Machine Learning II Project Proposal

Aisha, Mufaddal, Raju Ahmed

Maternal Health Risk

Abstract

Maternal health is a major challenge in Bangladesh, particularly in rural areas where healthcare is hard to access. Many pregnant women suffer from conditions like high blood pressure and infections that go unnoticed due to a lack of medical facilities and trained professionals. Early marriages, limited education, and poverty add to the problem. Women often cannot reach healthcare centers in time, leading to complications which increases the risk associated to the same. Sometimes this can lead to the deaths. Addressing this issue requires improving access to healthcare and supporting women throughout their pregnancies.

Problem Statement

- So, to calculate the risk associated to it, we have decided to covert that problem into Machine Learning task.
- We want to predict the Risk Intensity Level during pregnancy and classify who has risk

Machine Learning task

- This is a classification Problem.
- There are 3 classes in the target variable.
- Using classification algorithms like Random Forest/Decision tree and SVM.

Data Set

Maternal Health Risk Data

Data has been collected from different hospitals, community clinics, maternal health cares from the rural areas of Bangladesh through the IoT based risk monitoring system.
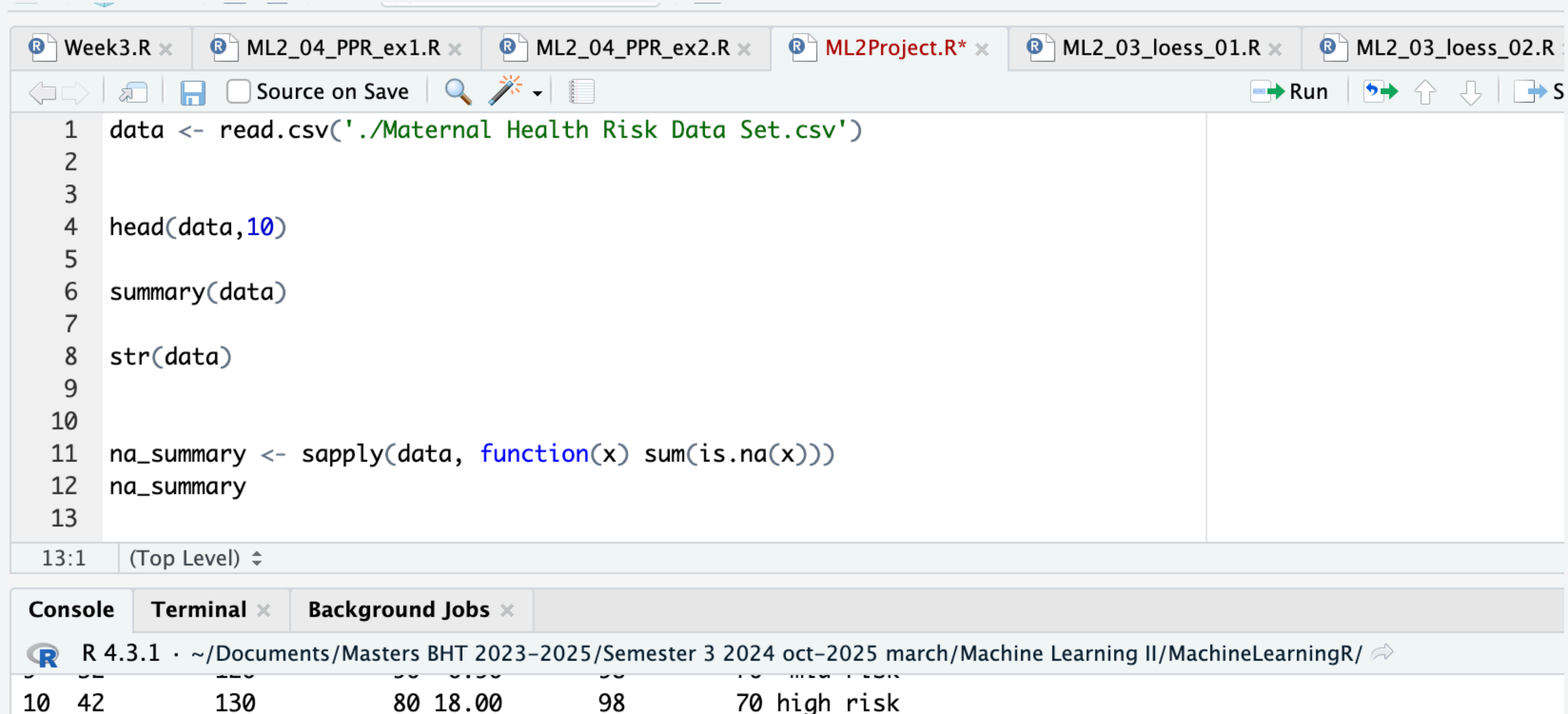
Dataset Information

Records : 1013
Features (Predictors) : 6
Predicted Variable : Risk Type (High, Mid and Low)

Loading dataset with R

Source on Save   Run   S

```r
1   data <- read.csv('./Maternal Health Risk Data Set.csv')
2
3
4   head(data,10)
5
6   summary(data)
7
8   str(data)
9
10
11  na_summary <- sapply(data, function(x) sum(is.na(x)))
12  na_summary
13
```

13:1   (Top Level) ⇕

Console   **Terminal** ×   **Background Jobs** ×

R 4.3.1 · ~/Documents/Masters BHT 2023-2025/Semester 3 2024 oct-2025 march/Machine Learning II/MachineLearningR/

```
10  42      130       80 18.00      98       70 high risk
```

Result

```
> head(data,10)
   Age SystolicBP DiastolicBP    BS BodyTemp HeartRate RiskLevel
1   25        130          80 15.00       98        86 high risk
2   35        140          90 13.00       98        70 high risk
3   29         90          70  8.00      100        80 high risk
4   30        140          85  7.00       98        70 high risk
5   35        120          60  6.10       98        76  low risk
6   23        140          80  7.01       98        70 high risk
7   23        130          70  7.01       98        78  mid risk
8   35         85          60 11.00      102        86 high risk
9   32        120          90  6.90       98        70  mid risk
10  42        130          80 18.00       98        70 high risk
```

```
> na_summary <- sapply(data, function(x) sum(is.na(x)))
> na_summary
       Age  SystolicBP DiastolicBP          BS    BodyTemp   HeartRate   RiskLevel
         0           0           0           0           0           0           0
```

Checking the data out

```
               10  42        130         80 18.00       98       70 high risk

> summary(data)
      Age          SystolicBP      DiastolicBP          BS          BodyTemp        HeartRate
 Min.   :10.00   Min.   : 70.0   Min.   : 49.00   Min.   : 6.000   Min.   : 98.00   Min.   : 7.0
 1st Qu.:19.00   1st Qu.:100.0   1st Qu.: 65.00   1st Qu.: 6.900   1st Qu.: 98.00   1st Qu.:70.0
 Median :26.00   Median :120.0   Median : 80.00   Median : 7.500   Median : 98.00   Median :76.0
 Mean   :29.87   Mean   :113.2   Mean   : 76.46   Mean   : 8.726   Mean   : 98.67   Mean   :74.3
 3rd Qu.:39.00   3rd Qu.:120.0   3rd Qu.: 90.00   3rd Qu.: 8.000   3rd Qu.: 98.00   3rd Qu.:80.0
 Max.   :70.00   Max.   :160.0   Max.   :100.00   Max.   :19.000   Max.   :103.00   Max.   :90.0
  RiskLevel
 Length:1014
 Class :character
 Mode  :character


> str(data)
'data.frame':    1014 obs. of  7 variables:
 $ Age        : int  25 35 29 30 35 23 23 35 32 42 ...
 $ SystolicBP : int  130 140 90 140 120 140 130 85 120 130 ...
 $ DiastolicBP: int  80 90 70 85 60 80 70 60 90 80 ...
 $ BS         : num  15 13 8 7 6.1 7.01 7.01 11 6.9 18 ...
 $ BodyTemp   : num  98 98 100 98 98 98 98 102 98 98 ...
 $ HeartRate  : int  86 70 80 70 76 70 78 86 70 70 ...
 $ RiskLevel  : chr  "high risk" "high risk" "high risk" "high risk" ...
> na_summary <- sapply(data, function(x) sum(is.na(x)))
> na_summary
       Age  SystolicBP DiastolicBP          BS    BodyTemp   HeartRate   RiskLevel
         0           0           0           0           0           0           0
```

Thank You!