Delivery of a Pregnant women

PROJECT ATTRIBUTES

- Project Title

  Delivery of a pregnancy women
- Tech Stack

  Python, Machine Learning
- Domain

  Health Care
- Programming Language Used

  Python

  Machine Learning
- Project Difficulty level

  Veteran
- Tools used

  Jupitor Note Book, MS - Excel

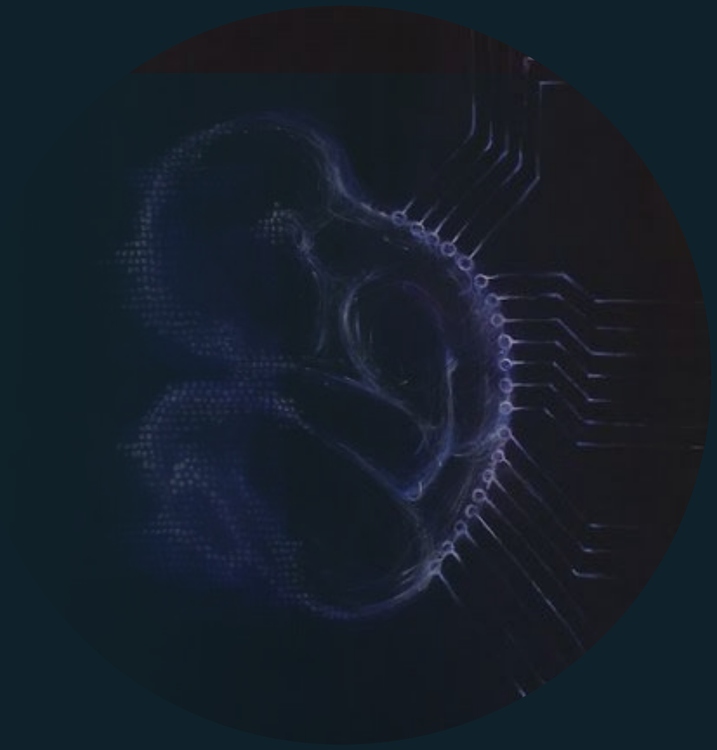
OBJECTIVE

To predict whether an expecting pregnant woman will deliver her baby through c-section mode or through normal delivery.

Background

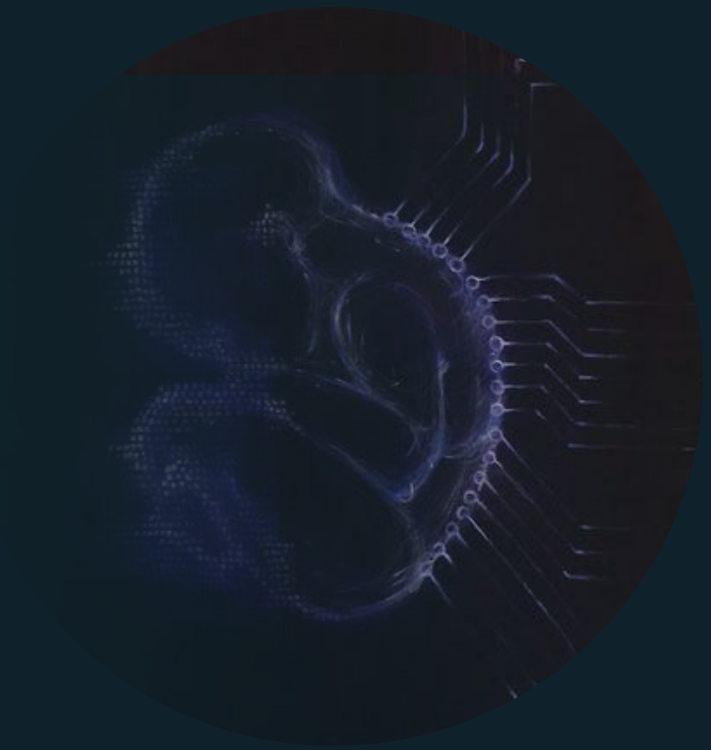
As and when pregnant women near the date of delivery, it becomes a challenge to the treating gynaecologist to wait for normal delivery to take place.

Hence, they induce pain (by administering some injection) and speed up the delivery process. However, sometimes the induction fails.
Even though it is successful, there is no guarantee that there will be a normal delivery.



Scope

Analyze the different factors that influence the delivery mode and come out with a predictive model which can help the treating doctors to find out whether a given sample (expecting woman) will give birth to her baby through C-Section or Normal Delivery mode



Data Set Information

- **SL No & Sample ID:**
  Represents Index.

- **BMI:**
  Represents body fat based on Height and Weight.

- **OB Score:**
  Represents a patient obstetric history simply at a glance.

- **Del Mode:**
  Mode of delivery <- Target Variable

Step by Step Approach to follow:

1) We need to understand what are the necessary packages to import for our given data set.

2) We need to learn the business objective of the given dataset.

3) We need to learn how a typical classification problem we need to work with.

4) We need to learn how to use concatenation to merge for a given two datasets.

Step by Step Approach to follow:

5) We need to check the null values in a given dataset.

6) We need to learn how to check correlation features using heatmap for X-X and X-Y features.

7) We need to check some unique values for ['Consistency score'] variable.

8) We need to learn how to check the total number of variables.

Step by Step Approach to follow:

9) We need to understand which are the features not important for us and we need to decide to drop the features.

10) We need to understand how to use label encode for encoding certain features.

11) We need to understand when to use label encoding and one hot encoding.

12) We need to understand how to use KNN imputer to impute the missing values.

Step by Step Approach to follow:

13) We need to learn how to plot all the features using boxplot and check its outliers.

14) We need to learn how to build a function and treat the outliers.

15) We need to then check if it has removed all the outliers.

16) We need to learn and understand how to use chisquare test and select k best to decide which are the important features for our model.

Step by Step Approach to follow:

17) We need to learn different feature selection process to decide the importance of the features for our model building.

18) We need to learn how mutual info classifier helps us to decide the important features as well.

19) We need to discuss on the VIF (variance inflation factor) where suing VIF and we need to understand which are the features that are important for our model building.

20) We need to create a new dataframe where we are going to share only the important features to consider for our model building.

Step by Step Approach to follow:

21) We need to consider 8 independent features to consider our model building.

22) We need to discuss on Standard scaler how using standard scaler we scale down our data into one format for further model building.

23) Further we need to split our data into train and test set where we are considering 20% as test set and 80% as train set.

24) We need to discuss on how to install SMOTE to balance our data for imbalance class.

Step by Step Approach to follow:

25) We need to discuss on what are the different algorithms we need to build which includes SVM, Decision tree, KNN, etc.

26) We need to learn that our target variable is having an imbalance class and we need to balance our target variable for the same.

27) We need to learn the difference between Under sampling and Oversampling techniques.

28) We need to decide to go ahead with SMOTE (Synthetic Minority Oversampling Technique) since our data is small if we go ahead with Under sampling we will loose data and our model wont be able to give good results.

Step by Step Approach to follow:

29) We will be going ahead with all the machine learning algorithms and apply to build our model and decide the best fitted model.

30) We need to learn how to apply SMOTE and balance our data.

31) We need to learn how to use Logistic regression model using smote and build our model.

32) We need to apply KNN-classifier algorithim using smote and built our model.

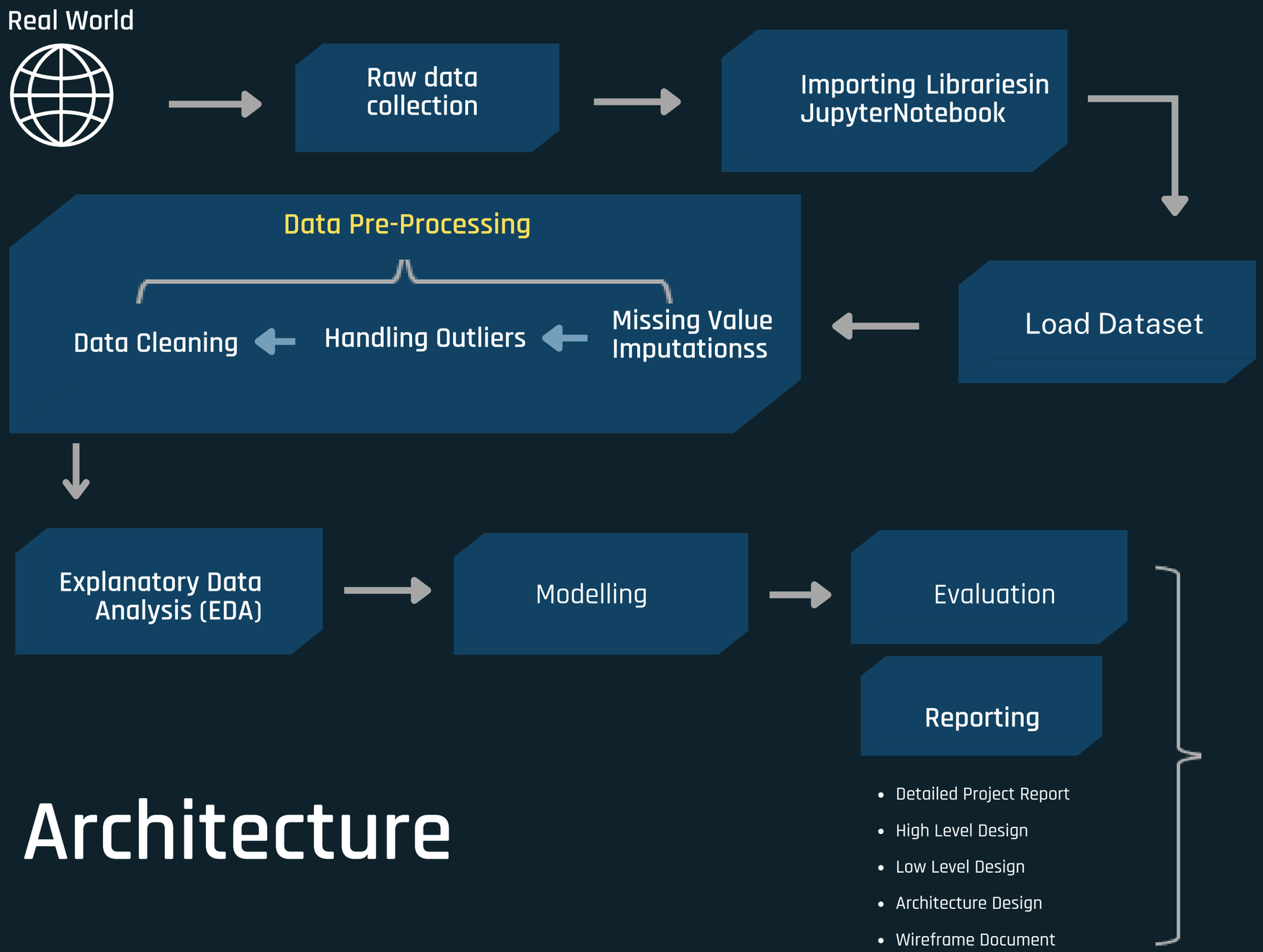
Step by Step Approach to follow:

33) We need to learn how to use Grid search CV for our model.

34) We need to learn how to perform random forest classifier for our model building and checked its accuracy and plot the same.

35) We need to learn how to use Support Vector machine and understood its different kernels including 'linear' ,'poly', 'rbf'.

36) We need to learn how to perform gridsearch cross validation technique on our support vector machine model.



Architecture

Benefits

Gives doctors prior information about the mode of delivery and all medical arrangements can be made before hand and also prepare (mentally and physically) the expecting woman to undergo c-section delivery

Thank You!

For Any Queries, Connect to:

Dhrub@geeklurn.com
preethamtg@geeklurn.com