risk, Mid Risk, and High Risk classes was kept in both subsets. This made sure that the evaluation was fair and reliable. We chose to give 70% of the data to training because the dataset is small and a larger training portion helps the model learn more stable decision boundaries. We used stratified 10-fold cross-validation on the training set, though, to avoid overfitting and make the model more generalizable. Each fold kept the class distribution the same and made sure that each sample was used once for validation and nine times for training. It is important to note that SMOTE-ENN was only used on the training folds, which kept data from leaking into the test set. Using both stratified splitting and cross-validation gives a more reliable and repeatable evaluation than just using one split.

4.3 Individual Model Performance

Table 2. Individual model performance

Model	Accuracy	Precision	Recall	F1-Score
KNN	86.55%	88.00%	87.00%	87.00%
Random Forest	92.07%	92.00%	91.00%	92.00%
XGBoost	96.21%	96.00%	96.00%	96.00%
Stacking Ensemble	97.59%	98.00%	98.00%	98.00%

Table 2 shows that the Stacking Ensemble operated the most accurately with 97.59% accuracy, as compared to other models of KNN and RF, and XGBoost in precision, recall, and F1-score. This validates the usefulness of the ensemble in correct prediction of maternal health risks [7,8,9].

4.4 10-Fold Cross-Validation Accuracy Comparison

Fig 3 shows that the performance of KNN models in terms of accuracy is high,similar to Random Forest and XGBoost due to fold-wise distribution for

10-fold CV as well as the Stacking Ensemble process. The Stacking Ensemble reaches the highest median accuracy and 8% less performancevariance over folds compared to the single models. Such low spread means the ensemble holds itsperformance, when the training and validation subsets are varied, hence can be interpreted as stronger generalization and lower sensitivity to fluctuations of sampling.

KNN and RandomForest, on the other hand, present wider variations of performance that seem to be directly influenced by the characteristics of each fold. XGBoost has less variability, butstill more than the Stacking Ensemble. These findings consolidate the idea that stacking multiple learners improve robustnessand establish a reliable predictive model for maternal health risk stratification.

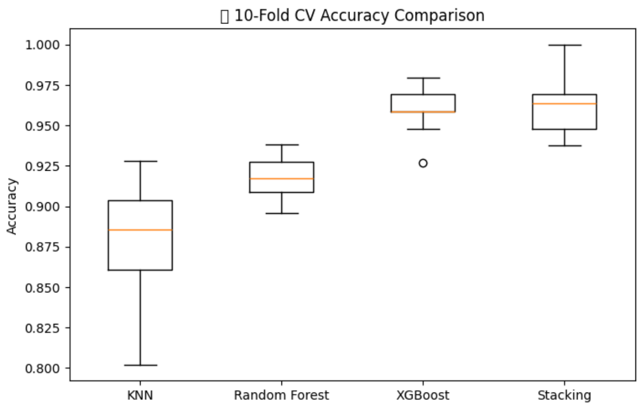

Fig. 4. 10-fold cross-validation accuracy comparison of all models.

4.5 Confusion Matrix Analysis

The confusion matrices present a more specific depiction of the performance of each model for the three risk categories (Low Risk, Mid Risk and High Risk), concerning their correct classificationas well as misclassification. One of the most noticeable observations is that all thesingle models have difficulty in differentiating Mid Risk the least represented. KNN and RandomForest learing models exhibit a confusion between the Mid Risk and High Risk classes demonstrating their sensitivity to overlapping decision boundaries effectuated by class imbalance. ei section ta r ektu shundhor kore likhe erokom r ekta pera baniye, total duita pera deo float subfigure

4.6 ROC Curve Analysis

Figure 5 shows the multi-class ROC curve for the Stacking Ensemble model. It shows that the model does a great job of classifying all three risk groups:

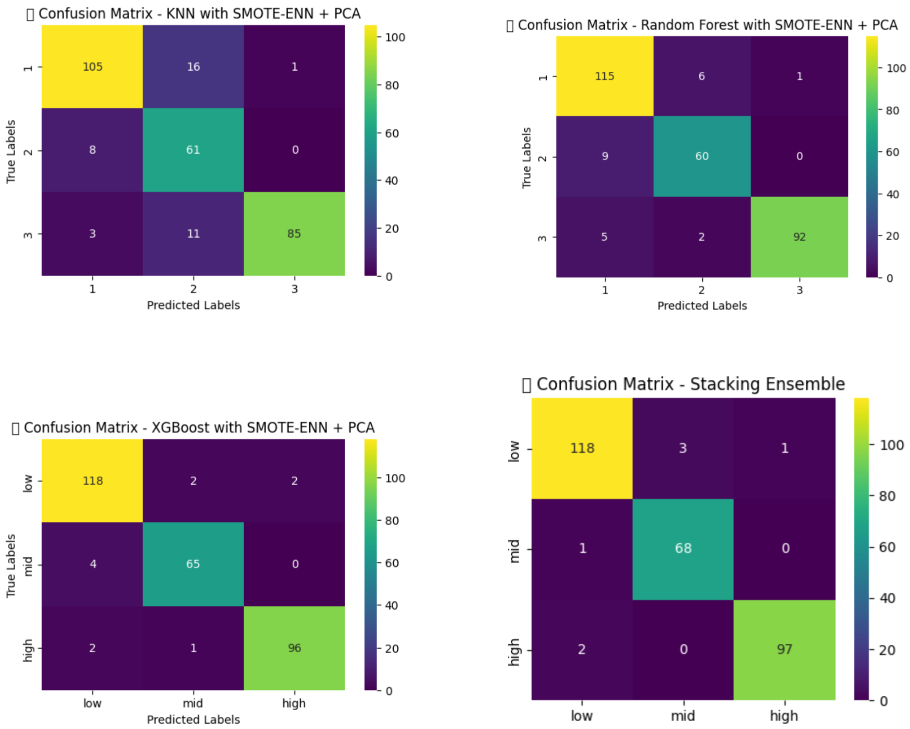

Fig. 5. Confusion matrices for all models (Random Forest, KNN, XGBoost, Stacking Ensemble

Low Risk, Mid Risk, and High Risk. The AUC values for Low Risk (0.99), Mid Risk (0.99), and High Risk (1.00) show that the model is very good at correctly classifying both the majority and minority classes, with a lot of space between each class. This shows that the ensemble method works well for dealing with class imbalance and making accurate predictions.

5 Discussion

This study's results show that combining PCA, SMOTE-ENN, and ensemble learning makes it much easier to predict maternal health risks than using just one model. The combined workflow cuts down on noise, fixes a big class imbalance, and makes the model better at finding nonlinear relationships between physiological indicators. The experimental results clearly show how each part of the pipeline makes the system more stable, accurate, and robust. There were especially big improvements in finding the underrepresented Mid Risk category.

5.1 Comparing Models

The Stacking Ensemble always beat KNN, Random Forest, and XGBoost on every test. This improvement is quite important for the Mid Risk group, where many previous studies have had trouble. The stacking method finds complicated patterns and lowers model variance better than single classifiers by using the varied strengths of several learners [3,8,9].

5.2 Effectiveness of Class Balancing

The SMOTE-ENN combination approach effectively mitigated thesevere class imbalance of the dataset. Synthetically constructing minority instances and deleting noisy majority samples, SMOTE-ENN enhanced the separation between classes and led to significant improvement of MidRisk recall. This two-level effect is more powerful compared to over-samplingor under-sampling methods, which frequently introduce artificial noise or get rid of informative samples [11,12].

5.3 Feature Reduction using PCA

The PCA decreased redundancy and retained more than 95% of the total variance. Not only did it speed up training, but also reduced overfitting potential—which was important with such a small dataset. Comprehensive PCA: PC A, where coupled with SMOTE-ENN can result in a cleaner and more discriminative feature space that improves the stability and consistency of ensemble model [13,15].

5.4 Practical Implications

With intensive examination, the proposed framework showsa great potential for decision support in maternal health risk diagnosis. By having the ability to detect more of these Mid Risk cases, this functionality allows doctors to screen populations and identify patients who may otherwise be missed in typical population screenings. Incorporation of this model into hospital Triage or community health based programs can assist healthcare services to prioritize high-risk individuals, undertake early interventions and minimize avoidable complications. This is particularly useful in resource-limited settings, which have limited available diagnostic tools and specialist access [2,7].

5.5 Constraints and Prospective Pathways

Despite its excellentperformance, the model was trained on a small dataset from a single source, indicating impaired generalization. The limited rangeof features also leads to the model's inability to capture behavioural, socio-demographic and environmental risk factors. Future works should validate theframework with larger scale, multi-center datasets; introduce more clinical variables; and apply explainable AI techniques aimed at better interpretability and clinician trust in models. Real-world implementation would also be beneficial to movethis approach beyond the laboratory benches [6,10,14].

6 Conclusion

This work contributes an overall sound and cohesive approach for predicting maternal health risk by incorporating PCA based dimensionality reduction with SMOTE-ENN classbalancing followed by a stacking ensemble classifier. The resulted pipeline shows significant improvements of predictionperformance with above all the identification of under-represented Mid Risk category, which jointly verifies that carefully noise reduction and hybrid resampling in conjunction with ensemble learning have potential to address both calibration-limited conditions (low-dimensionality/multiple imbalance).

Apart from its technical merit..., the framework also hassome non-trivial practical consequences. As it can facilitate the early and effective detection of high- and mid-risk pregnancies, themodel may help healthcare providers to identify patients who should be prioritized along with providing guidance for timely interventions and decision-making when resources are scarce.

The results are nonetheless limited given the source and nature ofthe data — as a single, restrictive dataset, with few features available. In thefuture, the proposed method can be extended to larger and many center datasets and more clinical or demographic variables in a multi-institutional manner and further combined with explainable AI approaches for better interpretability and adoption. Assessing the system in clinical settings is also necessaryto enable clinical translation of this research to practical maternal health support tools.

References

1. Lubano, K.: Bridging the gap: Implementation science, knowledge translation and innovation for advancing maternal health sustainable development goals. J. Obstet. Gynaecol. East Cent. Africa (2024). https://doi.org/10.59692/jogeca.v36i1.158
2. Ojong, S.A., Wamakima, B., Moyer, C.A., et al.: Maternal Health and Well-Being. Oxford Research Encyclopedia of Global Public Health. Oxford University Press (2023)
3. Lee, H.: Introducing maternal oral health as global health and public health agenda. Eur. J. Public Health 32 (2022)
4. Drife, J.O., Lewis, G., Neilson, J.P., et al.: International maternal health: Global action. Why mothers died and how their lives are saved (2023)
5. Deardorff, J., Tissue, M.M., Elliott, P., et al.: The critical value of maternal and child health (MCH) to graduate training in public health. Matern. Child Health J. 26, 121–128 (2022)
6. Sajedinejad, S., Majdzadeh, R., Vedadhir, A., et al.: Maternal mortality: A cross-sectional study in global health. Global Health 11, 4 (2015)
7. Mu, C., Yan, Z., Zhu, Y.: Prediction of maternal health risk based on physiological indicators. ACM Int. Conf. Proc. Series (2023)
8. Raza, A., Siddiqui, H.U.R., Munir, K., et al.: Ensemble learning-based feature engineering to analyze maternal health during pregnancy and health risk prediction. PLoS One 17, e0276525 (2022)
9. Yang, Y., Wu, N.: Gestational diabetes mellitus and preeclampsia: Correlation and influencing factors. Front. Cardiovasc. Med. 9, 831297 (2022)
10. Sandsæter, H.L., Horn, J., Rich-Edwards, J.W., et al.: Preeclampsia, gestational diabetes and later risk of cardiovascular disease. BMC Pregnancy Childbirth 19, 448 (2019)
11. Khanlou, N., Haque, N., Skinner, A., et al.: Maternal health among immigrant and refugee women in Canada: A scoping review. J. Pregnancy 2017, 8783294 (2017)
12. Phillimore, J.: Delivering maternity services in an era of superdiversity: Challenges of novelty and newness. Ethn. Racial Stud. 38, 568–582 (2014)
13. Arshad, H., Tayyab, M., Bilal, M., et al.: Trends and challenges in harnessing big data intelligence for healthcare transformation. In: Artificial Intelligence for Intelligent Systems: Fundamentals, Challenges, and Applications, pp. 220–240 (2024)
14. Basile, L.J., Carbonara, N., Pellegrino, R., et al.: Business intelligence in the healthcare industry: A data-driven approach to support clinical decision making. Technovation 120, 102482 (2023)
15. Zitnik, M., Nguyen, F., Wang, B., et al.: Machine learning for integrating data in biology and medicine. Inf. Fusion 50, 71–91 (2019)

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (http://creativecommons.org/licenses/by-nc/4.0/), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.