Smart Maternity: Machine Learning for Safer Prenatal Clinic

Dr.Jade Abuga
University of Eastern Africa, Baraton,
Kenya
jade@ueab.ac.ke

May 16, 2024

Abstract

Maternal health is a major public health concern because of its far-reaching implications for the well-being of both the mother and the child. Most maternal deaths can be prevented if there is a timely intervention that is offered to the mothers. It is therefore important to be able to predict if a mother is classified as being in high risk, low risk, and mid-risk to enable prompt attention to be given to the mother.
In this study, we are using Machine learning to train a maternal data set having seven attributes and divided into three categories, high-risk, mid-risk, and low-risk pregnancies. The main aim of this study is to develop and evaluate machine learning models for predicting maternal risk levels, categorized as high risk, mid-risk, and low risk, based on a dataset containing seven attributes related to maternal health.
The method involved training three Machine Learning algorithms, Logistic Regression, Random Forest and Support Vector Machine (SVM) using the dataset. The data had a significant difference in the categories thus, Synthetic Minority Over-sampling Technique (SMOTE) was used to address the class imbalance. Each algorithm was trained and evaluated on both the imbalanced and balanced datasets.
To train the model, the data was divided into the training and testing sets split into 80 and 20 percent for the train and test data respectively to evaluate the model's

performance on unseen data. The performance of the algorithms was compared based on their accuracy in predicting maternal risk levels. Additionally, the study assessed the effectiveness of each algorithm in predicting risk levels for randomly entered datasets.
The Random Forest achieved the highest accuracy of 85.71 and 81.77 percent for the balanced and imbalanced dataset respectively. Generally, algorithms trained with the smote-balanced dataset performed much better than with the imbalanced dataset. The risk level for a randomly entered dataset was predicted and Random Forest and Support Vector Machine predicted accurately.

Keywords

Logistic Regression, Support Vector Machine(SVM), Random Forest, SMOTE

Introduction

Maternal mortality is a major public health problem globally, with over 287,000 women dying each year worldwide from pregnancy or childbirth-related complications. Almost 95 percent of these deaths occurred in low and lower-middle-income countries, with most of them potentially preventable. Sub-Saharan Africa and Southern Asia accounted for 87 percent of the estimated global maternal deaths. In Kenya, more than 6000 maternal deaths, and 35,000 stillbirths occur each year [1].
The most common causes of maternal death are hemorrhage, infection, hypertensive disorders, obstructed labor, and unsafe abortion. Other factors that can contribute to maternal death both direct and indirect include eclampsia, obstructed labor, infections, malnutrition, age, parity, a lack of proper health care and preexisting medical conditions [2].
Machine learning has the potential to play a significant role in alleviating maternal deaths. By giving prenatal clinics a tool to track ans spot high-risk pregnancies, machine learning has the potential to reduce the number of maternal fatalities. The usage of this tool might be utilized to prioritize patients and to make sure that those who are risk receive the quickest and most suitable care by providing prenatal clinics a tool to monitor and identify high-risk pregnancies.
Maternal death according to World Health Organization (WHO) [3] is defined as the death occurring during pregnancy, at childbirth or within 42 days of termina-

tion of pregnancy, irrespective of the cause of death (obstetric and non-obstetric). This definition takes into account both unintentional/accidental and incidental causes. There are women who succumb to death due preventable causes related to pregnancy and childbirth annually [4]. The applications of Artificial Intelligence is emerging as a promising tool which can be used to address these challenges and improve maternal health outcomes. The vast datasets of clinical and demographic data can be analyzed using the ML algorithms, patterns identified and risks associated with maternal morbidity and mortality predicted. This information can be used to guide targeted interventions, reduce risk, and help in clinical decision-making.

Numerous studies have been done and they show the AI and ML has been used in the health care industry. Some of the examples in include: improving in hospital mortality prediction of diabetes using deep learning architecture [5], [6], predicting coronary artery disease from single photon emission computed tomography MPI using Deep Convolution Neural Networks [7], predicting the postoperative mortality and outcomes of EGS patients and predicting individual surgeons' risk [8], detecting CoronaVirus (Covid 19)[9], [10] and [11] among many other applications.

In maternal health applications, Assaduzzaman *et al.* [12] used machine learning algorithms to identify the maternal health risk factor. They used several machine learning algorithms including Cat Boost, Random Forest, XGB, Decision Tree, and Gradient Boost. They concluded the best ML algorithm was Random Forest an accuracy score, precision, recall all of 90 percent. Using these models, high risk pregnancies can be identified and necessary interventions undertaken to reduce risk. Machine learning has been used to predict the likelihood of either vaginal delivery or cesarean section by analyzing factors as maternal age, parity, and fetal presentation. The prediction can be used to make in decision making and in fostering the development of well-informed and tailored care strategies.

Marvin *et al.* [13] used AI to predict preterm births. They used Random Forest and KNeighbors and obtained an accuracy of 100 percent and 78 percent respectively with Synthetic Minority Oversampling Technique (SMOTE) and Adaptive Synthetic (ADASYN) class balancing techniques. They were able to provide useful automated data driven insights to maternal healthcare management stakeholders and policy makers for a sustainable healthcare system in developing countries. Satoshi *et al.* [14] did a study that established early prediction models of low-birth-weight reveal influential genetic and environmental factors. The AI-based models used genetic and environmental factors to determine the influential variables that influenced low birth weight. An early prediction model was developed and could be used to assess risk of low birth weight during pregnancy. Another study was done by Sahithi *et al.* [15] where they investigated the use of Logistic Regression, Random Forest, Adaboost, KNN, Catboost, XGboost, 1-D CNN and ANN to predict and analyze the challenges faced by women. Their results indicated that 1D-CNN based prediction models perfomed better with an accuracy of 99.53 percent when compared to the other approaches.

A systematic review was done by Munetoshi and Kazunori [16] on the prediction of pre-term using Artificial Intelligence where the state of AI research was elucidated and predictive values and accuracy clarified. Electrohysterogram images were mostly used, followed by the biological profiles, the metabolic panel in amniotic fluid or maternal blood, and the cervical images on the ultrasound examination and the conclusion of the review was that accuracy was better in the studies using the metabolic panel and electrohysterogram images.

Ayesha and Mim [17] used five ML methods to forecast five methods forecast maternal health risk. They employed the Xgboost, Decision Tree, Random Forest, Support Vector Machine, and Naive Bayes algorithms where it was shown that Xgboost, Decision Tree, and Random Forest Algorithm had a higher accuracy rate of 94 percent as compared to SVM accuracy of 72 percent, and Naive Bayes accuracy is 64 percent.

Apart from predicting risks and outcomes, Machine Learning can be used to look for pattern that are not obvious to the human eye which might lead to new insights as to what causes maternal mortality and morbidity. This in turn will guide decision making of preventative strategies and treatments as done by Aphinyanaphongs *et al.* [18].

In as much as AI and ML in maternal health have such encouraging prospects, there are some challenges such as potential bias [19] and interpretability [20]. But despite these challenges, ML possesses significant potential to enhance outcomes in maternal health. As machine learning algorithms advance in sophistication and interpretability, their incorporation into clinical settings is expected to grow and thus, in our study we are using ML for safer Prenatal clinic.

Methodology

0.1 Data Description

The maternal health risk data set that is being used in this study was obtained from Kaggle[21]. The Data was collected using IOT based risk monitoring system from different hospitals, community clinics, and maternal health cares. It contains columns on age in years, Systolic and Diastolic blood pressure, blood glucose levels, the normal resting heart rate and predicted risk level. The attributes and description of the attributes is summarized in table 1. There is a total of 1014 instances.

Name of the Variable	Data type	Description
Age	Integer	Age in years when a woman is pregnant.
SystolicBP	Integer	Upper value of Blood Pressure in mmHg.
DiastolicBP	Integer	Lower value of Blood Pressure in mmHg
BS	Float	Blood glucose levels is in terms of a molar concentration, mmol/L.
HeartRate	Integer	A normal resting heart rate in beats per minute.
Body Temp	Float	The woman's body temperature
Risk Level	Object	Predicted Risk Intensity Level.

Table 1: Table of Description of the attributes of the data set

Data Preprocessing

Data Cleaning and Analysis

Data cleaning involved checking the data for any errors, checking if it is incomplete or inconsistent due to errors or/and missing values and correcting the errors. The data set on being checked has shown that there is no missing data as shown in figure 1.

```
data.isnull().sum()
```

```
Age            0
SystolicBP     0
DiastolicBP    0
BS             0
BodyTemp       0
HeartRate      0
RiskLevel      0
dtype: int64
```

Figure 1: Figure of missing values

To further analyze the data, the histplots were obtained as shown in figure 2.

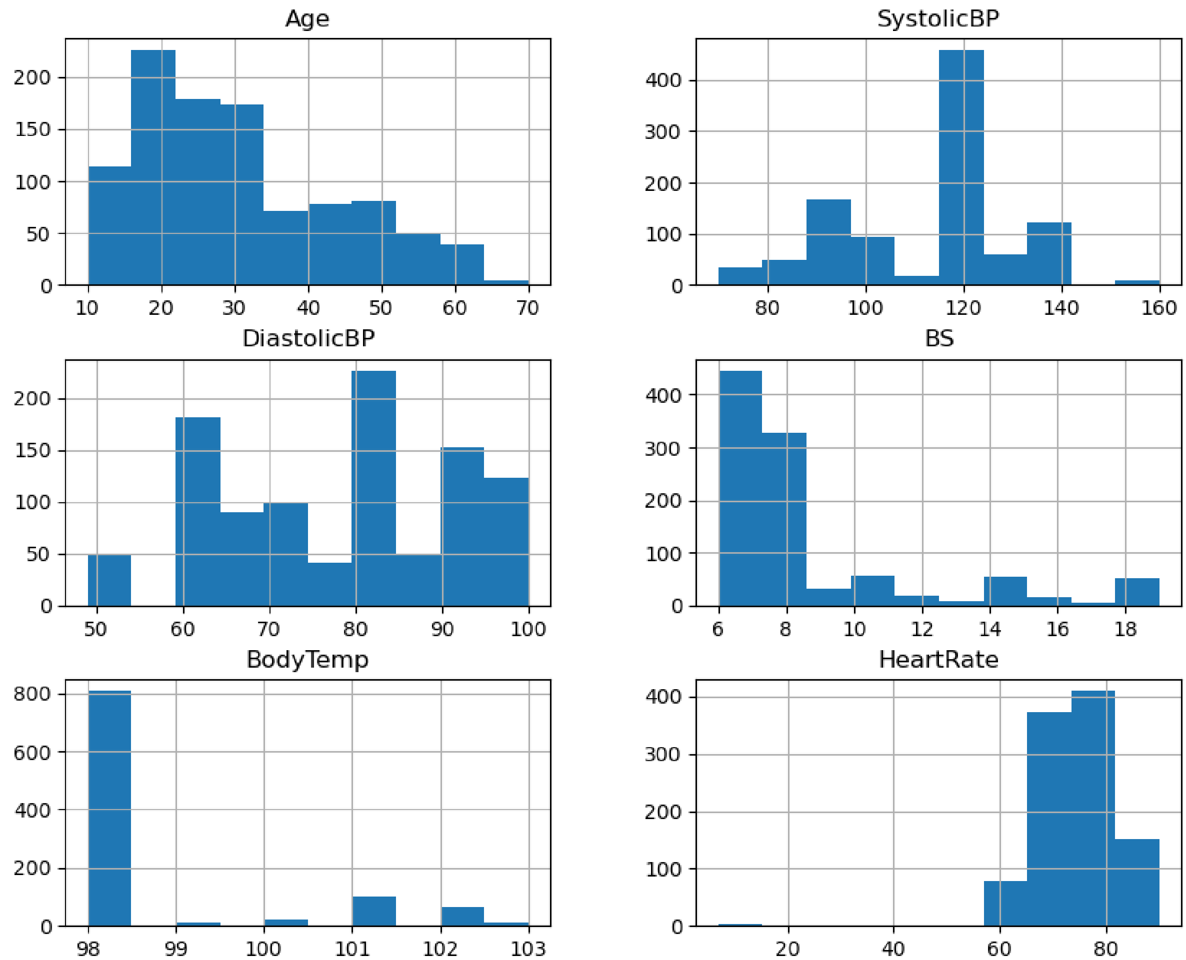


Figure 2: Histplot of the data set attributes

Further analysis was done to obtain the descriptive statistics as summarized in table 2.

	Age	SystolicBP	DiastolicBP	BS	BodyTemp	HeartRate
Count	1014.000	1014.000	1014.000	1014.000	1014.000	1014.000
Mean	29.871	113.198	76.460	8.726	98.665	74.301
Std	13.474	18.404	13.885	3.294	1.313	8.088
Min	10.000	70.000	49.000	6.000	98.000	7.000
25%	19.000	100.000	65.000	6.900	98.000	70.000
50%	26.000	120.000	80.000	7.500	98.000	76.000
75%	39.000	120.000	90.000	8.000	98.000	80.000
Max	70.000	160.000	100.000	19.000	103.000	90.000

Table 2: Table of descriptive statistics

From the table of the summary of statistics (table 2), it was seen that the minimum value for the heart rate is 7BPM. This is an unrealistic value for a human being since it is extremely low and is not compatible with life. It seems that there was an error when imputing the values since a typical value for heart rate for a human being at rest is between 60 to 100 beats per minute. To deal with this outlier value, we choose to remove the outliers which make for 0.2 percent of the data since it will not result to much loss of information.

Further analysis was done by plotting a correlation matrix and it was noted that the three variables with the highest correlation coefficient to the risk were the diastolic blood pressure, age and blood glucose level while heart rate had the lowest correlation coefficient of 0.11. This implies that the heart rate had very little impact on the risk level. Thus, the column of heart rate was excluded when training the data.

On further analysis, was found that 27 percent of the data set is categorized as high risk, 33 percent as mid risk and 40 percent as low risk. It was is clearly seen number of values in each classification risk level is significantly different. We chose to that deal with the imbalance by using Synthetic Minority Over-sampling Technique (SMOTE). To complete data preparation for data training, the data set categorical variables was turned into numerical variables using the encoder. High risk level was coded into 0, low riks level into 1 and mid-risk into 2.

Splitting the Data and Model Training

To train the model, the data was divided into the training and testing sets to evaluate the model's performance on unseen data. For our models, data was split into

80 and 20 percent for the train and test data respectively. The training data is used in the the machine learning model in order to predict the risk.

Results and Discussion

The whole dataset was split into the training and testing dataset using a ratio of 80:20 respectively for both the balanced and imbalanced. The performance of models is evaluated using both the imbalanced and balanced datasets. The performance of the results measured using Accuracy, F1- Score, Precision Score and Recall Score by using the confusion matrix and classification report.

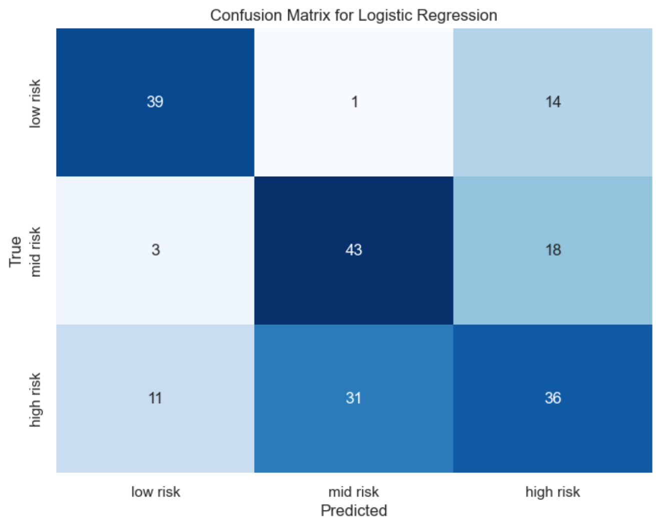
Logistic Regression

Logistic regression is a type of supervised learning algorithm that is used for binary classification problems that assumes a linear relationship between independent variables and the dependent variable. In this study, various strategies were employed for improving the model's accuracy. The strategies applied were Regularization, cross-validation and ensemble methods. However there was no improvement in the accuracy for the different strategies. For the balanced dataset, all the three methods had an accuracy of 60.2 percent and the same performance metrics score for the three risk categories. The performance metrics for Logistic Regression with regularization is shown figure 3.

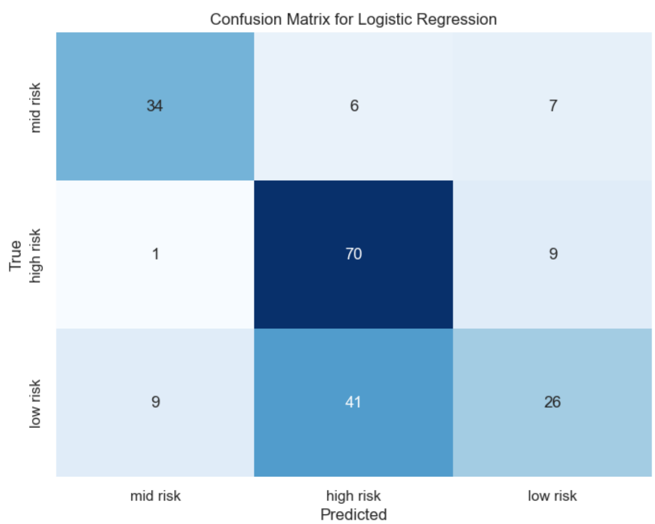
```
Accuracy for Logistic Regression: 60.20%
Classification Report for Logistic Regression:
              precision    recall  f1-score   support

   high risk       0.74      0.72      0.73        54
    low risk       0.57      0.67      0.62        64
    mid risk       0.53      0.46      0.49        78

    accuracy                           0.60       196
   macro avg       0.61      0.62      0.61       196
weighted avg       0.60      0.60      0.60       196
```

Figure 3: Performance Metrics for Logistic Regression for the balanced dataset

For the imbalanced data, the accuracy using different strategies varied. The Regularization had an accuracy score of 64.04 percent while the cross-validation linear regression had an accuracy score of 62.56 percent as shown in figure **??**. The confusion matrix for the both the balanced and imbalanced dataset for the logistic regression is plotted using the seaborn library as shown in Figure 4. From the confusion matrix in 4 a, it is clear that in the balanced dataset, only 11 cases of the high risk cases were predicted as low risk. This might have serious consequences in real cases situations. They might be overlooked and not get the timely medical assistance that they need leading to even loss of lives. The confusion matrix for the imbalanced train and test dataset of the logistic regression using regularization is shown in Figure 4 b. 9 of the high risk cases were predicted as low risk, an improvement from the 11 predicted when using balanced data. 41 of the low risk cases were predicted as high risk, a costly mistake with regards to time. The 41 will be assigned to emergency medical services yet they don't need.



(a) Balanced data



(b) Imbalanced data

Figure 4: Confusion Matrix for Logistic Regression for both the balanced and imbalanced dataset

Random Forest Classifier

It is is an ensemble learning method which means that multiple decision trees are built and merged together to get a more accurate and stable prediction. Random Forest classifier is not as prone to overfitting as the individual decision tree. The key hyperparameters used are the number of trees$(n-estimators)$, the maximum depth of each tree $(max - depth)$, the minimum number of samples required to be at a leaf node $min - samples - leaf$, and the minimum number of samples required to split an internal node $(min - samples - split)$. To get the best hyperparameters, we used hyperparameter tuning using grid search in $scikit - learn$. The code for that is shown below in figure (5)

```python
from sklearn.model_selection import GridSearchCV
param_grid = {
    'n_estimators': [50, 100, 200],
    'max_depth': [None, 10, 20],
    'min_samples_split': [2, 5, 10],
    'min_samples_leaf': [1, 2, 4]
}

# Initialize RandomForestClassifier
rf_model = RandomForestClassifier(random_state=42)

# Perform GridSearchCV
grid_search = GridSearchCV(estimator=rf_model, param_grid=param_grid, cv=5, scoring='accuracy')
grid_search.fit(X_train, y_train)

# Best parameters found
print("Best Parameters found by Grid Search:")
print(grid_search.best_params_)

# Get the best model
best_rf_model = grid_search.best_estimator_

# Predictions using the best model
best_predictions = best_rf_model.predict(X_test)

# Evaluate the performance of the best model
accuracy = accuracy_score(y_test, best_predictions)
print(f"Accuracy for the best Random Forest model: {accuracy:.2%}")
```

Figure 5: Code for Random Forest hyperparameter tuning

The results of the hyperparameter tuning shows that the besthyperparameter using grid search were: $max - depth$: None, $min - samples - leaf$: 1, $min - samples - split$: 2, $n - estimators$: 200. Using the said hyperparameters, the accuracy for the imbalanced dataset is found to be 81.77 percent (6 a), lower than for the balanced dataset at 85. 71 percent (6 b). The perfomance metrics for both the balanced and imbalanced dataset is shown in figure 6

Accuracy for Random Forest: 85.71%
Classification Report for Random Forest

	precision	recall	f1-score	support
high risk	0.90	0.96	0.93	54
low risk	0.85	0.83	0.84	64
mid risk	0.83	0.81	0.82	78
accuracy			0.86	196
macro avg	0.86	0.87	0.86	196
weighted avg	0.86	0.86	0.86	196

(a) Balanced data

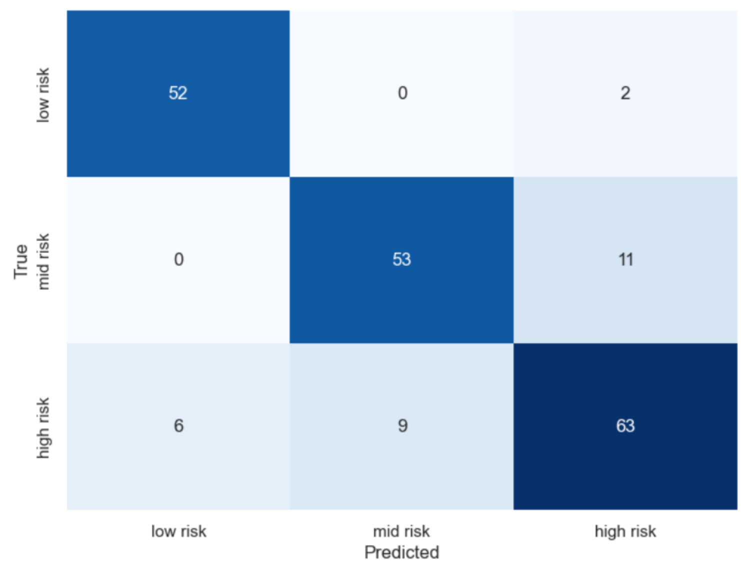
Accuracy for Random Forest: 81.77%
Classification Report for Random Forest

	precision	recall	f1-score	support
high risk	0.85	0.85	0.85	47
low risk	0.86	0.78	0.82	80
mid risk	0.76	0.84	0.80	76
accuracy			0.82	203
macro avg	0.82	0.82	0.82	203
weighted avg	0.82	0.82	0.82	203

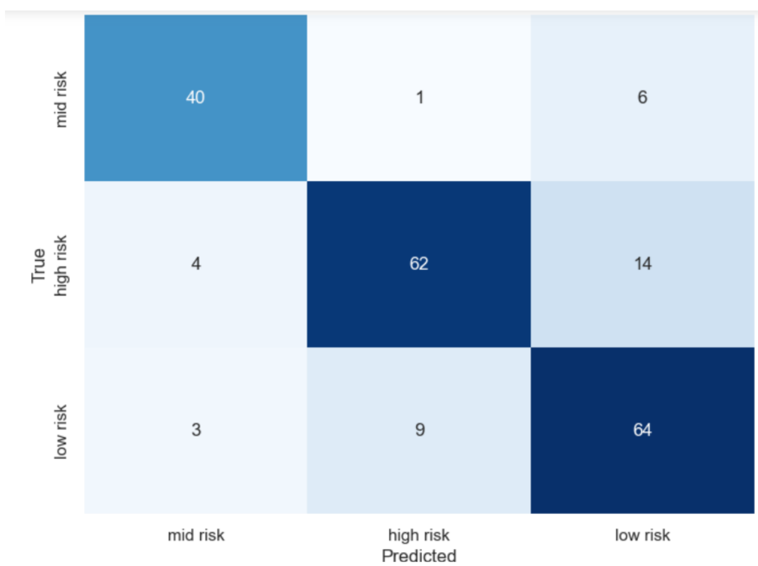
(b) Imbalanced data

Figure 6: Random Forest perfomance metrics for both the balanced and imbalanced dataset

The confusion matrix for both the balanced and imbalanced set is plotted in figure 7. For the balanced data, 6 of the high risk cases were predicted as low risk as shown in figure 7 a, compared to 11 of the imbalanced data set.



(a) Balanced data



(b) Imbalanced data

Figure 7: Random Forest Confusion Matrix for both the balanced and imbalanced dataset

Support Vector Machines

Support Vector Machine (SVM) is a supervised machine learning algorithm that is used for classification and regression tasks. We find the best hyper-parameters for the SVM by tuning. This was performed using GridSearchCV with the SVM model, parameter grid, 5-fold cross-validation, and accuracy as the scoring metric as shown in the code in figure 8. The best hyperparameter for regularization was $c = 100$, kernel is rbf and gamma is auto. Thee perfomance metrics for imbalanced and balanced data set is shown in figure 9

```
# Define the parameter grid
param_grid = {
    'C': [0.1, 1, 10, 100],  # Regularization parameter
    'kernel': ['linear', 'rbf', 'poly'],  # Kernel type
    'gamma': ['scale', 'auto']  # Kernel coefficient
}

# Instantiate the SVM model
svm_model = SVC()

# Perform Grid Search
grid_search = GridSearchCV(svm_model, param_grid, cv=5, scoring='accuracy')
grid_search.fit(X_train, y_train)

# Access the best hyperparameters
best_params = grid_search.best_params_
print("Best Hyperparameters:", best_params)
```

Figure 8: Code for Support Vector Machine hyperparameter tuning

```
Accuracy for SVM for balanced dataset: 82.65%
Accuracy: 82.65%

Support Vector Machines:
Classification Report:
              precision    recall  f1-score   support

   high risk       0.83      0.98      0.90        54
    low risk       0.77      0.80      0.78        64
    mid risk       0.88      0.74      0.81        78

    accuracy                           0.83       196
   macro avg       0.83      0.84      0.83       196
weighted avg       0.83      0.83      0.82       196
```

(a) Balanced data

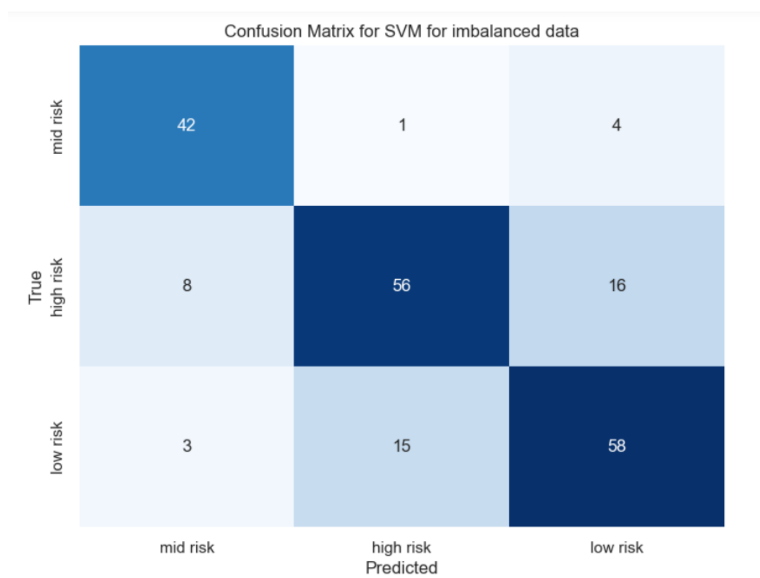
```
Accuracy for SVM for imbalanced dataset: 76.85%
Accuracy: 76.85%

Support Vector Machines:
Classification Report:
              precision    recall  f1-score   support

   high risk       0.79      0.89      0.84        47
    low risk       0.78      0.70      0.74        80
    mid risk       0.74      0.76      0.75        76

    accuracy                           0.77       203
   macro avg       0.77      0.79      0.78       203
weighted avg       0.77      0.77      0.77       203
```

(b) Imbalanced data

Figure 9: Support Vector Machine perfomance metrics

From Figure 10, the perfomance metrics is better for the balanced dataset at 82.65 percent as compared to the imbalanced dataset at 76.85 percent. The confusion matrix for the two dataset is presented in figure 10.

(a) Balanced data



(b) Imbalanced data

Figure 10: SVM confusion matrix for both the balanced and imbalanced dataset

From figure 10, it is clear that the balanced data performs better than the imbalanced data. Only 6 high cases were mislabeled as low risk as compared to 16 high cases for the imbalanced dataset. There were no low risk cases that were mislabeled high risk while 15 low risk cases were labeled as high risk in the balanced and imbalanced data respectively.

Predicting data set

A random data was entered for prediction as shown in table 3 and prediction done using the three machine learning algorithms. The age, Systolic BP, Diastolic BP, and body temperature was kept the same for the three patients. The BS was varied based what is considered as low, high and very high BS. For expectant mothers, blood glucose level of 7mmol/l and above is considered high and 15mmol/l and above is considered very high and dangerous. This is because ketone bodies may be produced which causes there to be an increased increased amounts of acid in the blood. This condition known as acidosis might be life-threatening for both mother and fetus. Thus, using our machine learning algorithms that was trained, we expect to see the patients with high and very high BS to be flagged as high risk.

Patient Data				
Age	**SystolicBP**	**DiastolicBP**	**BS**	**BodyTemp**
25	110	60	15	76
25	110	60	10	76
25	110	60	5.9	76

Table 3: Patients' data

The prediction output is given in figure 11.

We expect that patient 1 and 2 be flagged as high risk since we used BS values above the normal range. The predictions using random forest and SVM gave an accurate risk level for all the three patients. The Logistic regression only got one prediction correct.

Discussion and Conclusion

In conclusion, we used three machine learning algorithms for maternal health risk classification. The data was resampled using smote technique. The machine was trained on both the balanced and the imbalanced dataset. The algorithm that

```
# Predict the risk level for the patient using Logistic Regression
predicted_risk_level_P1 = LR.predict(patient_data1)
predicted_risk_level_P2 = LR.predict(patient_data2)
predicted_risk_level_P3 = LR.predict(patient_data3)
print("Predicted Risk Level for Patient 1 using Logistic Regression:", predicted_risk_level_P1)
print("Predicted Risk Level for Patient 2 using Logistic Regression:", predicted_risk_level_P2)
print("Predicted Risk Level for Patient 3 using Logistic Regression:", predicted_risk_level_P3)
```

```
Predicted Risk Level for Patient 1 using Logistic Regression: ['low risk']
Predicted Risk Level for Patient 2 using Logistic Regression: ['low risk']
Predicted Risk Level for Patient 3 using Logistic Regression: ['low risk']
```

```
# Predict the risk level for the patient using Random Forest
predicted_risk_level_P1 = rf_model.predict(patient_data1)
predicted_risk_level_P2 = rf_model.predict(patient_data2)
predicted_risk_level_P3 = rf_model.predict(patient_data3)
print("Predicted Risk Level for Patient 1 using Random Forest:", predicted_risk_level_P1)
print("Predicted Risk Level for Patient 2 using Random Forest:", predicted_risk_level_P2)
print("Predicted Risk Level for Patient 3 using Random Forest:", predicted_risk_level_P3)
```

```
Predicted Risk Level for Patient 1 using Random Forest: ['high risk']
Predicted Risk Level for Patient 2 using Random Forest: ['high risk']
Predicted Risk Level for Patient 3 using Random Forest: ['low risk']
```

```
# Predict the risk level for the patient using SVM
predicted_risk_level_P1 = svm_model.predict(patient_data1)
predicted_risk_level_P2 = svm_model.predict(patient_data2)

print("Predicted Risk Level for Patient 1 using Random Forest:", predicted_risk_level_P1)
print("Predicted Risk Level for Patient 2 using Random Forest:", predicted_risk_level_P2)
print("Predicted Risk Level for Patient 3 using Random Forest:", predicted_risk_level_P3)
```

```
Predicted Risk Level for Patient 1 using Random Forest: ['high risk']
Predicted Risk Level for Patient 2 using Random Forest: ['high risk']
Predicted Risk Level for Patient 3 using Random Forest: ['low risk']
```

Figure 11: Prediction outcome for Logistic Regression, Random Forest and Support Vector Machine

achieved highest performance in terms of accuracy, recall, precision and F1-score is the random forest with an accuracy of 85.71 and 81.77 percent for the balanced and imbalanced dataset respectively. Generally, algorithms trained with the smote-balanced dataset performed much better than with the imbalanced dataset. It should be noted that there were instances where the dataset were mislabeled. This is risky especially for cases of high risk being labeled as low risk. More improvement should be done to the algorithms to increase accuracy, precision, recall and the F1-score.

Recommendations

The different strategies for Logistic Regression had very little effect in improving the accuracy of the model. This could be attributed to the small size of the dataset.

Future research should aim to collect more data and test the how well the different models will perform on data collected from different geographical locations.
Future research should focus on linking wearable devices like smartwatches to these algorithms such that the device will collect data continuously. The data will be monitored and will give a warning to the pregnant women if the risk level is high depending on the data that has been collected. This will ensure that the high risk women go to get treatment immediately to reduce maternal mortality. Future research should also find ways to improve the accuracy. It is costly high risk expectant mother is labeled low risk since the effects are adverse and could even lead to maternal mortality. The goal should be to make the algorithms as accurate as possible to avoid maternal deaths.

References

[1] Marleen Timberman Evaline Langat, Lillian Mwanri. Effects of implementing free maternity service policy in kenya: An interrupted time series analysis. *BMC Health Services Research*, 2019.

[2] Kagia JW. Maternal mortality causality a kenyan experience. *Linacre Q*, 2011.

[3] Maternal mortality. retrieved from: https://www.who.int/data/gho/indicator-metadata-registry/imr-details/4622: :text=a

[4] Maternal mortality. retrieved from: https://www.who.int/news-room/fact-sheets/detail/maternal-mortality.

[5] Julian Theis, William L. Galanter, Andrew D. Boyd, and Houshang Darabi. Improving the in-hospital mortality prediction of diabetes icu patients using a process mining/deep learning architecture. *IEEE Journal of Biomedical and Health Informatics*, 26(1):388–399, 2022.

[6] Shane D. Sykes, Eugenie R. Lumbers, Kirsty G. Pringle, Tamas Zakar, Gustaaf A. Dekker, and Claire T. Roberts. Predicting gestational diabetes from maternal angiotensin ii and angiotensin 1–7 levels at 15 weeks gestation. *Reproduction, Fertility and Development*, 22:106–106, 2010.

[7] Commandeur F. Motlagh M. Sharir T. Einstein A. J. Bokhari S. Fish M. B. Ruddy T. D. Kaufmann P. Sinusas A. J. Miller E. J. Bateman T. M. Dorbala S. Di Carli M. Germano G. Otaki Y. Tamarappoo B. K. Dey D. Berman

D. S. Betancur, J. and P. J. Slomka. Deep learning for prediction of obstructive disease from fast myocardial perfusion spect: A multicenter study. *Cardiovascular imaging*, 2018.

[8] Mohamad El Moheb, Anthony Gebran, Lydia R. Maurer, Leon Naar, Majed W. El Hechi, Kerry Breen, Ander Dorken-Gallastegi, Robert D. Sinyard, Dimitris Bertsimas, George Velmahos, and Haytham M.A. Kaafarani. Artificial intelligence versus surgeon gestalt in predicting risk of emergency general surgery. *Journal of Trauma and Acute Care Surgery*, 95:565 – 572, 2023.

[9] Harsh Panwar, P.K. Gupta, Mohammad Khubeb Siddiqui, Ruben Morales-Menendez, and Vaishnavi Singha. Application of deep learning for fast detection of covid-19 in x-rays using ncovnet. *Chaos Solitons Fractals*, 2020.

[10] Mojtaba Mohammadpoor, Mehran Sheikhi Karizaki, and Mina Sheikhi Karizaki. A deep learning algorithm to detect coronavirus (covid-19) disease using ct images. *PeerJ. Computer science*, 7:e345, 2021.

[11] Mojtaba Mohammadpoor, Mehran Sheikhi Karizaki, and Mina Sheikhi Karizaki. A deep learning algorithm to detect coronavirus (covid-19) disease using ct images. *PeerJ. Computer science*, 7:e345, 2021.

[12] Md Assaduzzaman, Abdullah Al Mamun, and Md Zahid Hasan. Early prediction of maternal health risk factors using machine learning techniques. In *2023 International Conference for Advancement in Technology (ICONAT)*, pages 1–6, 2023.

[13] Ggaliwango Marvin, Joyce Nakatumba-Nabende, Nakayiza Hellen, and Md. Golam Rabiul Alam. Responsible artificial intelligence for preterm birth prediction in vulnerable populations. *2022 IEEE Asia-Pacific Conference on Computer Science and Data Engineering (CSDE)*, pages 1–6, 2022.

[14] Satoshi Mizuno, Satoshi Nagaie, Gen Tamiya, Shinichi Kuriyama, Taku Obara, Mami Ishikuro, Hiroshi Tanaka, Kengo Kinoshita, Junichi Sugawara, Masayuki Yamamoto, Nobuo Yaegashi, and Soichi Ogishima. Establishment of the early prediction models of low-birth-weight reveals influential genetic and environmental factors: a prospective cohort study. *BMC Pregnancy and Childbirth*, 23, 2023.

[15] P. Sahithi, S. Amulya, A. Sagar Gajapathi, S. Vishnu Vardhan Raju, K. Rukmini, and Dr. S. Vishnu Murthy. Deep learning based risk level prediction model for maternal mortality. *International Journal of Engineering Research and Applications*, 13:70–80, 2023.

[16] Munetoshi Akazawa and Kazunori Hashimoto. Prediction of preterm birth using artificial intelligence: a systematic review. *Journal of Obstetrics and Gynaecology*, 42(6):1662–1668, 2022. PMID: 35642608.

[17] Ayesha Siddika and Mim Sultana. Maternal health risk analysis by using exploratory data analysis and machine learning algorithms. 10:146, September 2023.

[18] Y. Aphinyanaphongs, S. Bossyut, T. Theeramunkong, and T Wilaiprasitporn. Explainable machine learning models to analyse maternal health. *Data in Brief*, 45, 2022.

[19] A. Rajkomar, M. Hardt, M. D. Howell, G. Corrado, and M. H Chin. Ensuring fairness in machine learning to advance health equity. *Annals of Internal Medicine*, 169:866–872, 2018.

[20] Rudin C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence*, 1(15):206–215, 2019.

[21]