# Instructions for Running the C Program Implementing Preterm Birth Prediction Algorithms

David Tian

## Introduction

The C program can be used as a decision support tool to predict the preterm birth (PTB) of a patient from her electrical impedance spectroscopy (EIS) measurement in mid-trimester. The prediction algorithm is shown in Figure 1. The algorithm consists of an EIS spectrum selection filter and a PTB predictor. Following data capture with the EIS device, firstly, all the EIS spectra of a patient from an EIS device are input to the filter which selects the best quality EIS spectrum. Secondly, the best quality spectrum is input to the PTB predictor which outputs the probability of PTB of the patient. If the probability is greater than or equal to a cut-off value e.g. 0.5, the predicted class is “preterm” (delivery before 37 weeks); otherwise, the predicted class is “onterm” (delivery ≥ 37 weeks).

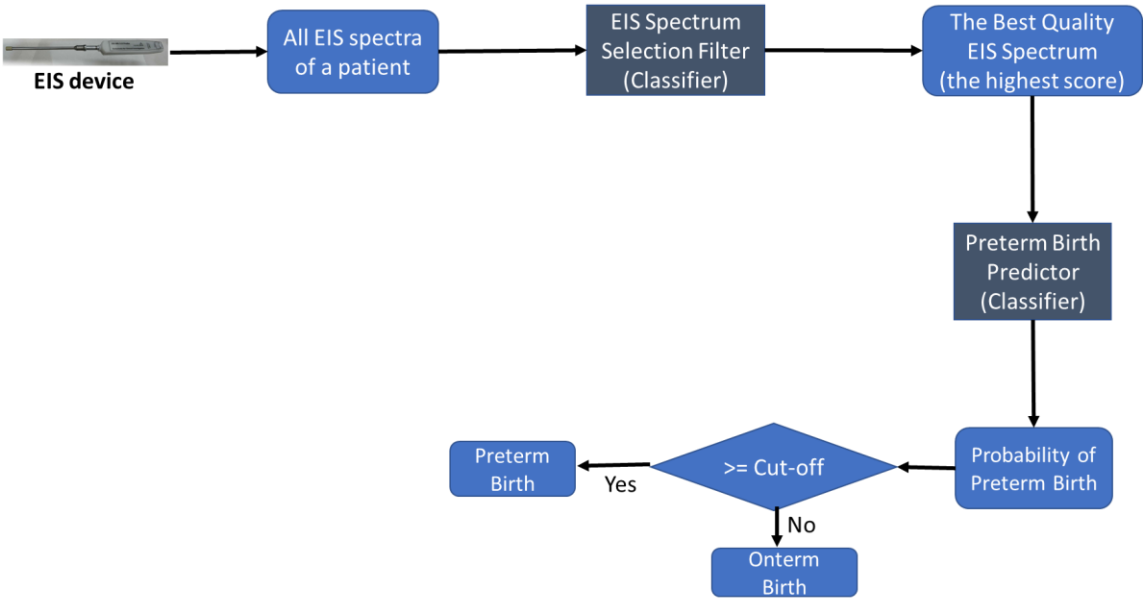


Figure 1: The Preterm Birth Prediction Algorithm

The C program provides 2 PTB prediction algorithms. **Algorithm1** consists of a random forest-based spectrum selection filter and a random forest-based PTB predictor. **Algorithm2** consists of the random forest-based spectrum selection filter and a logistic regression-based PTB predictor. The program files are as follows:

- main.c
- rf23.c (random forest based spectrum selection filter)
- rf92.c (random forest based PTB predictor)
- log_reg23.c (logistic regression based PTB predictor)
- header.h

- 438_V1_all_eis_readings_real_imag_with_dummy_ids.csv[1]

The program takes 3 arguments from the command line: 1) patient ID, 2) PTB predictor to use (random forest based PTB predictor or logistic regression based PTB predictor) and 3) a cut-off value for the probability of PTB, e.g. 0.5. Then, the program fetches, from the above csv file, the EIS data of the patient with the ID provided in the command line, work out a probability of PTB with the patient, and output this probability as well as the outcome of PTB prediction using the specified cut-off value.

## Installation of GCC Compiler and Compiling the C Program

1. Install the GCC compiler on your computer if it has not been installed. To install GCC on **Windows operating system**, download MingW from Downloads - Mingw-w64 and install it. After installation, add the bin subdirectory of your MingW to your PATH environment variable, so that you can run gcc on the command line. Run the following command to check whether GCC is installed:

   C: \>gcc -v

   The following message should be printed if GCC is installed on your system and bin subdirectory is added to the PATH environment variable:

   Using built-in specs.
   …
   gcc version 8.1.0 (x86_64-win32-seh-rev0, Built by MinGW-W64 project)

   If you are using **Linux or UNIX**, then check whether GCC is installed on your system by entering the following command from the command line:

   $ gcc -v

   If you have GNU compiler installed on your machine, then it should print a message as follows:

   Using built-in specs.

   Target: i386-redhat-linux Configured with: ../configure --prefix=/usr .....

   Thread model: posix gcc version 4.1.2 20080704 (Red Hat 4.1.2-46)

   If GCC is not installed, then you will have to install it yourself using the detailed instructions available at https://gcc.gnu.org/install/

   More detailed installation instructions are on the following website:

[1] The EIS data of 438 patients are saved in the CSV file in the folder 'C code' where Patient IDs are from 1 to 438.

C - Environment Setup (tutorialspoint.com)

2. Unzip “C code.zip” into a folder **“C code”** on your computer.

3. Open a command prompt on Windows OS or a shell on Linux and change to the **“C code”** directory using cd command. For example:

   cd “D:\C code”

4. Run the following command to compile the program into an executable file named ptbpredictor:

   gcc main.c rf23.c rf92.c log_reg23.c -o ptbpredictor

## Examples of Running the Program for PTB prediction

1. To predict PTB of the patient with ID 25 using the **random forest based PTB predictor** and the cut-off value 0.5, enter the following command at the command prompt:

   ptbpredictor 25 rf 0.5

```
C:\Users\dtian\OneDrive\Desktop\C code>ptbpredictor 25 rf 0.5
###Predict PTB of ID 25 using a random forest based PTB predictor.###
###1. Select the best EIS spectrum of the ID 25###
score of spectrum 1: 0.8333
score of spectrum 2: 0.0000
score of spectrum 3: 0.0667

best spectrum: 1, score: 0.8333

###2. Predict PTB of the ID 25 using a random forest based PTB predictor and the best spectrum.###

probability of PTB: 1.0000
class: preterm (cut-off value=0.50)
```

2. To predict PTB of the patient with ID 25 using **the logistic regression based PTB predictor** and the cut-off value 0.5, enter the following command at the command prompt:

   ptbpredictor 25 lr 0.5

```
C:\Users\dtian\OneDrive\Desktop\C code>ptbpredictor 25 lr 0.5
###Predict PTB of ID 25 using a logistic regression based PTB predictor.###
###1. Select the best EIS spectrum of the ID 25###
score of spectrum 1: 0.8333
score of spectrum 2: 0.0000
score of spectrum 3: 0.0667

best spectrum: 1, score: 0.8333

###2. Predict PTB of the ID 25 using a logistic regression based PTB predictor and the best spectrum.###

probability of PTB: 0.6669
class: preterm (cut-off value=0.50)
```

It is worth noting that 0.5 is a standard cut-off value. Different cut-off values can be used. An optimal cut-off value maximizes the sensitivity (accuracy of predicting preterm patients) and the specificity (accuracy of predicting onterm patients) of PTB prediction.