Maternal Health Risk.

By Hilary Henshaw, Emmanuel Adeoti and Carolyn Bollerman

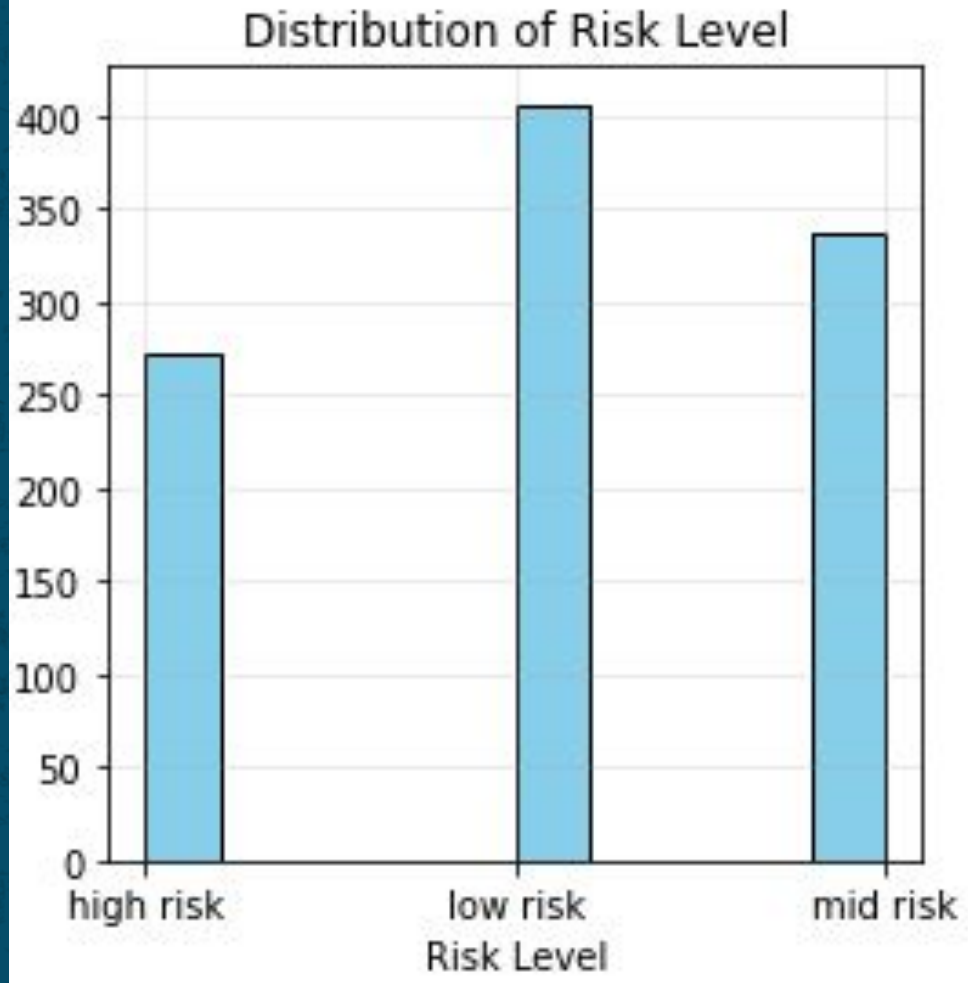
Introduction

Maternal health refers to the physical and mental well-being of women during and after pregnancy. It includes comprehensive health care from preconception through postpartum.

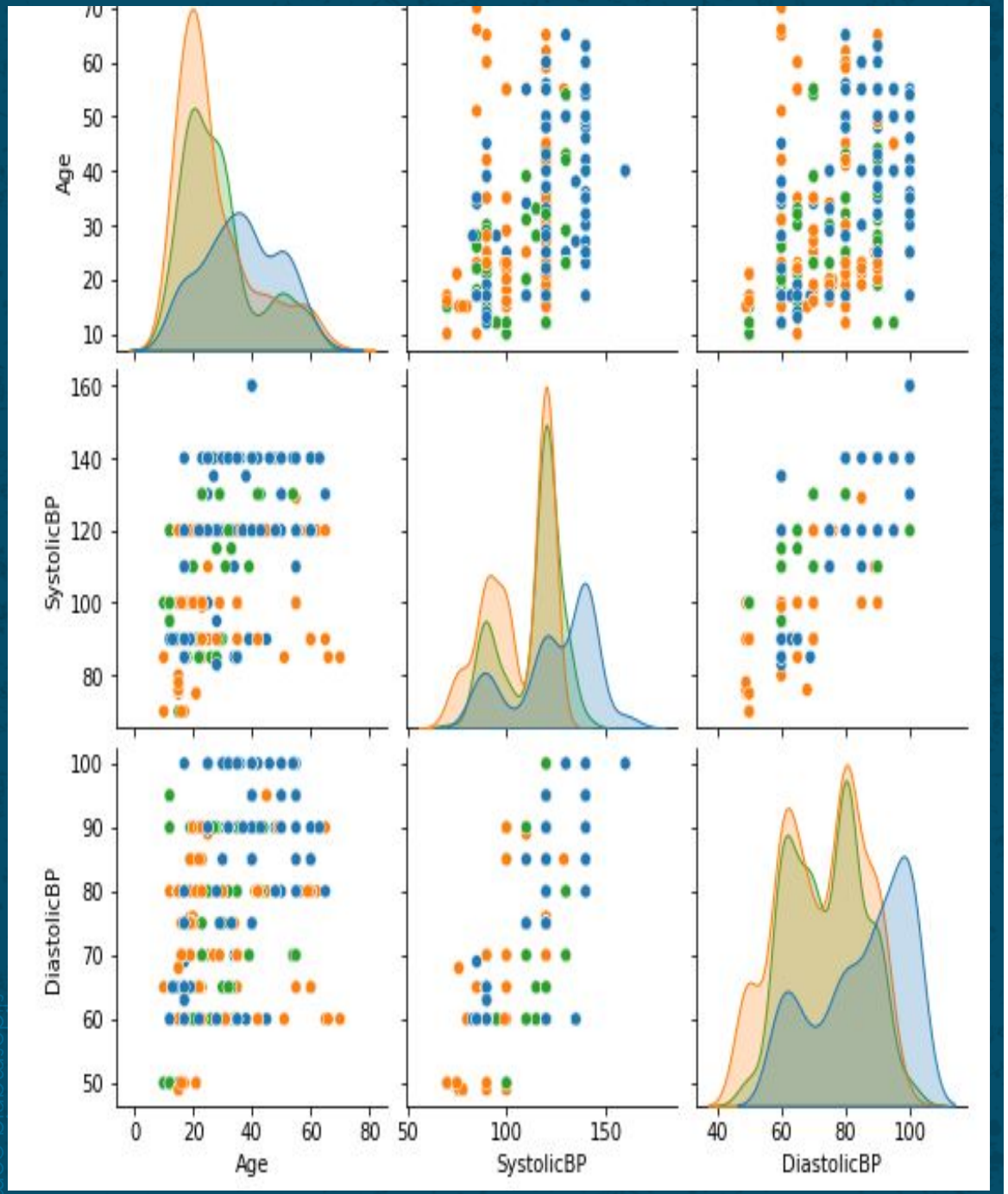
Understanding the health factors of pregnant women enables us to identify signs that predict pregnancy risks and to develop effective strategies to support both mothers and their babies.

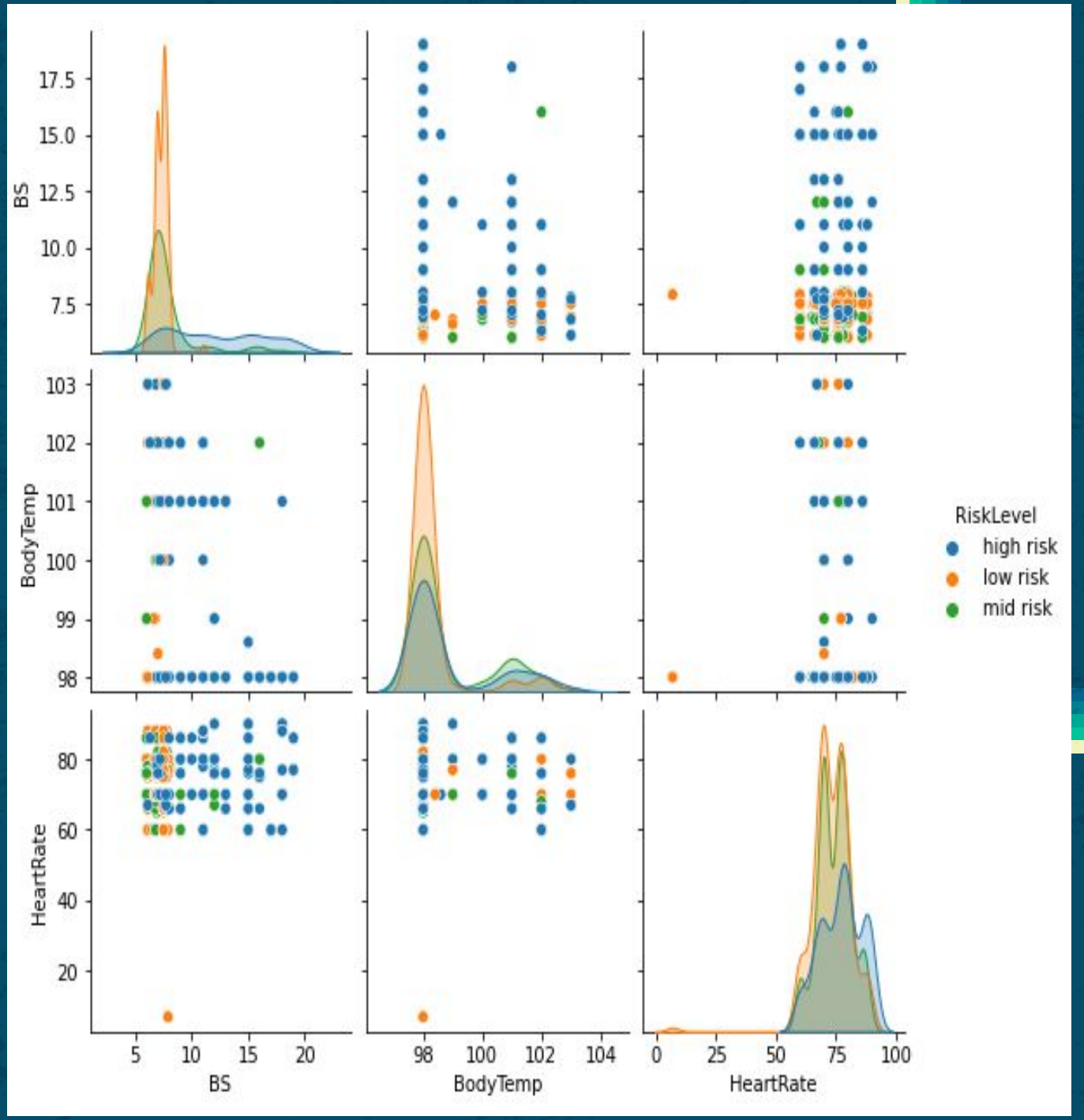
Research question: In what ways can health factors be used to predict maternal survival during pregnancy?

- **Dataset:** Maternal Health Risk from the UCI Machine Learning Repository.
- Real-time health monitoring of pregnant women using wearable devices that alert them and their families about critical health conditions.
- The dataset contains 1014 instances with 7 attributes, including a class attribute (low, medium and high risk).
- The study also incorporates data collected via IoT-enabled devices from select pregnant women in Bangladesh, and a comparison between the existing dataset and real-world data demonstrates consistent risk prediction results.



	Age	SystolicBP	DiastolicBP	BS	BodyTemp	HeartRate
count	1014.000000	1014.000000	1014.000000	1014.000000	1014.000000	1014.000000
mean	29.871795	113.198225	76.460552	8.725986	98.665089	74.301775
std	13.474386	18.403913	13.885796	3.293532	1.371384	8.088702
min	10.000000	70.000000	49.000000	6.000000	98.000000	7.000000
25%	19.000000	100.000000	65.000000	6.900000	98.000000	70.000000
50%	26.000000	120.000000	80.000000	7.500000	98.000000	76.000000
75%	39.000000	120.000000	90.000000	8.000000	98.000000	80.000000
max	70.000000	160.000000	100.000000	19.000000	103.000000	90.000000





Models and Prediction Metrics

Models (classifiers)	Training Accuracy	Test Accuracy	Precision (Test)	Recall (Test)	F1-score (Test)
NN (deep learning)	66.5	67.5	69	66	67
Decision Tree	93	84	85	84	85
Random Forest {'max_depth': 7, 'min_samples_split': 2, 'n_estimators': 100}	78	72	71	72	71
QDA	64	68	70	68	66
NB (Naive Bayes)	61	60	58	60	54
KNN (knn= 1)	81	82	83	82	82
Stacking (final =lda) (base= DT, KNN, RF)	93	82.3	82	81	81
Stacking (final = SVM) (base= DT, KNN, NN)	92.2	83	84	84	84

MAIN FINDINGS! Top 3 models

- Stacking classifier with LDA as the final estimator demonstrates a strong performance in predicting classes 0 and 2, with room for improvement in predicting class 1. 82.3% of instances in the test set are correctly classified.

```
Confusion Matrix using predict:
[[69 12  0]
 [15 51  1]
 [ 6  4 45]]
Confusion Matrix assuming perfection:
[[81  0  0]
 [ 0 67  0]
 [ 0  0 55]]
```

- Stacking classifier with SVC as the final estimator demonstrates a strong performance in predicting classes 0 and 2, with room for improvement in predicting class 1. 83% of instances in the test set are correctly classified.

```
Confusion Matrix using predict:
[[69 12  0]
 [13 51  3]
 [ 1  3 51]]
Confusion Matrix assuming perfection:
[[81  0  0]
 [ 0 67  0]
 [ 0  0 55]]
```

- Decision Tree Classifier demonstrates a strong performance in predicting classes 0 and 2. Has the highest percentage of TP in predicting class 1. 84 % of instances in the test set are correctly classified.

```
Confusion Matrix using predict:
[[64 17  0]
 [ 4 57  6]
 [ 1  4 50]]
Confusion Matrix assuming perfection:
[[81  0  0]
 [ 0 67  0]
 [ 0  0 55]]
```

Surprisingly Neural Network did not do very well for this dataset.

Progression: Over the 10 epochs, both training and validation accuracy improve, with final training accuracy reaching 52.15% and validation accuracy at 50.25%.

Test Performance: The test accuracy of 50.74% indicates that the model performs slightly better than random guessing (which would be around 33.33% for three classes). The cross-entropy loss of 1.0017 suggests that while the model is learning, there is still significant room for improvement.

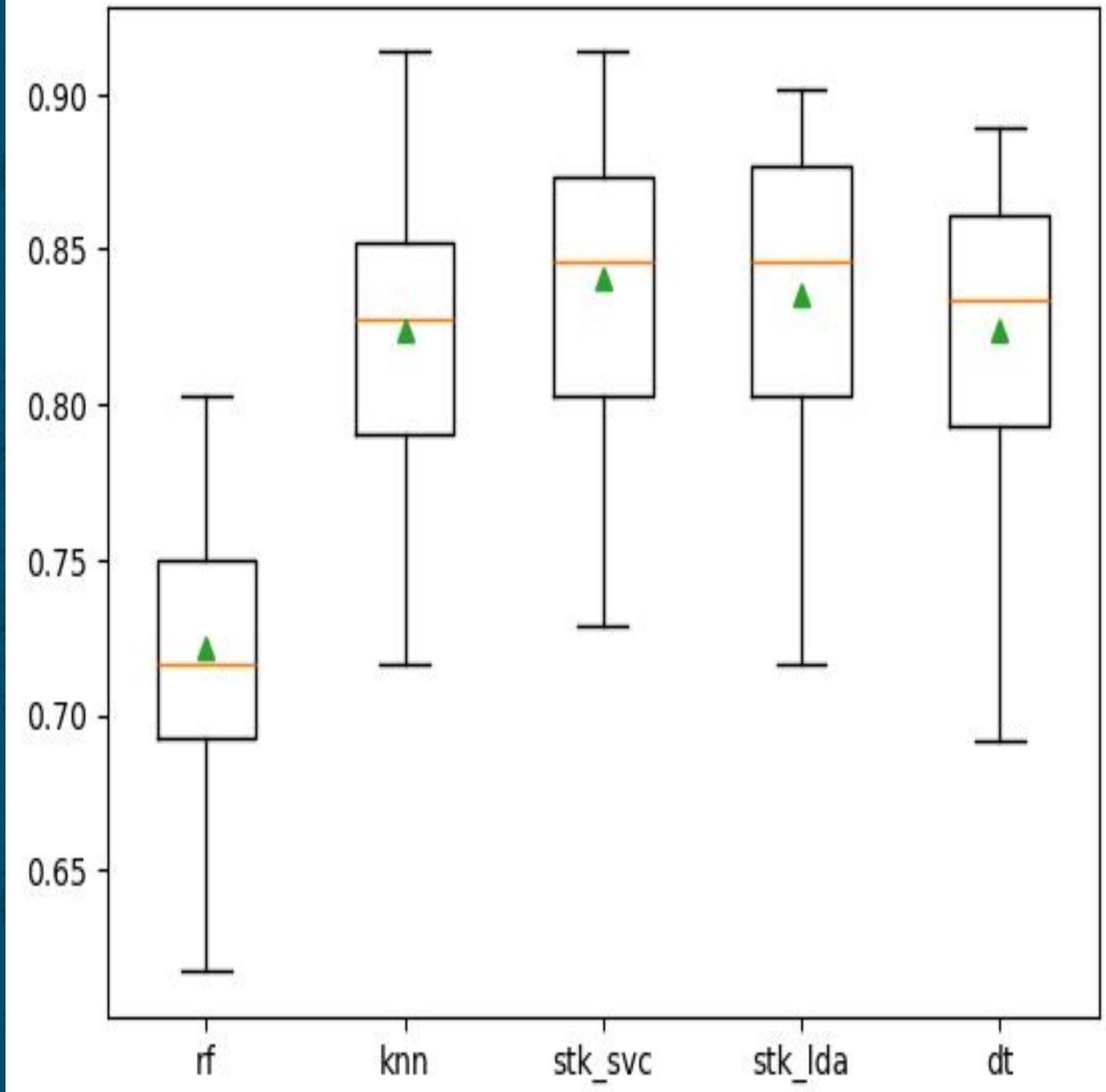
Progression: Over the 20 epochs, both training and validation accuracy slightly improved, with final training accuracy reaching 66.5% and validation accuracy at 68.5%. This shows that the model is learning, but perhaps more epochs or enhanced hyperparameters might further improve performance.

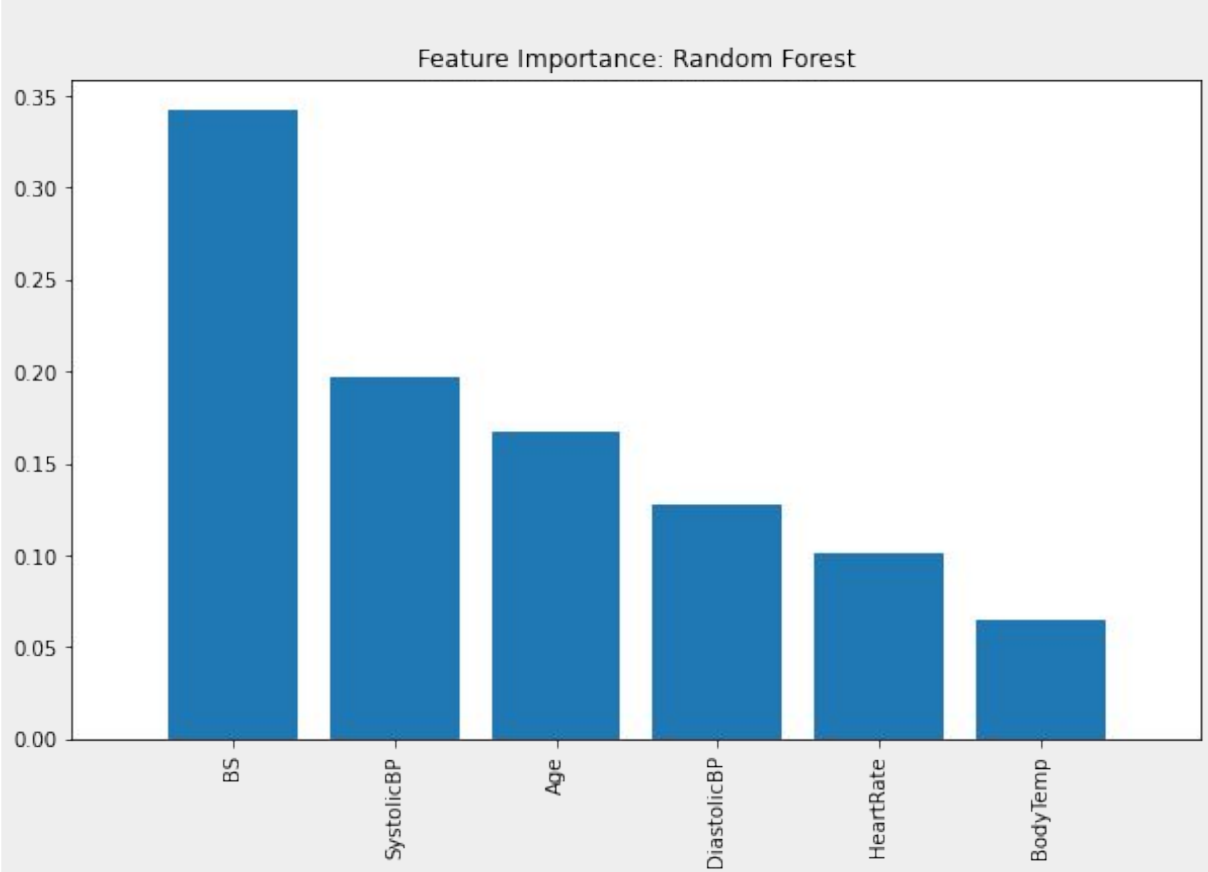
Test Performance: The test accuracy of 67.00% Higher than NB and QDA. The cross-entropy loss of 0.7366 suggests that while the model is learning, there is still significant room for improvement.

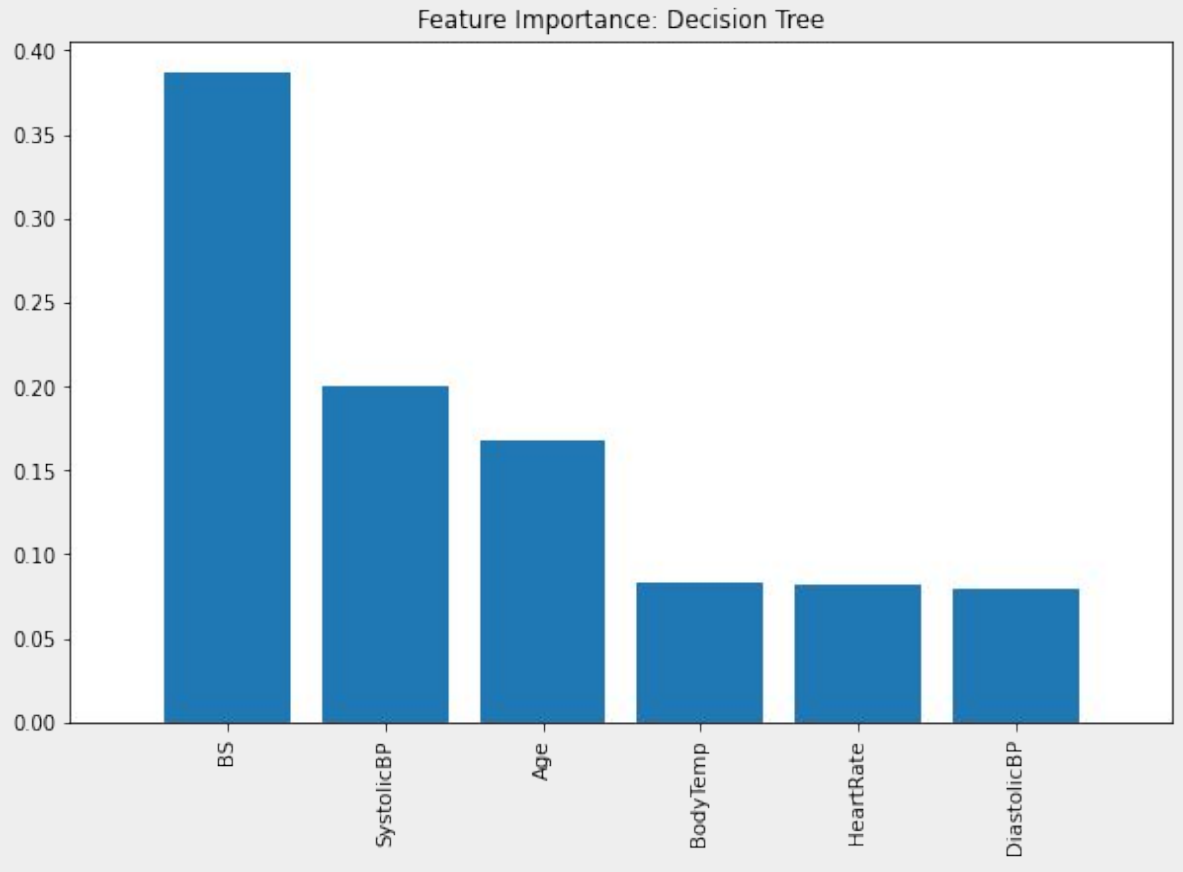
We compare the models that produced the best training and test accuracy with a box plot displaying their accuracy and error distribution.

The box plot visualization suggests that Stacking with SVC has a more stable performance compared to the other models and is highly effective in classifying maternal health risks.

All models have the same mean score (0.824) and the same standard deviation (0.047), suggesting similar performance on the dataset.







- This visual shows how much each variable contributes to the top performer model's decision-making process.
- In identifying the most important features, healthcare practitioners could focus on monitoring those variables more closely in expectant mothers to better assess and manage maternal health risks.

Conclusion

- **Decision Tree**: Base model with no hyperparameter tuning. in particular, with its tree approach, provided robust metrics and balanced precision-recall scores. It did okay in handling the complexities and variances in the data.
- **Stacking with SVC**: Visually depicted better consistency in box plots, suggesting stable performance, although quantitative metrics were on par with other models.
- **Neural Networks**: Show potential but require further tuning, more epochs, or possibly more complex architectures to achieve better performance.

Implications for Maternal Health Prediction

The models provide a framework for identifying maternal health risks using various health factors. By leveraging the strengths of stacking and random forest models:

-**Risk Prediction**: Effective prediction of maternal health risks can lead to timely interventions and better healthcare strategies.

-**Decision Support**: Healthcare providers can use these models to support decision-making processes, improving maternal and fetal health outcomes.

-**Future Improvements**: A Global data, Continuous model refinement, including deeper neural networks and more sophisticated stacking methods, could further enhance prediction accuracy.