

# PRETERM BIRTH PREDICTION OF PREGNANT WOMEN IN POST-CONIZATION PERIOD USING MACHINE LEARNING TECHNIQUES

## Authors



Mian Ahmed Jamiul Bari  
Mohammad Intiaz Faisal  
Mahmud Hasan  
Labiba Islam  
Md. Sabbir Hossain  
Dr. Sifat Momen



**Department of Electrical and Computer Engineering, North South University,  
Bangladesh.**

## TABLE OF CONTENTS

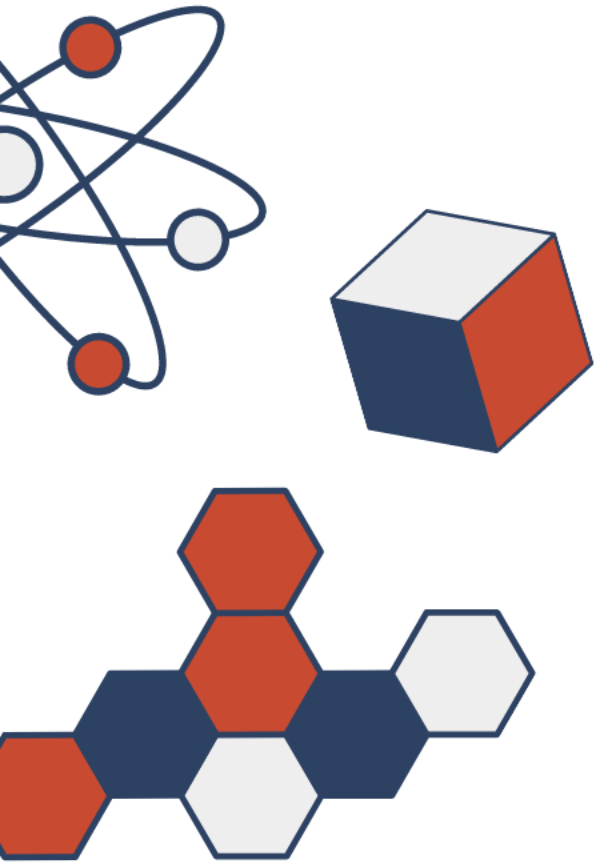
I	WHAT IS PRETERM BIRTH	3
II	EFFECTS OF CONIZATION ON PRETERM BIRTH	4
III	RELATED WORKS	5
IV	METHODOLOGY	6
V	CLASSIFIERS USED AND HYPER-PARAMETER TUNING	8
VI	RESULT	10
VII	CONCLUSION	12

A baby born without completing a full 40 weeks of gestation period (born before 37 weeks) is considered to be premature baby and the term of the delivery is called preterm birth.

The severity of preterm birth depends on how early the delivery occurs.



**According to World Health Organization approximately 15 million preterm deliveries happen every year. About 1 million babies die due to the consequences of preterm birth.**



Women who have been diagnosed with Cervical Intra epithelial Neoplasia (a pre-cancerous condition) undergo conization surgeries. These conization procedures leave a negative impact on a woman's pregnancy and greatly increases the risk of preterm birth.

As a result, pregnant women who underwent excisional surgeries (conization) for cervical intraepithelial neoplasia (CIN) display high risks of preterm birth.

Therefore, it is crucial to predict the risks of preterm birth amongst women in their post conization periods. **We applied machine learning approaches to correctly predict preterm birth delivery of women in their post conization.**



### Nam et. al

conducted studies of Korean patients who had cervical conization and delivered singleton infants. Their observation included type of excisional surgery performed and second trimester cervical length at different gestational periods.



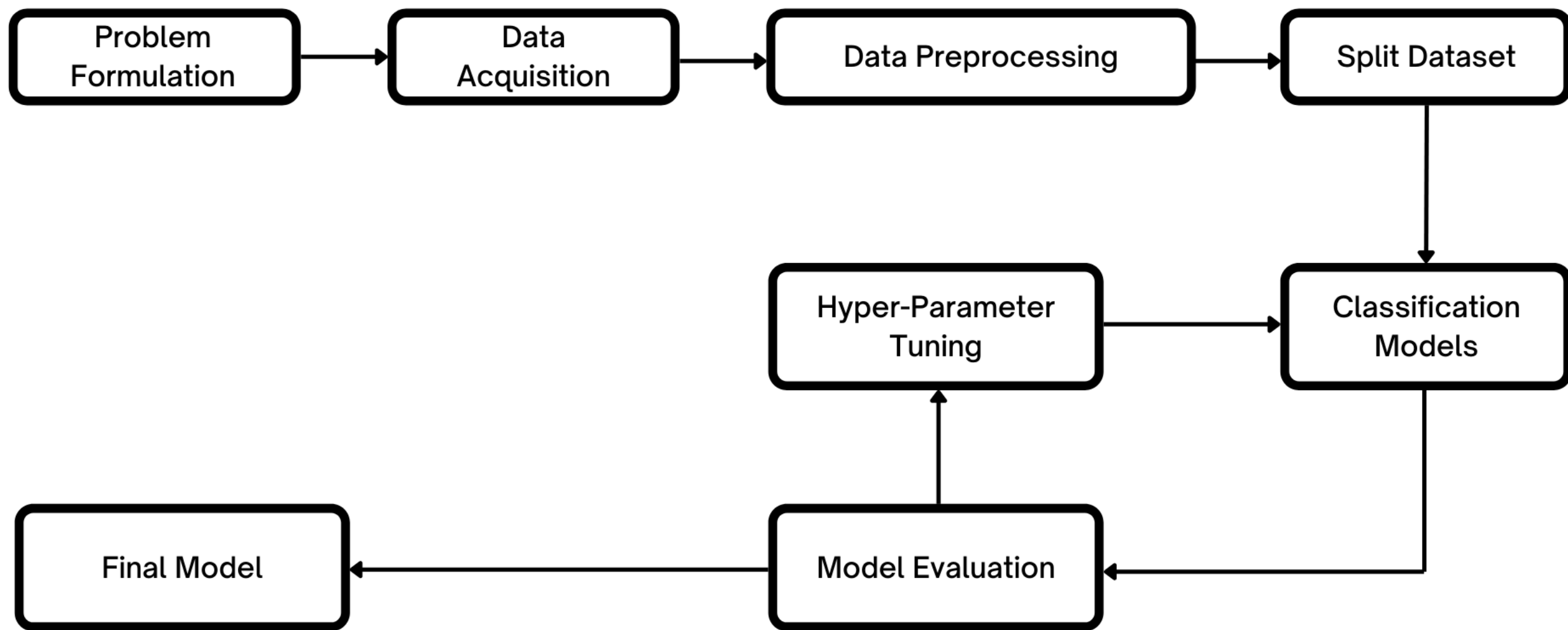
### Kindinger et. al

supervised a retrospective analysis in clinics that specializes in caring pregnant women of post conization phase. Their research discovered that the risk of preterm birth can be reduced with a targeted cervical cerclage.



### Crane et. al

conducted a prospective cohort study in a healthcare center which included women undergone surgical treatments for CIN such as CKC, LEEP, or cryotherapy earlier and with singleton pregnancies. Based on their research, LEEP and CKC lead to spontaneous preterm delivery. A cutoff of <3.0 cm was found the best for predicting spontaneous preterm delivery in LEEP.



## Data Acquisition:

- Dataset Details
  - Source: Dryad
  - Instance: 730, Features: 7
  - Missing Values: Yes
  - Data Type: Nominal, Numerical
- Features
  - age
  - ethnicity
  - cervical length during weeks 13 -15
  - cervical length during weeks 16 -18
  - cervical length during weeks 20 -22
  - gestation at spontaneous delivery
  - cerclage

## Data Preprocessing:

- Preprocess predictor attributes
- Preprocess class attributes
- Scaling (Min-Max)
- Class Imbalance Issue: 613 Term / 71 Preterm; SMOTE



## Split Dataset:

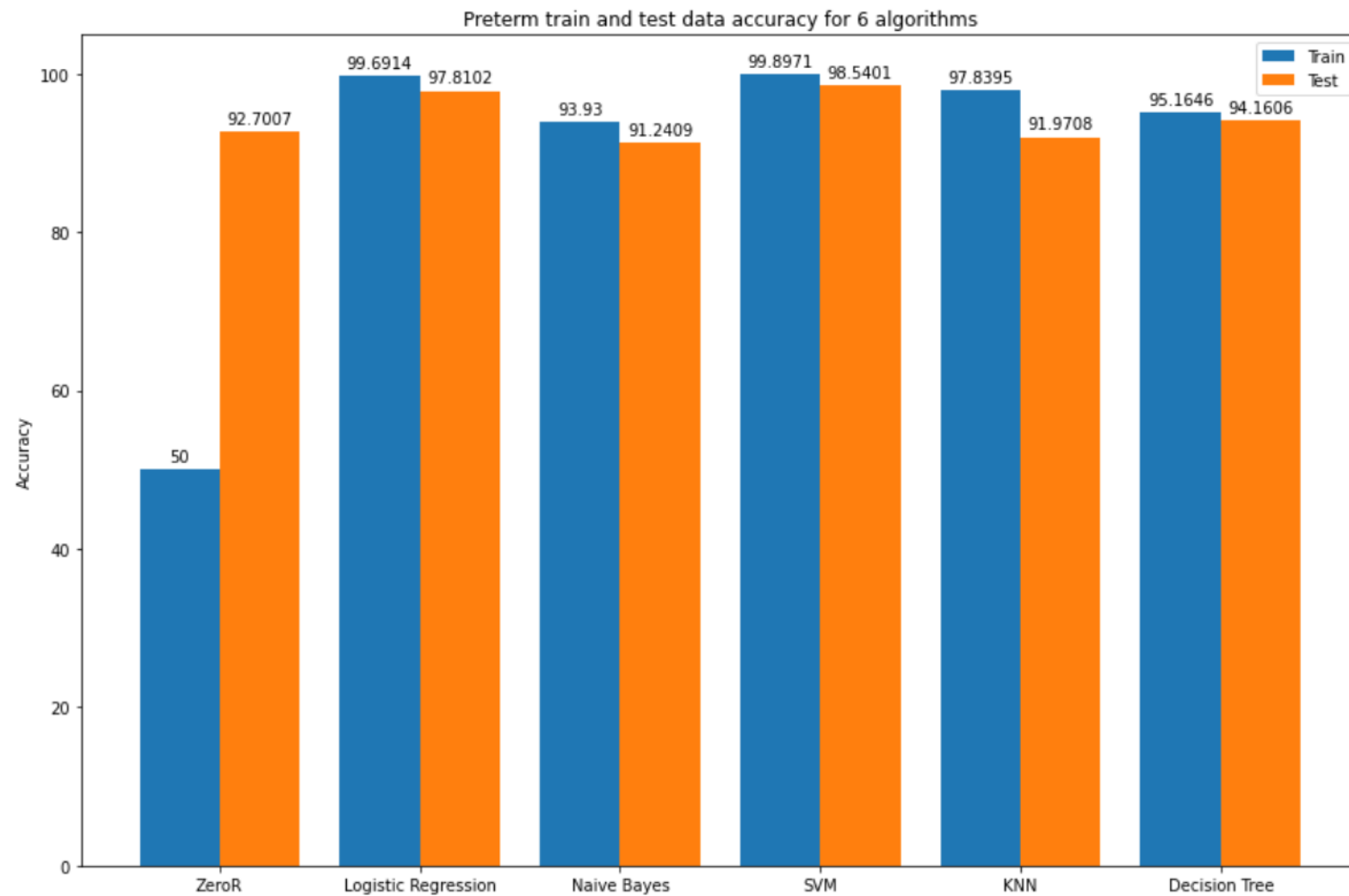
- Train set: 80%
- Test set: 20%

- ZERO-R
- LOGISTIC REGRESSION
- NAIVE BAYES
- SUPPORT VECTOR MACHINE
- K-NEAREST NEIGHBOR
- DECISION TREE



10-fold cross validation was done on each of the classifiers. Hyper-parameters were tuned accordingly to achieve the best results.





Equations of the performance metrics used:

$$\text{False Positive Rate} = \frac{FP}{TN + FP}$$

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

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

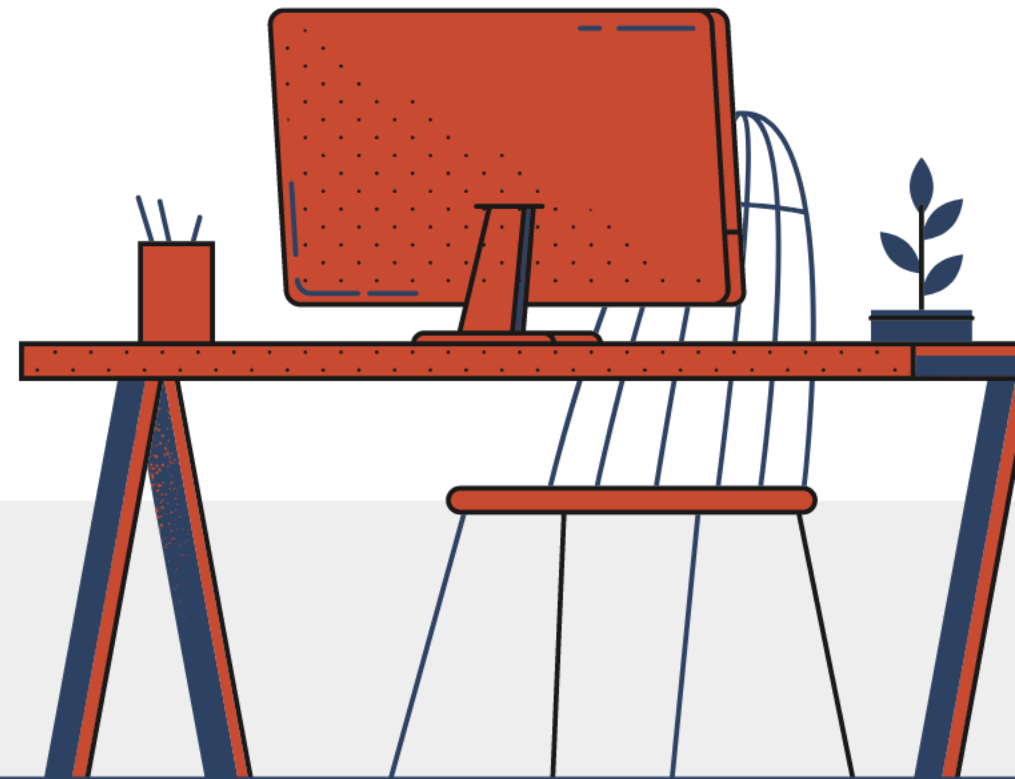
$$\text{F-Measure} = \frac{2 \times \text{recall} \times \text{precision}}{\text{precision} + \text{recall}}$$

Algorithm	Train Accuracy	Test Accuracy	FP Rate	Precision	Recall	F-Measure	Area Under ROC
Zero R	50%	92.7%	0.091	0	0	0	0.500
Logistic Regression	99.63%	98.5%	0	1	0.984	0.992	1
Naive Bayes	93.31%	88.3%	0.500	0.959	0.929	0.944	0.954
SVM	100%	97.8%	0	1	0.984	0.992	0.992
K-Nearest Neighbour	98.14%	92.70%	0.700	0.947	0.984	0.965	0.875
<b>Decision Tree</b>	<b>99.90%</b>	<b>99.30%</b>	<b>0</b>	<b>1</b>	<b>0.988</b>	0.988	0.859

The acquired dataset was trained with six machine learning algorithms (including that of Decision Tree, KNN, Logistic Regression, SVM, Naive Bayes, and ZeroR).

Our experiments have found Decision Tree algorithm attaining the best accuracy score with 99.3% accuracy on the test data.

In future, we intend to use explainable AI techniques to better comprehend the models' prediction.



The End

**Thank you  
for listening**