Maternal Health Risk Classification

Machine-LearningSemester Project

Class: BSCS-VII (B)

Instructor: Dr. Faheem Akhtar

Dated: December 29, 2022.

Group Members:

Fasiha Naz (023-19-0046)

Fiza Fatima (023-19-0012)

Gorap Kumar (023-19-0011)

**Table of Contents**

[**1.** **Introduction:** 2](#_Toc123245720)

[**2.** **Problem Statement:** 2](#_Toc123245721)

[**3.** **Methodology:** 2](#_Toc123245722)

[**4.** **Dataset discussion:** 3](#_Toc123245723)

[**Attribute Information:** 3](#_Toc123245724)

[**Source:** 3](#_Toc123245725)

[**5.** **Major Outcomes:** 3](#_Toc123245726)

[**6.** **Project Timeline:** 4](#_Toc123245727)

[**7.** **Literature Review:** 4](#_Toc123245728)

[**8.** **Deployment of Dataset** 5](#_Toc123245729)

[**9.** **Conclusion:** 6](#_Toc123245730)

[**10. References:** 6](#_Toc123245731)

# **Introduction:**

Maternity is a phase, from, which almost every woman goes through. Women face different health-condition during it. Such as increased heartbeat, blood pressure, and other conditions which do not happen normally. Since there is no particular measurement scale for how intense the conditions normally. In this fast age, it is hardly possible to take time out for regular gynecologist checkups.

Most of the time, uncertain health conditions are completely normal except when they are not. Thus, there must be a system to determine when things go out of the normal pattern. We have designed this system in which you enter the details i.e., age, BP & Sugar level. As a result, the system will tell the level of risk that a woman holds. From this, we can analyze how safe the child and mother are.

# **Problem Statement:**

The fatal pregnancy rate is increasing day by day. Even in Pakistan, it has reached up to 14.47%. Most of the time it happens due to the inconsistent healthy routine that women practice. Pregnancy triggers many behaviors which make you sick. Some of them are increased sugar levels and high blood pressure. Sometimes, the sugar-level increases so much that it makes one diabetic. Most of the time these changes are considered normal. Which is partly true. However, there is a chance that extremes can hurt pregnancy (both mother and the child). To reduce the confusion about when these conditions are not normal, we have designed a system that checks these conditions and gives risk level as the output of it.

# **Methodology:**

For achieving the desired solution, we will apply the following machine learning techniques to our chosen dataset which is the “Maternal health risk Dataset” To achieve the desired results, we have first converted the output classification column into numerical data. Secondly, we have removed duplicates and highly correlated columns. Duplicates were removed and we applied Smote to increase the number of values in dataset. Then we have seen the data distribution between 3 output classes. Then, we performed normalization. After that, we applied the classifiers such as Random-Forest, Logistic-regression, XGBoost Classifier (Extreme Gradient Boosting), KNN (K-Nearest Neighbors Algorithm), and Decision tree Classifier. The Random-Forest Algorithm worked best without preprocessing, but after preprocessing was applied it didn’t work well. The maximum accuracy was achieved through XGBoost Classifier. XGBoost worked well after the dataset had been preprocessed and its score was better. After Preprocessing many duplicates were removed hence the improvement in XGBoost performance.

# **Dataset discussion:**

## **Attribute Information:**

1. Age: Age of the mother in years.
2. Systolic BP
3. Diastolic BP
4. Blood Sugar: The blood sugar level of the mother
5. Body temperature: Current body temperature in Fahrenheit
6. Heart rate

Lastly, there is one output classifier as Risk Level.

## **Source:**

This dataset is the processed information of major datasets used in maternity clinics. However, we have found it on UCI. An online site that provides public datasets.

<https://archive.ics.uci.edu/ml/datasets/Maternal+Health+Risk+Data+Set>

# **Major Outcomes:**

Considering the number of pregnancies happening in this world. There must be a proper channel through which the phase can be for women. Our major outcome(s) are based upon:

* Increasing accuracy so that the system can become reliable.
* Making sure that our intended accuracy provides the help needed from the system.
* Easy to use the dataset through flask implementation.

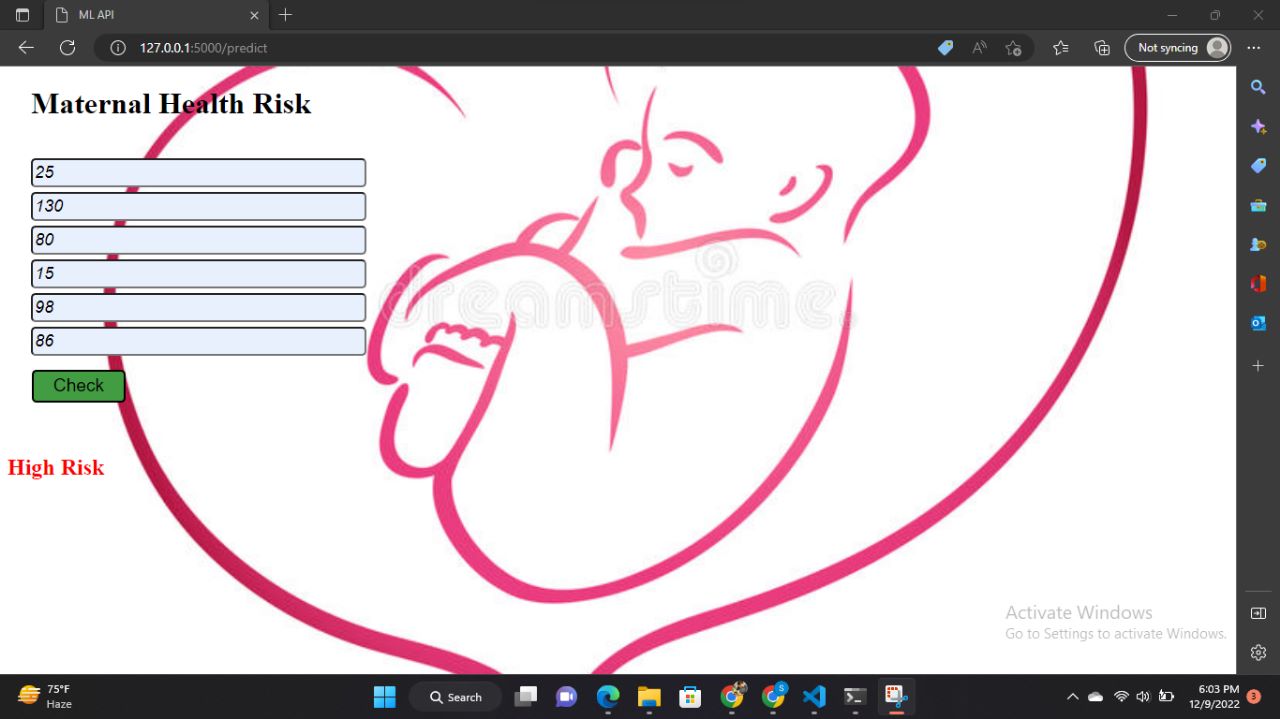
# **Project Timeline:**

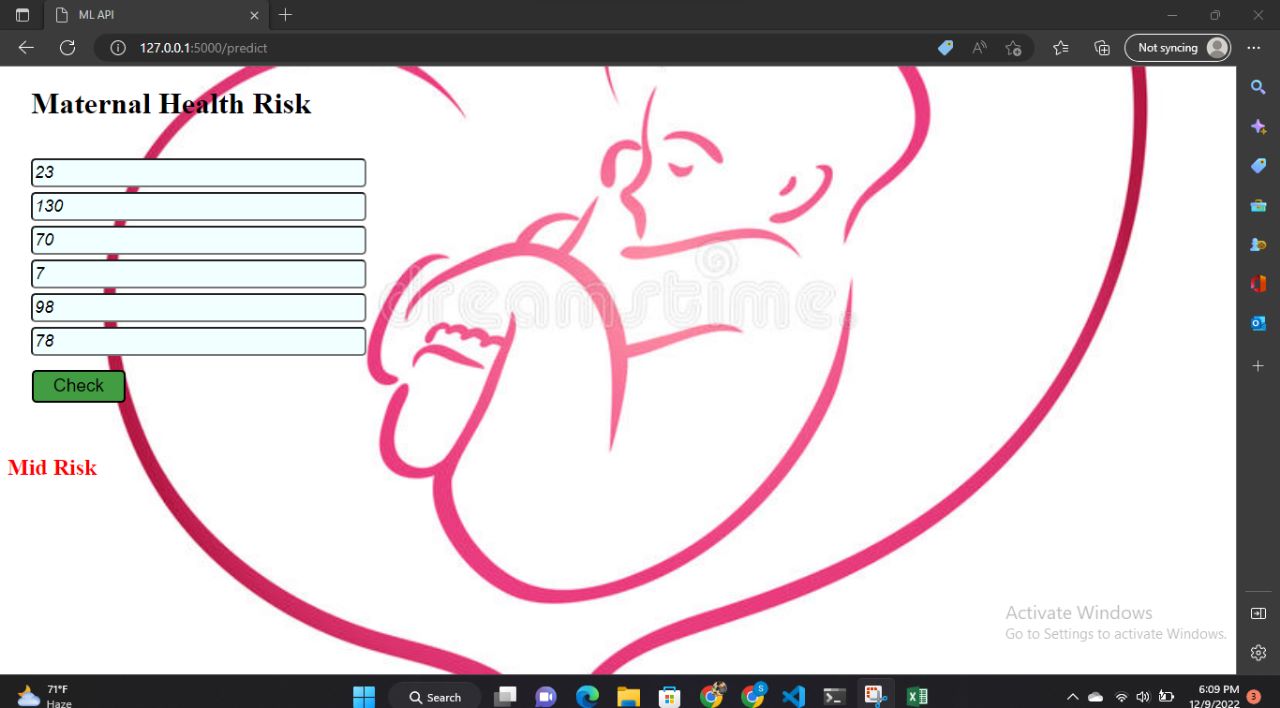
|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Activity** | **November** | | | **December** | | | |
| 12-16 | 17-21 | 22-25 | 22 | 24 | 29 | 29 |
| Literature Review |  |  |  |  |  |  |  |
| Problem Identification |  |  |  |  |  |  |  |
| System Requirements |  |  |  |  |  |  |  |
| System Design |  |  |  |  |  |  |  |
| System Development |  |  |  |  |  |  |  |
| System Testing |  |  |  |  |  |  |  |

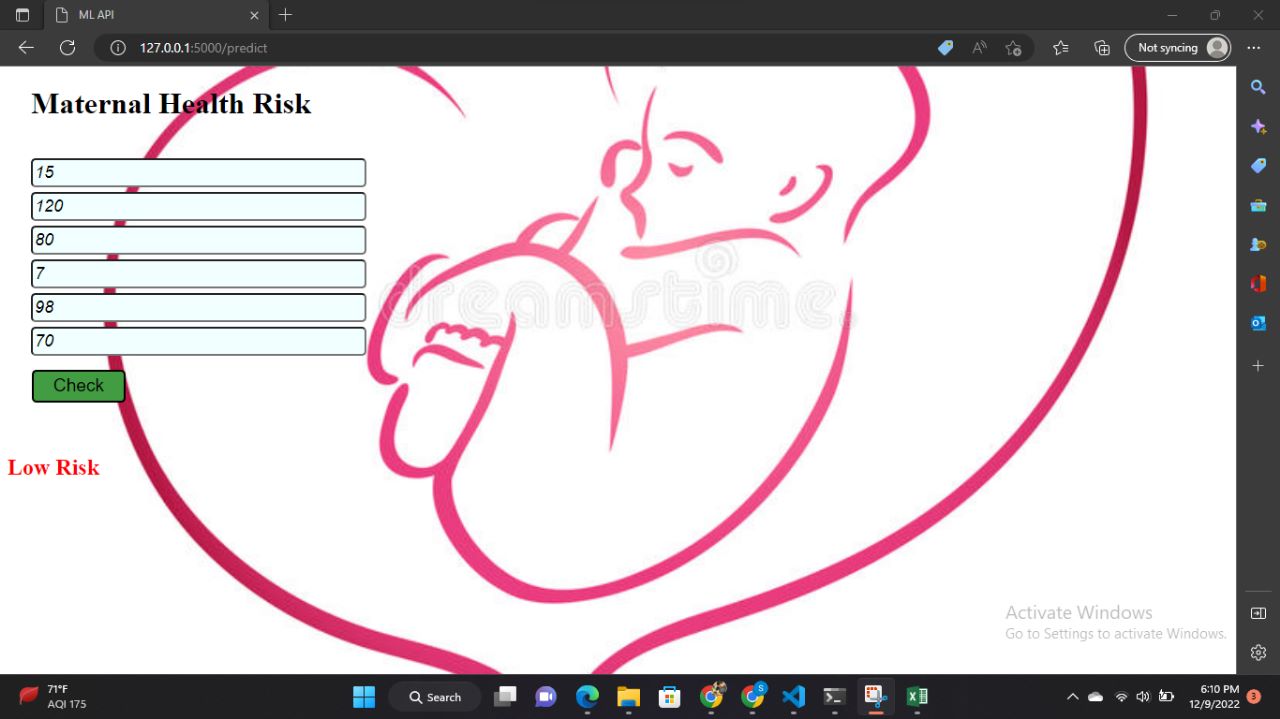
# **Literature Review:**

When we took this data set without preprocessing only random forest algorithm was applied on it. people only work on this data set without applying preprocessing on it, so they got better results even with random forest and other algorithms. without preprocessing it would only give wrong results and results which is of no use. The data set contains too many duplicates, which without prepossessing will only result in bad results which will impact on people's health. Hence any result found on the Internet without preprocessing is of no use and is very dangerous to the people who are using it. We have seen this data on UCI machine learning repository. people had worked on this data set and taken it from here.

# **Deployment of Dataset**







# **Conclusion:**

As discussed earlier, it has been extremely important that every person can self-assist themselves during any minor issues they have been facing in the day-to-day life. Pregnancy, for women, brings these uncertain conditions at hand. There must be a system through which women can regularly check to keep the child safe. With the proposed system, we aim to provide a safe maternity scale that can easily be checked from the comfort of their homes.

# **10. References:**

UCI Repository [accessed 29 December 2022],

<https://archive.ics.uci.edu/ml/datasets/Maternal+Health+Risk+Data+Set>