Informatics in Medicine Unlocked 28 (2022) 100771

Contents lists available at ScienceDirect

Informatics in Medicine Unlocked

journal homepage: www.elsevier.com/locate/imu

Automated labour detection framework to monitor pregnant women with a high risk of premature labour using machine learning and deep learning

Hisham Allahem, Srinivas Sampalli *

Faculty of Computer Science, Dalhousie University, 6050 University Avenue, Halifax, Nova Scotia, B3H 1W5, Canada

ARTICLE INFO

Keywords:
Uterine contractions
Labour
Premature birth
Machine learning
Deep learning
Artificial intelligence
Electrohysterography

ABSTRACT

It is estimated that more than 1 in 10 babies are born prematurely worldwide. Moreover, premature babies can face lifelong health-related disabilities, such as difficulties in learning or hearing and vision loss. By monitoring uterine contractions, physicians can evaluate the health and progress of the pregnancy and determine if the pregnant woman is in labour, thus assisting them to go to the hospital and help reduce the complications associated with premature birth. In this paper, we aim to mitigate the consequences of premature birth for pregnant women and the foetus using machine learning and deep learning approaches to detect and predict labour. The proposed system was tested for reliability and accuracy. The results show that the deep learning approach achieved the best results with a 0.98 accuracy rate.

1. Introduction

1.1. Overview of premature birth

Premature birth is a serious health issue for the foetus and the mother. It is a consequence of premature or preterm labour, it can be a major pregnancy complication [1,2], and it is the most common cause of death for neonates, even with current advanced healthcare technologies [3–8]. Moreover, there are 15 million premature births per year worldwide, representing 9.6% to 11% of total births [7,9–12]. This percentage makes premature birth the top cause of death of children under five years [13].

Premature birth is defined as birth before completing the 37th week of gestation [5,10,14–17]. Babies born prematurely are not usually fully developed, which can lead to severe perinatal healthcare issues [1,2,12,15,17–23]. For example, as a result of being born prematurely, babies can experience lifelong disabilities such as hearing and vision loss or learning and cognitive impairments [10,20,22,24]. Babies can also suffer from growth impediments and mental health issues [25]. Other health issues can include respiratory, cardiovascular, and neuro-developmental impairments [15,18].

Premature birth can also have an economic impact on healthcare systems, healthcare providers, families, and society [8,15,26,27]. For instance, the cost of premature births is estimated to be between $26 billion and $50 billion per year in the United States alone [17,22,23,28], and £2.95 billion per year in England and Wales [16,29].

1.2. Causes and solutions of premature births

Determining the cause of premature birth is difficult. It can occur spontaneously for many reasons, and the cause is usually unidentified [4,10,11]. However, a pregnant woman can be diagnosed as at a high risk for premature birth if she or one of her family members has a medical history of premature labour [13,20].

1.3. Early detection and intervention in premature birth

Studies have shown that long distances from hospitals, such as in rural areas, are associated with higher mortality and a lack of healthcare access [30]. Furthermore, since it is difficult to prevent premature births, mitigating the effects and consequences of premature births on the foetus and the mother is the best solution found so far. To prevent premature births and mitigate and reduce their consequences, we need a better understanding of this issue. One of the key factors in mitigating premature birth issues is the early detection of labour [5,15,31]. Early detection can help provide medical intervention and achieve the best childbirth outcome and treatment [7,8]. Early premature birth can be detected by monitoring the symptoms of early labour as soon as possible via the pregnant woman's biochemical or biophysical signals [4,9,27]. Furthermore, early detection helps mitigate the health risks for the foetus and the mother and reduces the treatment cost of premature birth complications [32].

* Corresponding author.
E-mail addresses: allahem@dal.ca (H. Allahem), srini@cs.dal.ca (S. Sampalli).

https://doi.org/10.1016/j.imu.2021.100771
Received 27 June 2021; Received in revised form 6 October 2021; Accepted 21 October 2021
Available online 16 December 2021

With practical and cost-effective care, researchers suggest that over three-quarters of babies born prematurely could be saved [10]. Therefore, there is a crucial need for an automated approach to detect and predict labour for pregnant women with a high risk of premature birth [33,34]. Such an approach can help mitigate the consequences of premature birth and provide better healthcare for both the pregnant woman and the foetus [31,35].

In this paper, we propose an approach to automatically detect and predict labour by classifying uterine electrohysterography (EHG) signals using ML and DL algorithms.

The paper is organized as follows. Section 2 lists the related work and literature review. Section 3 details the proposed system and evaluation methodology. Section 4 presents and discusses the results. Section 5 investigates big data and synthetic data. Section 6 presents the experts' domain. Section 7 provides the concluding remarks and scope for future work.

2. Related work and literature review

In this section, we will outline thirteen ML-focused studies and one DL-focused study.

2.1. Machine learning studies

The first study by [29] compared the performance of three classifying algorithms: the RF classifier, Rule-based classifier and Penalized Logistic Regression classifier. They used the Term-Preterm EHG database [36] from PhysioNet [37]. The RF classifier demonstrated the best performance with a specificity of 86%, sensitivity of 97%, AUROC of 94% and mean square error rate of 14%. No accuracy rate, false positive rate (FPR) or false negative rate (FNR) were reported.

The second study by [38] proposed a system for classifying term and preterm deliveries using uterine EHG. They used the Term-Preterm EHG database [36] from PhysioNet [37]. The authors used seven advanced artificial neural network classifiers. Radial Basis Function Neural Network (RBNC) performed the best with 90% accuracy, 85% sensitivity, 80% specificity, 90% area under the ROC curve (AUC), and a 17% mean error rate. No results were reported on the FPR or FNR.

The third study by [39] proposed a new algorithm to use to predict premature labour for women with a high risk of premature labour using the term and preterm EHG signals. They used the Term-Preterm EHG database [36] from PhysioNet [37]. SVM was used to classify uterine EHG signals into term and preterm and reported an accuracy of 96.25%, sensitivity of 95.08%, and specificity of 97.33%.

The fourth study by [40] proposed a method for labour detection using both EHG and foetus heart rate (HR). They used a Bloomlife wearable sensor to collect uterine EHG and FHR from 37 pregnant women (19 labour and 18 non-labour). The authors then summarized the features in terms of the mean and standard deviation over 20-minute windows into the power of the EHG and the frequency and amplitude of the main peak. They used the RF classifier and achieved an accuracy of 79% for EHG only, 81% for HR only and 87% for both. No results were reported for FPR or FNR.

The fifth study by [18] proposed a system to distinguish between premature and term birth patterns using recurrence quantification analysis (RQA) and principal component analysis (PCA). They used EHG signals from 20 pregnant women with a risk of premature birth between the 24th and 28th weeks of pregnancy. The authors used Support Vectors Machine classifications (multiclass SVM) to train the classifiers. The accuracy of the classification was 83.32%. On the other hand, the authors did not report other essential results, such as FPR and FNR, on how the classifiers performed.

The sixth study by [35] proposed a machine learning model to non-invasively detect labour in unsupervised free-living settings using a Bloomlife wearable body sensor device. The proposed model has two phases. In the first phase, the researchers used data collected under supervised laboratory settings to develop artefact and labour probability estimation models. They combined EHG and HR data collected at different gestational ages and showed high accuracy in identifying artefacts and labour. In the second phase, they deployed the models on 142 pregnant women from week 22 to delivery. Pregnant women had an average of seven hours of data.

The authors concluded that within the last 24 h of pregnancy, labour probability is consistently higher than in any other gestational week. The authors recruited an adequate number of pregnant women for the study; however, the proposed system cannot be applied for monitoring premature labour since it can detect labour only 24 h before delivery, and premature labour can occur spontaneously at any pregnancy week. Furthermore, the authors did not provide any results on the accuracy, sensitivity, specificity, false positive (FP) or false negative (FN) rates.

The seventh study by [41] proposed a system using the SVM classifier with RBF kernel function to classify EHG signals into term and preterm delivery. The Term-Preterm EHG database [36], which contains 300 records, from PhysioNet [37] was used in the study. The authors used the autoregressive model (AR) for feature extraction and used the particle swarm optimization (PSO) algorithm to find the optimal features. The authors reported a 97.1% accuracy rate, 95% sensitivity, and 99% specificity. No results were reported regarding FPR or FNR.

The eighth study by [42] proposed a novel and highly efficient approach that uses one feature to classify EHG signals. It is based on the centroid frequency, a single time-varying feature extracted using spectral analysis. They used the Term-Preterm EHG database [36] from PhysioNet [37]. The authors used four classification algorithms. SVM performed the best with a 99.74% accuracy rate, 99% sensitivity, and 99% F score (F). They used only one feature that is used in the classification algorithms, which could imply a bias in the results since EHG signals have several features that could change how the algorithm would classify the signals. Furthermore, no FPR or FNR were reported.

The ninth study by [43] proposed a study to evaluate EHG signals for term and preterm delivery recognition using the RF classification algorithm. They used the Term-Preterm EHG database [36], which contains 300 records, from PhysioNet [37]. The authors reported an accuracy of 93%, sensitivity of 89%, specificity of 97%, and AUC of 80%. However, neither FPR nor FNR was reported.

The tenth study by [44] proposed a low computational and efficient algorithm to detect premature labour using EHG signals. The authors used the Term-Preterm EHG database (TPEHGDB) with 300 datasets of pregnant women's EHG signals. Two hundred sixty-two of these datasets are from term delivery and 38 are from premature delivery. SVM was used for classification.

The authors reported an accuracy rate of 99.5%, sensitivity rate of 98.9% and specificity rate of 99.3%. It should be noted that the authors did not provide any results on the FPR or FNR. Moreover, the achieved rates could have some bias towards the term or premature datasets because, from what we learned from research papers, each woman's EHG is different and detecting premature labour is highly complex. More rates such as FPR and FNR could clarify these viewpoints. Finally, the rates are from one small database only; hence, more testing on other available databases is required.

The eleventh study by [45] proposed a system to monitor pregnant women's uterine EHG signals and alert them when the initial contractions of labour occur. The system consists of AgC12 wearable electrodes. EHG data are captured by the electrodes and a smartphone application, and the pregnant woman's health status is monitored and sent to the healthcare provider. The authors used the Term-Preterm EHG Database (TPEHGDB) with 300 datasets to extract features and classify the term and preterm datasets using the following machine learning classifiers: SVM, Naïve Bayes classifier, K-NN classifier, Gradient Boost and Decision Tree.

The SVM performed the best, with an accuracy rate of 96%, a specificity rate of 94% and a sensitivity rate of 92%. Naïve Bayes performed the worst, with an accuracy rate of 85%, a specificity rate of 84% and a sensitivity rate of 80%. There was no reporting of FPR or FNR. The datasets are relatively small, and more testing is needed with available databases.

The twelfth study by [46] proposed a novel method to classify EHG signals into pregnancy or labour signals. They built their method based on wavelet transform, sample entropy and stacked sparse autoencoder. The authors used the Icelandic 16-electrode Electrohysterogram Database [47] of PhysioNet [37].

A two-hidden-layer stacked sparse autoencoder (SSAE) deep neural network with a softmax classifier network was chosen to automatically classify EGH signals. The SSAE algorithm was compared with two widely used classification algorithms: the extreme learning machine (ELM) and SVM. The results show that SSAE performed the best, with 90% accuracy, 92% sensitivity and 88% specificity. However, there were no results reported for the FPR and FNR. Furthermore, the number of datasets for training and testing is too low (300 in total), which can cause biases or overfitting. More datasets are needed to further confirm and generalize the results.

The thirteenth study by [48] proposed a methodology to automatically detect EHG contractions using signal envelope features. They evaluated two main features: contraction detection and its related delineation accuracy.

The authors used the Icelandic 16-electrode Electrohysterogram Database [47] of PhysioNet [37]. They selected datasets of gestational age between 29 and 41 with 20 subjects, which represent 16% of the datasets. From those 20 subjects, the authors obtained 2383 contractions. The five methods were then applied to detect contractions using the associated energy bursts as a feature with a sample window of 70 s.

Of the five methods, the squared root mean square (RMS) gave the best results with an accuracy rate of 97.15% for contractions detection and 89.43% for delineation. The paper is very well explained, but no specificity or sensitivity rates were given. Also, the authors discussed the FPR in the abstract; however, they reported the same rates in the results section as FNR, not FPR. Another issue is that the datasets are too small, which can cause biases or overfitting. There are other available datasets the authors could use to further confirm and generalize the results. In the next section, we will present the DL-related work.

2.2. Deep learning studies

The study by [49] used the convolutional neural network (CNN) to identify uterine contractions in EHG signals from two databases. The first database is the Icelandic 16-electrode Electrohysterogram database [47] from PhysioNet [37]. The second database was collected by recruiting 20 pregnant women with singleton pregnancies at the Department of Gynaecology and Obstetrics, Peking Union Medical College Hospital, Beijing. The authors created 14,016 45-second-length segments from the first database. The author extracted a total of 308 segments from the second database. All EHG segments were saved and normalized by resizing them to 482 X 482 pixel images.

Furthermore, CNN's ability to recognize uterine contractions was evaluated using five-fold cross-validation using the first database. The CNN model was then applied to the second database for testing the model. The results for the first database are 98% accuracy, 87% sensitivity, 93% specificity, 92% AUC, FP value of 121 and FN value of 903. For the second database, there was 93% accuracy, 88% sensitivity, 97% specificity, 87% AUC, FP value of 5 and FN value of 18. The study is well written and designed. The authors were able, to an extent, to increase the input datasets to the CNN models from the first database since the size of EHG data from pregnant women is hard to collect in large quantities to use in deep learning approaches. However, the authors did not provide any information on the average recording time; hence, more testing is needed to evaluate the proposed approach's ability to detect if a pregnant woman is in labour. Furthermore, the proposed approach needs tocodynamometer signals, which will make it more expensive and unusable in home-monitoring scenarios due to the high demand for medical devices and medical staff for data collection. In the next section, we will detail the proposed system and evaluation methodology.

The above previously proposed systems are limited in terms of:

- Did not report important evaluation criteria such as misclassification rate (MisCl), F, FPR, FNR and AUC.
- Do not have a practical system to solve the issue of premature birth.
- Some studies used only one database.
- One study used only one feature to solve the complex issue of detecting premature labour.

Our proposed system improves on the previously proposed systems in the following ways:

1. Imitates obstetricians' labour detections in terms of the uterine contractions monitoring process.
2. Has improved accuracy, FPR and FNR.
3. Uses the proposed system in a practical application for monitoring pregnant women.
4. The proposed system is applicable in developing countries.
5. Applicable for premature and normal births.

3. Materials and method

3.1. Overview

The two main challenges of premature birth are detection and prediction [1]. Currently, there is no permanent treatment for preventing premature birth. Furthermore, while preventing premature births is difficult, labour itself is detectable. Moreover, the health of the pregnant woman and the foetus is the central issue in premature birth alongside the resulting financial burden. Therefore, the early detection of premature birth through monitoring systems is the best option for preventing or mitigating the consequences of premature birth [50].

We need a system that continuously monitors the uterine contractions of the pregnant woman and detects if she is in labour, which would allow for the early detection of premature birth and provide an opportunity for medical intervention that could reduce premature birth outcomes. In our previous work [51], we proposed a framework for monitoring pregnant women with a high risk of premature birth. Although our previous work results were promising for labour detection [51], the previously proposed system was lacking in two areas. The first is that we need to extract more features from EHG signals to increase the detection accuracy. The second is that we need to predict premature birth occurrence using ML and DL approaches. Therefore, we will be using ML and DL approaches to overcome the limitations of the previously discussed ML and DL studies and improve the framework's reliability and accuracy we previously proposed [51].

3.2. Data selection and preprocessing

3.2.1. Data selection

To test the framework, we used the following five publicly available uterine contraction EHG signal databases of pregnant women from the PhysioNet repository [37]:

1. Icelandic 16-electrode Electrohysterogram Database [47].
2. Term-Preterm EHG Database [36].
3. Term-Preterm EHG DataSet with Tocogram [52].
4. CTU-CHB Intrapartum Cardiotocography Database [53].
5. OB-1 Fetal ECG Database [37].

The first database is the Icelandic 16-electrode Electrohysterogram database [47], which has 45 recordings of pregnant women performed between 2008 and 2010. Each recording consists of 16 datasets of uterine contraction EHG signals.

The second database is the Term-Preterm EHG database [36], which contains 300 uterine EHG recordings of pregnant women obtained from 1997 to 2005. Each record is 30 min long and has 12 datasets of uterine contraction EHG signals.

The third database is the Term-Preterm EHG DataSet with Tocogram [52], which contains 24 uterine EHG recordings of pregnant women and another five uterine EHG recordings of non-pregnant women.

The fourth database is the CTU-CHB Intrapartum Cardiotocography database, which contains 552 cardiotocography (CTG) recordings collected between 2010 and 2012 from the Czech Technical University (CTU) in Prague and the University Hospital in Brno (UHB) [53]. Each recording has a foetal heart rate (FHR) time series and a uterine contraction (UC) signal. The length of the recording is 90 min at most.

The fifth database is the OB-1 Fetal ECG database, which contains five hours and 42 min of foetal ECG signals and uterine contraction signals [37]. Next, we will explain data preprocessing and how we extracted the final datasets for testing.

3.2.2. Data preprocessing

The databases downloaded from the PhysioNet repository [37] are in the WFDB signal files format. To read and convert them into CSV format, we used a tool called **rdsamp** (see Appendix A).

Moreover, the main goal of our proposed framework is to detect and predict premature births by analysing EHG signals for labour contraction patterns for 30-minute monitoring sessions. To do so, we need to extract the datasets from the five databases and prepare them to be 30-minute-long single columns of readings. To further illustrate, each dataset in the five databases has several columns of EHG readings. Based on that, the datasets extraction process has the following steps:

1. Extract each column separately from each dataset.
2. Split each column to the length of 30 min following these rules:
   (a) If the length of the dataset is less than 30 min, then discard the dataset.
   (b) If the length of the dataset is 30 min, then extract the dataset as it is.
   (c) If the length of the dataset is more than 30 min, then extract the first 30 min.
   (d) If the dataset has more readings after the first 30 min, then repeat the steps starting from process number (a).

We extracted a total of 7271 30-minute-long datasets. The datasets we have considered are derived from three different pregnant women's uterine contraction databases. The pregnant women were either in the labour or non-labour stage when they were monitored. The ratio between the pattern and non-pattern datasets is 60% pattern to 40% non-pattern. This ratio of recordings makes the data balanced to test and evaluate the machine learning and deep learning algorithms for labour detection. It also minimizes the bias in the datasets towards pattern or non-pattern datasets. In the next section, we will present both ML and DL approaches implementation in detail.

3.3. Machine learning approach

In this section, we will illustrate feature selection, the chosen classifiers and our selection of hyper-parameters. Note that for all the classifiers, the split data ratio is 70% for training to 30% for testing. The implementation of the classifiers was done on a Windows 10 operating system (OS) with a 3.60-GHz 8-Core Intel i7-9700KF processor, 128 GB of RAM and NVIDIA Quadro RTX 8000 48-GB GDDR6 memory GPU.

3.3.1. Feature selection

EHG is used to record uterine contractions, and it is the most consistent and efficient factor in predicting and detecting labour [1,54–58]. To use uterine contractions' EHG in predicting and detecting premature birth, we need to extract uterine contractions' quantitative features from EHG readings [1,59,60].

Furthermore, feature extraction can generate detailed and improved information from raw EHG signals [61]. Moreover, many studies extracted EHG features based on the time domain, frequency domain, and time–frequency domain [61]. We will be using the following features:

- Mean frequency (MNF), which represents the frequency.
- Peak frequency (PF), which represents the amplitude.

In the next section, we will illustrate the hyper-parameters search for the ML classifiers.

3.3.2. Hyper-parameters search

Hyper-parameters are numeric or boolean values the user can set and configure before the classifier learning process. They help to improve the training time, performance and prediction of the classifier. Examples of hyper-parameters are learning rate, number of epochs or random state. Each hyper-parameter directly affects the classifier learning process, which means we need to make sure we choose the best hyper-parameters according to the data we are using.

To choose the optimum hyper-parameters for the classifiers, we will use a tuning technique called Grid Search. Grid Search is a technique that calculates the optimum hyper-parameters for a given classifier by an exhaustive search of a given hyper-parameter. This technique can save time and resources when training the classifier. Refer to B for the grid search results of the hyper-parameters. Next, we will present ML classifiers.

3.3.3. Decision tree

DT is a commonly used classifier tree that starts at the root, then each node represents an attribute test, and sorting down occurs until reaching a leaf node, which will provide a classification [62] (see B).

3.3.4. Random forest

RF was proposed by Breiman in 2001 [63]. It is made of many decision trees. More trees lead to more accurate results [63,64] (see B).

3.3.5. Support vector machine

The SVM classifier was introduced in 1995 [65]. It became popular due to its classification accuracy, robustness and indifference towards the input data type [66]. It is based on Vapnik statistical learning theory and can be applied to many classification and regression problems [67,68] (see B).

3.3.6. Naïve Bayes

NB is one of the most widely used classifiers for many applications with a fast learning and testing process [69]. It is designed based on the Bayesian rule and probability theorems [69,70] (see Appendix B). In the next section, we will illustrate our choice of the DL model.

3.4. Deep learning approach

Deep learning is a subfield of machine learning under the umbrella of artificial intelligence. Deep learning models, such as artificial neural network (ANN) and convolutional neural network (CNN), learn from complex relationships or high-level features [71]. The difference between deep learning and machine learning is that deep learning is a more complex and advanced technique. It utilizes big data to extract and form complex relationships between features to obtain a deeper understanding of data [72].

By utilizing the power of deep learning, we aim to personalize the scheme to fit pregnant women's different needs and characteristics. For example, a pregnant woman could be diabetic, could be a smoker and could have had a premature birth before. Another pregnant woman could have hypertension and could have gone through her first pregnancy with no history of premature labour. In this case, deep learning could be fed with personal features, alongside the EHG signal features such as amplitude, to predict if the pregnant woman is in premature labour.

The ANN model primarily consists of several layers: an input layer, several hidden layers and output layers [73,74]. Each layer is comprised of neurons (also called nodes) that form parallel-connected neural networks to achieve superior computation speed and power [73]. A layer also has activation functions that are critical in designing the neural network. It is used to decide the neuron's output to the next layer depending on the activation function type. Two commonly used activation functions are:

- Rectified Linear Activation (ReLU): It makes the network training fast due to its simple and easy computation. If the input x is less than 0, then set the input to 0. Otherwise, set the input to x. ReLU is given by the following equation:

$$relu(x) = max(0, x) \tag{1}$$

- Logistic (Sigmoid): It is best used for prediction since it produces a value between 0 and 1. Sigmoid is given by the following equation:

$$S(X) = \frac{1}{1 + e^{-x}} \tag{2}$$

When compiling the model, we use some parameters such as optimizers, loss functions and learning rate. Optimizers are methods used to minimize the losses by changing the attributes of the neural network. Commonly used optimizers are Adam, RMSprop and Stochastic Gradient Descent (SGD).

Loss functions are methods used to predict errors in the neural net and update their weights. Examples of loss functions are Mean Squared Error (MSE), Binary Cross-entropy (BCE) and Categorical Cross-entropy (CC). Learning rate refers to how much to update the neurons' weights when training the model, which ranges from 0.0 to 1.0. A lower learning rate means more training time. Finally, ANN has two phases: forward pass and backtrack. During the forward pass, the neural network outputs from the input layer to the output layer using activation functions. During the backtrack phase, we fine-tune the weights of the neural network using optimizers, loss functions and learning rates. One forward pass and backtrack is called an iteration, during which one batch of data is passed through. Finally, the implementation of the DL model was done on a Windows 10 OS with a 3.60-GHz 8-Core Intel i7-9700KF processor, 128 GB of RAM and NVIDIA Quadro RTX 8000 48-GB GDDR6 memory GPU.

3.4.1. Feature selection

We will be using the mean frequency (MNF), peak frequency (PF) and median frequency (MDF) features. Furthermore, to personalize the features for each pregnant woman, we need to select features that are common between the five databases. Each database has a description record of the pregnant women, such as their gestational age, weight, or maternal risk factors. However, not all databases have the same record information. For example, one database has information about pregnant women's diabetes status, while other databases lack such information.

To avoid any bias towards these features in the ANN model, we will only select features collected among all five databases alongside the three features: MNF, and PF and median frequency (MDF). We can choose more features; however, we need to test if more features will give us more accuracy and, most importantly, the new improved results could indicate bias since not all the records will have the same features. For example, if we add obesity to the features, some patients are either not obese or their obesity status is not reported. In this case, we cannot be sure if obesity plays a role in premature birth or not. In total, we will have six selected features:

- **PF**, which represents the amplitude.
- **MNF**, which represents the frequency.
- **MDF**, which represents the frequency.
- **Pregnancy gestational age**.
- **The pregnant woman's age**.
- **Parity**, which is the number of times a woman has given birth with a gestational age of 24 weeks or more and whether the child was born alive or was stillborn.

In the next section, we will illustrate our choice of the DL model.

3.4.2. Artificial neural network

ANN is a data processing archetype that mimics the biological nervous system of the human brain [73]. Its base is composed of neurons (also called nodes) that form parallel neural networks that can achieve superior computation speed and power [73]. ANN primarily consists of an input layer with several neurons, several hidden layers and output layers [73,74].

To the best of our knowledge, only one study used EHG signals to identify uterine contractions using CNN [49]. The issue with this study is that it cannot be applied to a continuous monitoring scenario. In the next section, we will illustrate the hyper-parameter searching and setting for the ANN model.

3.4.3. Hyper-parameter searching and setting

To choose the optimum hyper-parameters for the ANN model, we will use the Grid Search tuning technique. We will search for the best number of layers, neurons, activation functions, optimizers and loss functions.

The search range was as follows:

- The number of layers: 2, 3, 4 and 5.
- Neurons: 16, 32, 64, 128 and 256.
- Activation functions: ReLU and sigmoid.
- Optimizers: SGD, RMSprop and Adam.
- Loss functions: Binary cross-entropy and Mean Squared Error.
- Learning rate: 1e-2, 1e-3 and 1e-4.

In the grid search experiments, the network has shown excellent accuracy and precision results based on the following parameters:

- Four hidden layers.
- The number of nodes: 128.
- The activation functions: ReLU for the hidden layers and sigmoid for the output layer.
- Adam for the optimizer.
- The binary cross-entropy for the loss functions.
- Batch size of 20 with 50 epochs.
- Learning rate of 1e-3.

ReLU is expected to perform the best for the hidden layers as it is one of the most common and effective activation functions. Moreover, the sigmoid activation function is used for binary classification. Therefore, it performed the best for the output layer. Finally, the binary cross-entropy loss function is the best for binary classification. In the next section, we will illustrate the evaluation methodology.

Table 1
Confusion matrix.

	Classified labour	Classified no labour
Labour	True positive (TP)	False negative (FN)
No labour	False positive (FP)	True negative (TN)

3.5. Evaluation methodology

We will evaluate the framework's ML and DL approaches separately in terms of:

- The framework algorithms' reliability and accuracy in detecting labour.
- Smartphone resources performance.

3.5.1. Reliability and accuracy analysis

We will be using a confusion matrix, which is a two-dimensional matrix table used to rate the performance of a classifier based on test data [75], for the framework algorithms' performance evaluation. Table 1 represents the confusion matrix we will be using.

From this table, we will calculate the following equations to measure the framework algorithms' performance:

- Accuracy (AC) is the percentage of how often the classifier is correct overall to identify the dataset as labour or no labour. AC is given by Eq. (3):

$$AC = \frac{truepositive(TP) + truenegative(TN)}{TP + TN + FP + FN} \tag{3}$$

- MisCl is the percentage of how often the classifier is wrong overall in identifying the dataset as labour or no labour. MisCl is given by Eq. (4):

$$MisCl = \frac{FP + FN}{TP + TN + FP + FN} \tag{4}$$

- Recall (R) is the percentage of how often the classifier identifies the dataset as labour when it is labour. R is given by Eq. (5):

$$R = \frac{TP}{TP + FN} \tag{5}$$

- Precision (P) is the percentage of how often the classifier is correct when it identifies the dataset as labour. P is given by Eq. (6):

$$P = \frac{TP}{TP + FP} \tag{6}$$

- F is the balance between the P and the R values. A higher F value means a better classifier, especially since we care more about labour datasets in our results. F is given by Eq. (7):

$$F = 2 * \frac{P * R}{P + R} \tag{7}$$

- FNR is the percentage of how often the classifier identifies the dataset as no labour when it is labour. This evaluation parameter is important because if it is high, then it means the application will not trigger an alarm even though the dataset indicates there is labour. FNR is given by Eq. (8):

$$FNR = \frac{FN}{TP + FN} \tag{8}$$

- FPR is the percentage of how often the classifier identifies the dataset as labour when there is no labour. This evaluation parameter is important because if it is high, then it means the application will trigger an alarm even though the dataset does not indicate there is labour, which will result in pregnant women having more unscheduled visits to the hospital for no reason. FPR is given by Eq. (9):

$$FPR = \frac{FP}{FP + TN} \tag{9}$$

Table 2
DT classifier confusion matrix.

	Classified pattern	Classified no pattern
Pattern	785	49
No pattern	52	1296

Table 3
RF classifier confusion matrix.

	Classified pattern	Classified no pattern
Pattern	814	31
No pattern	62	1275

Table 4
SVM classifier confusion matrix.

	Classified pattern	Classified no pattern
Pattern	806	28
No pattern	102	1246

Table 5
NB classifier confusion matrix.

	Classified pattern	Classified no pattern
Pattern	811	23
No pattern	192	1156

- The area under the ROC curve (AUC) is a statistical value between 0 and 1 that measures the performance of a classifier. It is the probability of the classifier ranking a randomly chosen positive instance higher than a randomly chosen negative instance [76]. A higher AUC value means better performance of the classifier. AUC is given by the following Eq. (10), where TP stands for true positive, FP stands for false positive, P stands for positive and N stands for negative:

$$AUC = \int_0^1 \frac{TP}{P} d\left(\frac{FP}{N}\right) \tag{10}$$

- Receiver operating characteristic (ROC) is a performance evaluation measurement for binary classification algorithms. It is represented in a graphic plot between the true positive rate (TPR) and the false positive rate (FPR) within different thresholds.

The following section will present the results.

4. Results and discussion

In this section, we will present the performance evaluation results and discussion of the ML and DL approaches. Moreover, we will present the smartphone resource consumption during the beginning, middle, and end of the monitoring session. The monitoring session length was 30 min, during which Bluetooth communication was sending EHG signals from one tablet to the other.

First, we will represent the confusion matrix table for each ML algorithm. Table 2 represents the DT classifier's confusion matrix, and Table 3 represents the RF classifier's confusion matrix.

Table 4 represents the SVM classifier's confusion matrix, and Table 5 represents the NB classifier's confusion matrix.

To compare the four ML classifiers (DT, RF, SVM and NB) and determine which is the best, we need to compare the top significant rates for our proposed framework: Higher AC, lower FPR and lower FNR. For example, a classifier could be better in AC; however, the FNR is higher than the classifier. In this case, the decision will be based on the user's preferences and whether they need higher accuracy in general or prefer to reduce the chance of the application not giving a warning while the pregnant woman is already in labour.

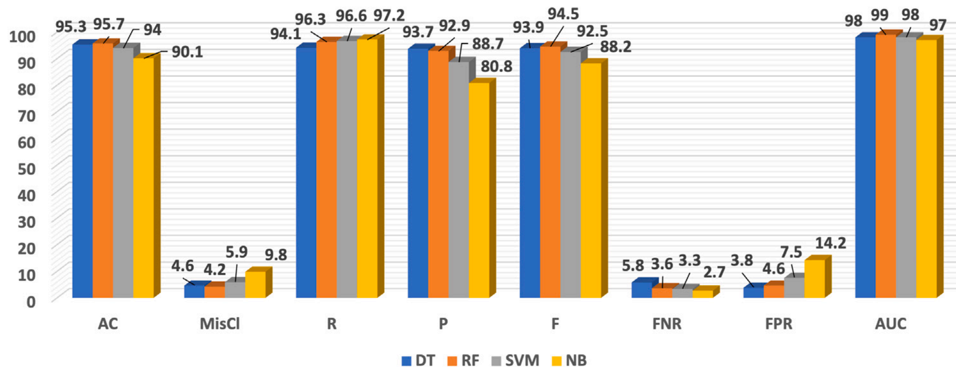

Fig. 1. Machine learning classifiers' results in comparison.

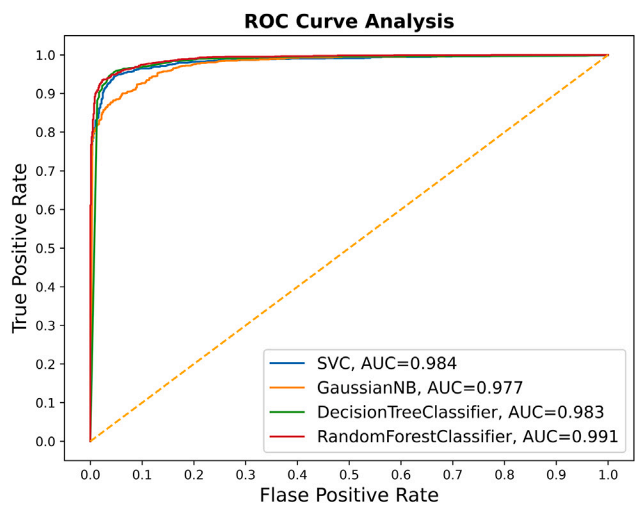

Fig. 2. All ML classifiers AUROC plot.

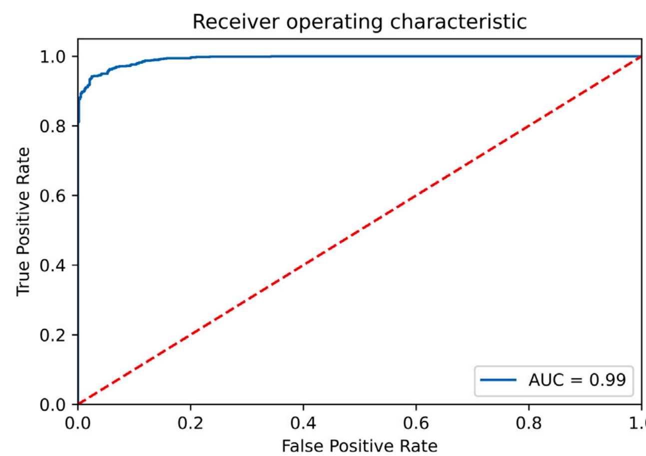

Fig. 3. ANN AUROC plot.

The RF performed the best in terms of the AC rate and the FPR rate among all the four machine learning classifiers. In terms of FNR rate, the NB classifier performed the best. Although NB has the lowest FNR (i.e. there is a labour pattern, but no warning is given), RF is the best choice for our framework because it has higher AC and lower FPR. Moreover, the FNR difference between RF and NB is insignificant, making the RF our best choice of the four classifiers. We highlighted the differences in FNR and FNR between the RF and NB algorithms since it is the only comparison in which NB is better than RF. Some other criteria, such as AC, are also insignificant; however, RF is better than NB in these criteria.

The chart in Fig. 1 visualizes the evaluation results. Moreover, Fig. 2 shows the AUROC plot for all the ML classifiers. Notice that the RF AUROC plot has the best curve.

4.1. ML Algorithms' results comparison with the literature review studies

We will compare our chosen RF classifier to the studies from the literature review. Table 6 shows the results of the RF classifier and the studies from the literature review.

As seen in the table, the studies did not provide some of the criteria. This makes the evaluation of their work more difficult since we cannot measure all the classifiers' performance results. For example, none of the studies have reported MisCl, P, FNR or FPR rates. In this case, we cannot compare our proposed work to the literature review studies in terms of labour EHG signals being categorized by the classifier as non-labour, for example, since there is no FNR reported.

In terms of AC rate, six studies are better than our proposed scheme. However, when we perform a deeper analysis of the studies' design and implementation, we can see some advantages of our scheme compared with these studies. For instance, none of the studies used the same number of databases we have used in our proposed scheme. Some of the studies used only one database. This could lead to a bias in the results due to a lack of diversity in the databases. The same concept applies to the R rate. In our datasets, the ratio is 60% pattern datasets to 40% non-pattern datasets, which implies no bias towards any dataset group in our datasets collection. Finally, our proposed scheme achieved the highest AUC among the literature review studies.

Next, Table 7 presents the confusion matrix table for the ANN DL model. Fig. 3 shows the AUROC plot for the ANN model. We aim to improve the prediction and recognition of labour by personalizing the needs of each pregnant woman. Moreover, we want to use deep learning to further improve the outcome of the framework. As seen from the results in Fig. 4, the ANN improved the prediction in all the criteria compared to the RF algorithm from the machine learning scheme. The AC has improved from 0.95 to 0.98, MisCl from 0.04 to 0.01, R from 0.96 to 0.98, P from 0.92 to 0.97, F from 0.94 to 0.98, FNR from 0.03 to 0.01 and FPR from 0.04 to 0.01. FNR and FPR, as among the top priority aspects in the framework, have been reduced by about 50% for the FNR and 75% for the FPR.

4.2. DL model results comparison with the literature review studies

We will compare the ANN model to the study from the literature review. Although the authors [49] used a different approach in using the CNN model for image recognition of uterine contractions, their

Table 6
Summary of the ML related work studies.

Proposed systems	AC	MisCl	R	P	F	FNR	FPR	AUC
Proposed RF classifier	0.95	0.04	0.96	0.93	0.94	0.03	0.04	0.99
Idowu et al. [29]	x	x	0.97	x	x	x	x	0.94
Fergus et al.[38]	0.90	x	0.85	x	x	x	x	0.90
Acharya et al. [39]	0.96	x	0.95	x	x	x	x	x
Altini et al. [40]	0.87	x	x	x	x	x	x	x
Borowska et al. [18]	0.83	x	x	x	x	x	x	x
Altini et al. [35]	x	x	x	x	x	x	x	x
Hoseinzadeh and Amirani [41]	0.97	x	0.95	x	x	x	x	x
Degbedzui and Yüksel [42]	0.99	x	0.99	x	0.99	x	x	x
Peng et al. [43]	0.93	x	0.89	x	x	x	x	0.80
Shahbakhti et al. [44]	0.99	x	0.98	x	x	x	x	x
Sheryl Oliver et al. [45]	0.96	x	0.92	x	x	x	x	x
Chen et al. [46]	0.90	x	0.92	x	x	x	x	x
Esgalhado et al. [48]	0.97	x	x	x	x	x	x	x

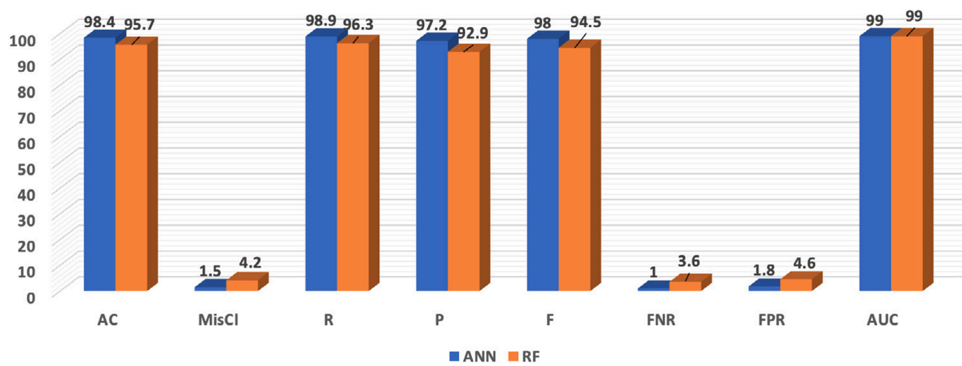

Fig. 4. Results comparison between the ANN model and RF classifier.

Table 7
ANN confusion matrix.

	Classified pattern	Classified no pattern
Pattern	839	9
No pattern	24	1306

goal is similar to our approach, which is to detect labour. Table 8 summarizes the comparison of the results between the two approaches.

Our ANN and the CNN model have the same AC rate. Furthermore, our ANN model has a significantly better R rate and AUC value. The authors of the CNN model have not provided information about the MisCl, P, F, FNR and FPR rates. Although both models have a similar AC rate, other critical criteria such as FNR and FPR are missing, making the complete comparison incomplete. Moreover, the R rate for the CNN model is below 90%. That means the probability of the model identifying a dataset as labour is lower than our ANN model by 11.2%. This further supports that we need more information on the CNN model's performance to analyse it comprehensively. Finally, although we achieved good results using the deep learning approach, the size of our data could make the problem too easy for the deep learning models. Deep learning is powerful and needs big data to produce excellent and reliable results. We need more data to confirm our results for the deep learning approach, which we will discuss later in the chapter. In the next section, we will discuss big data and the issue of synthesizing EHG data.

5. Big data and synthetic data

One of the aspects of deep learning is big data. It is one of the main differences that distinguishes it from machine learning. When it comes to medical data, obtaining big data on uterine EHG signals for pregnant women in general and pregnant women with a high risk of premature labour, in particular, is difficult. Moreover, the publicly available uterine EHG databases are limited. The vast majority of available databases, even by request, collect data using the intrauterine pressure catheter (IUPC) technique, which does not apply to our framework. For example, we contacted a pregnancy research group from Oxford University, and they agreed to share their big data with us; however, their data were on IUPC, not EHG. Furthermore, the lack of big data in our scheme is a limitation of our work. More data are needed to thoroughly utilize big data in our work.

To solve this issue, we proposed to generate our synthetic data using the five databases as seeds. We found Gretel Labs, Inc. (https://gretel.ai/), which provides such a service for free using Python. We were able to synthesize a few EHG datasets. However, after analysing the synthesized datasets and consulting with our deep learning expert, Dr. Jaume Manero, about the quality of the synthesized datasets, he recommended that we would not be able to verify if the synthesized datasets are similar to the actual datasets on pregnant women. It is difficult to produce high-quality bio-synthetic data that are similar to the original data. This could be a limitation of the company's algorithms since they do not specialize in synthesizing biodata. Dr. Manero has also recommended that if we need to synthesize uterine EHG data, we would have to build our own deep learning synthetic model.

6. Experts domain

In this section, we will present obstetricians' medical opinions on some aspects of our framework. We have contacted two obstetricians: Dr. Ebtehal Hussein Aljumaai, obstetric and gynaecology registrar, Abha Maternity Children's Hospital, Saudi Arabia; and Dr. Mohamed Elsheikh, obstetrician and foetal medicine consultant, National Guard Hospital, MRCOG, SSCOG, DIP Prenatal genetics, Riyadh, Saudi Arabia. To validate the medical aspects of the framework, we asked them the following questions:

Table 8
Summary of the DL related work studies.

Proposed systems	AC	MisCl	R	P	F	FNR	FPR	AUC
Proposed ANN model	0.98	0.01	0.98	0.97	0.98	0.01	0.01	0.99
CNN model [49]	0.98	x	0.87	x	x	x	x	0.92

1. Does the 30-minute monitoring session mimic the obstetrician's uterine contractions analysis to determine if a pregnant woman is in labour?
2. Are amplitude and frequency the most important parameters to use to decide if a pregnant woman is in labour?
3. Do you consider age, gestation and parity when evaluating the status of pregnant women with a high risk of premature labour?
4. Would you integrate our monitoring system to assist your patients?
5. How many stages does labour involve?
6. What is the window of time before the pregnant woman goes into labour?

In the following sections, we will provide the obstetricians' answers to these questions.

6.1. Question 1

The obstetrician will monitor the pregnant woman's contractions for a minimum of 30 to 40 min, then repeat the process. This process can take from 4 h to 24 h, depending on the pregnancy status. This supports our framework design, as the monitoring session is set to be 30 min long, and it is repeated if no alarm is triggered.

6.2. Question 2

Both obstetricians acknowledged that amplitude and frequency are the top vital parameters when analysing uterine contractions. That means our choice of having these two parameters as the basis for the machine learning and deep learning approaches is correct and mimics what an obstetrician would use to analyse uterine contractions.

Dr. Aljumaai also added that if the contraction pattern is regular or the amplitude is 60 millimetres of mercury (mmHg) or 200 Montevideo units (MVU) for IUPC, the pregnant woman is in labour. This is similar to our amplitude–frequency scheme, as we designed it to recognize regular contraction patterns.

6.3. Question 3

Of the three parameters, age is the most important factor, especially if the pregnant woman's age is less than 18 or over 40. Gestation and parity are also important, especially if the pregnant woman has a history of premature labour.

6.4. Question 4

Both obstetricians agreed that such a system is needed, and they would use it with their patients.

6.5. Question 5

There are four stages of labour:

1. Stage 1 (Latent): When the contractions start until the cervix widens to 6 cm.
2. Stage 2 (Active): When the cervix widens to 10 cm.
3. Stage 3: Starts when the baby is born until the placenta detaches.
4. Stage 4 (Afterbirth): Starts after the placenta comes out.

According to the obstetricians, our framework comes before stage 1 when the contractions start.

6.6. Question 6

The window before the pregnant woman goes into stage 1 of labour depends on how many pregnancies the woman has had (parity). If the pregnant woman is primigravida (it is her first pregnancy), the window is around 16 h. If the pregnant woman is multigravida (it is her at least second pregnancy), the window is around 12 to 14 h.

In the next section, we will present the concluding remarks and future work on the framework.

7. Conclusion and future work

Premature birth is a global issue for pregnant women and babies that can lead to death or life-long health problems. Furthermore, there is no permanent solution or cure for premature birth because the cause is usually unknown. We have proposed a framework for monitoring pregnant women with a high risk of premature birth using a wireless body sensor (WBS) and a smartphone. We designed the framework to be continuous, home-comfortable, cost-effective, and reliable. The framework aims to detect labour using ML and DL algorithms and send a warning to the pregnant woman by monitoring uterine EHG contractions.

We used 7271 datasets of uterine EHG contractions obtained from the physioBank depository [37] to evaluate the framework for accuracy and reliability. The deep learning ANN model achieved better results in comparison to the machine learning RF classifier. In the next section, we will list the framework's future work.

7.1. Future work

In future work, we will expand our smartphone usage to include iPhones. We will also develop a deep learning model to synthesize EHG signals. Furthermore, we will design and implement a user study to recruit pregnant women for EHG data collection.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The first author gratefully acknowledges the support of Jouf University, Saudi Arabia for funding his Ph.D. programme, which made the study reported in this paper possible. The authors also thank Dr. Jaume Manero, from Universitat Politècnica de Catalunya, Spain; Dr. Ebtehal Hussein Aljumaai, Obstetric and Gynaecology Registrar, Abha Maternity Children's Hospital, Saudi Arabia; Dr. Mohamed Elsheikh, obstetrician and foetal medicine consultant, National Guard Hospital, MRCOG, SSCOG, DIP Prenatal Genetics, Riyadh, Saudi Arabia; and Darshana Upadhyay, MCS, Faculty of Computer Science, Dalhousie University, Canada, for their help and support.

Table B.9
Hyper-parameters definitions.

Hyper-parameter	Definition	Values
criterion	The function is used to measure the quality of the split.	"gini" or "entropy'. default = "gini"
max_depth	Represents the maximum depth of the tree.	int. default = None
min_samples_split	The minimum number of samples needed to split an internal node	int, float. default = 2
min_samples_leaf	The minimum number of samples required to be at a leaf node (i.e. a node that has no children).	int or float. default = 1
random_state	Controls the randomness of the estimator.	int, RandomState instance or None. default = None
n_estimators	The number of trees in the random forest.	int. default = 100
kernel	The kernel to be used in the algorithm.	"linear", "poly", "rbf", "sigmoid" or "precomputed". default = "rbf"
C	The regularization parameter.	float. default = 1.0
gamma	The kernel coefficient for 'rbf', 'poly'and 'sigmoid'.	"scale", "auto" or float. default = "scale"
probability	To enable or disable the probability estimates.	bool. default = False

List of acronyms

AC	accuracy
ANN	artificial neural network
AUC	area under the ROC curve
AUROC	area under the receiver operator curve
CNN	convolutional neural network
DL	deep learning
DT	decision tree
EHG	electrohysterography
ELM	extreme learning machine
EMD	empirical mode decomposition
F	F score
FN	false negative
FNR	false negative rate
FP	false positive
FPR	false positive rate
GLMs	generalized linear machine learning models
HR	heart rate
IMF	intrinsic mode function
IUPC	intrauterine pressure catheter
MDF	median frequency
MisCl	misclassification rate
ML	machine learning
mmHg	millimetres of mercury
MNF	mean frequency
MVU	Montevideo units
NB	Naïve Bayes
OS	operating system
P	precision
PCA	principal component analysis
PF	peak frequency
R	recall
RBNC	Radial Basis Function Neural Network
RF	random forest
RMS	root mean square
ROC	receiver operating characteristic
RQA	recurrence quantification analysis
SSAE	stacked sparse autoencoder
SVM	support vector machine
TP	true positive
TPR	true positive rate
TN	true negative
WPD	wavelet packet decomposition
WBS	wireless body sensor

Appendix A. Data preprocessing

The following equation is from the **rdsamp** tool to read the physioBank databases [37]:

$$rdsamp\ -r\ record\ -c\ -H\ -f\ 0\ -v\ -pe\ > output\_file.csv \tag{11}$$

Where:

- **-r** reads the record.
- **-c** produces output in CSV.
- **-H** reads the signal files in high-resolution mode.
- **-f 0** begins at the specified time (0 means at the beginning of the record).
- **-v** prints column headings.
- **-pe** prints the elapsed time from the beginning of the record as hh:mm:ss.

Appendix B. Machine learning classifiers' hyper-parameters

Table B.9 defines the hyper-parameters used in the classifiers.

B.1. The DT classifier

```
criterion = 'entropy', max_depth = 8,
min_samples_split = 5, min_samples_leaf=10
random_state = 0
```

B.2. The RF classifier

```
n_estimators = 100, criterion = 'entropy', max_depth
= 10, min_samples_leaf = 1, min_samples_split = 15,
random_state = 2
```

B.3. The SVM classifier

```
C = 1000000, kernel = 'rbf', gamma = 1000,
random_state = 0, probability=True
```

B.4. The NB classifier

```
GaussianNB()
```

References

[1] Mischi M, Rabotti C, Vosters LP, Oei SG, Bergmans JW. Electrohysterographic conduction velocity estimation. In: Proceedings of the 31st annual international conference of the IEEE engineering in medicine and biology society: engineering the future of biomedicine, EMBC 2009. IEEE; 2009, p. 6934–7. http://dx.doi.org/10.1109/IEMBS.2009.5333636.

[2] Garcia-Gonzalez MT, Charleston-Villalobos S, Vargas-Garcia C, Gonzalez-Camarena R, Aljama-Corrales T. Characterization of EHG contractions at term labor by nonlinear analysis. In: proceedings of the annual international conference of the IEEE engineering in medicine and biology society, EMBS. IEEE; 2013, p. 7432–5. http://dx.doi.org/10.1109/EMBC.2013.6611276.

[3] Hassan M, Terrien J, Karlsson B, Marque C. Application of wavelet coherence to the detection of uterine electrical activity synchronization in labor. Irbm 2010;31(3):182–7. http://dx.doi.org/10.1016/j.irbm.2009.12.004, URL http://dx.doi.org/10.1016/j.irbm.2009.12.004.

[4] Shahrdad M, Amirani MC. Detection of preterm labor by partitioning and clustering the EHG signal. Biomed Signal Process Control 2018;45:109–16. http://dx.doi.org/10.1016/j.bspc.2018.05.044, URL https://doi.org/10.1016/j.bspc.2018.05.044.

[5] Nader N, Hassan M, Falou W, Diab A, Al-Omar S, Khalil M, Marque C. Classification of pregnancy and labor contractions using a graph theory based analysis. In: Proceedings of the annual international conference of the IEEE engineering in medicine and biology society, EMBS. IEEE; 2015, p. 2876–9. http://dx.doi.org/10.1109/EMBC.2015.7318992.

[6] Hassan M, Terrien J, Muszynski C, Alexandersson A, Marque C, Karlsson B. Better pregnancy monitoring using nonlinear correlation analysis of external uterine electromyography. IEEE Trans Biomed Eng 2013;60(4):1160–6. http://dx.doi.org/10.1109/TBME.2012.2229279.

[7] Batista AG, Najdi S, Godinho DM, Martins C, Serrano FC, Ortigueira MD, et al. A multichannel time–frequency and multi-wavelet toolbox for uterine electromyography processing and visualisation. Comput Biol Med 2016;76:178–91. http://dx.doi.org/10.1016/j.compbiomed.2016.07.003, URL http://dx.doi.org/10.1016/j.compbiomed.2016.07.003.

[8] Sadi-Ahmed N, Kacha B, Taleb H, Kedir-Talha M. Relevant features selection for automatic prediction of preterm deliveries from pregnancy ElectroHysterograhic (EHG) records. J Med Syst 2017;41(12). http://dx.doi.org/10.1007/s10916-017-0847-8.

[9] Yochum M, Laforêt J, Marque C. An electro-mechanical multiscale model of uterine pregnancy contraction. Comput Biol Med 2016;77:182–94. http://dx.doi.org/10.1016/j.compbiomed.2016.08.001, URL http://dx.doi.org/10.1016/j.compbiomed.2016.08.001.

[10] World Health Organization. Preterm birth. 2015, URL http://www.who.int/mediacentre/factsheets/fs363/en/, [Accessed 23 August 2016].

[11] López Bernal A. Overview. Preterm labour: mechanisms and management. BMC Pregnancy Childbirth 2007;7(Suppl 1):S2. http://dx.doi.org/10.1186/1471-2393-7-S1-S2, URL http://bmcpregnancychildbirth.biomedcentral.com/articles/10.1186/1471-2393-7-S1-S2.

[12] Vogel JP, Chawanpaiboon S, Watananirun K, Lumbiganon P, Petzold M, Moller AB, et al. Global, regional and national levels and trends of preterm birth rates for 1990 to 2014: Protocol for development of world health organization estimates. Reprod Health 2016;13(1):1–9. http://dx.doi.org/10.1186/s12978-016-0193-1, URL http://dx.doi.org/10.1186/s12978-016-0193-1.

[13] Ganer Herman H, Miremberg H, Dekalo A, Barda G, Bar J, Kovo M. Preterm uterine contractions ultimately delivered at term: Safe but not out of danger. Eur J Obstet Gynaecol Reprod Biol 2016;199:1–4. http://dx.doi.org/10.1016/j.ejogrb.2016.01.019, URL http://dx.doi.org/10.1016/j.ejogrb.2016.01.019.

[14] Nader N, Marque C, Hassan M, Falou W, Diab A, Khalil M. Pregnancy monitoring using graph theory based analysis. In: 2015 international conference on advances in biomedical engineering, ICABME 2015. IEEE; 2015, p. 73–6. http://dx.doi.org/10.1109/ICABME.2015.7323254.

[15] Hussain AJ, Fergus P, Al-Askar H, Al-Jumeily D, Jager F. Dynamic neural network architecture inspired by the immune algorithm to predict preterm deliveries in pregnant women. Neurocomputing 2015;151(P3):963–74. http://dx.doi.org/10.1016/j.neucom.2014.03.087, URL http://dx.doi.org/10.1016/j.neucom.2014.03.087.

[16] Nader N, Hassan M, Falou W, Marque C, Khalil M. A node-wise analysis of the uterine muscle networks for pregnancy monitoring. In: Proceedings of the annual international conference of the IEEE engineering in medicine and biology society, EMBS. 2016, p. 712–5. http://dx.doi.org/10.1109/EMBC.2016.7590801.

[17] Zhang M, Tidwell V, La Rosa PS, Wilson JD, Eswaran H, Nehorai A. Modeling magnetomyograms of uterine contractions during pregnancy using a multiscale forward electromagnetic approach. PLoS ONE 2016;11(3):1–23. http://dx.doi.org/10.1371/journal.pone.0152421.

[18] Borowska M, Brzozowska E, Kuć P, Oczeretko E, Mosdorf R, Laudański P. Identification of preterm birth based on RQA analysis of electrohysterograms. Comput Methods Programs Biomed 2018;153:227–36. http://dx.doi.org/10.1016/j.cmpb.2017.10.018.

[19] Alamedine D, Khalil M, Marque C. Parameters extraction and monitoring in uterine EMG signals. Detection of preterm deliveries. Irbm 2013;34(4–5):322–5. http://dx.doi.org/10.1016/j.irbm.2013.08.003, URL http://dx.doi.org/10.1016/j.irbm.2013.08.003.

[20] American College of Obstetricians and Gynecologists. Preterm birth definition. 2016, URL http://www.acog.org/Patients/FAQs/Preterm-Premature-Labor-and-Birth, [Accessed 25 April 2017].

[21] Dyson DC, Danbe KH, Bamber JA, Crites YM, Field DR, Maier JA, et al. Monitoring women at risk for preterm labor. N Engl J Med 1998;338(1):15–9. http://dx.doi.org/10.1056/NEJM199801013380103, URL http://www.nejm.org/doi/full/10.1056/NEJM199801013380103%5Cnhttp://www.nejm.org/doi/pdf/10.1056/NEJM199801013380103.

[22] La Rosa PS, Eswaran H, Preissl H, Nehorai A. Multiscale forward electromagnetic model of uterine contractions during pregnancy. BMC Med Phys 2012;12:4. http://dx.doi.org/10.1186/1756-6649-12-4, URL http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3605117&tool=pmcentrez&rendertype=abstract.

[23] Kuon R-J, Garfield RE. Actions of progestins for the inhibition of cervical ripening and uterine contractions to prevent preterm birth. Facts Views Vision ObGyn 2012;4(2):110–9, URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3987504/.

[24] Rozenberg P, Chauveaud A, Deruelle P, Capelle M, Winer N, Desbrire R, et al. Prevention of preterm delivery after successful tocolysis in preterm labor by 17 alpha-hydroxyprogesterone caproate: A randomized controlled trial. Am J Obstet Gynecol 2012;206(3):206.e1–9. http://dx.doi.org/10.1016/j.ajog.2011.12.026, URL http://dx.doi.org/10.1016/j.ajog.2011.12.026.

[25] Lemancewicz A, Borowska M, Kuć P, Jasińska E, Laudański P, Laudański T, et al. Early diagnosis of threatened premature labor by electrohysterographic recordings - the use of digital signal processing. Biocybern Biomed Eng 2016;36(1):302–7. http://dx.doi.org/10.1016/j.bbe.2015.11.005, http://arxiv.org/abs/1312.1079.

[26] Mas-Cabo J, Ye-Lin Y, Garcia-Casado J, Alberola-Rubio J, Perales A, Prats-Boluda G. Uterine contractile efficiency indexes for labor prediction: A bivariate approach from multichannel electrohysterographic records. Biomed Signal Process Control 2018;46:238–48. http://dx.doi.org/10.1016/j.bspc.2018.07.018, URL https://doi.org/10.1016/j.bspc.2018.07.018.

[27] Horoba K, Jezewski J, Matonia A, Wrobel J, Czabanski R, Jezewski M. Early predicting a risk of preterm labour by analysis of antepartum electrohysterograhic signals. Biocybern Biomed Eng 2016;36(4):574–83. http://dx.doi.org/10.1016/j.bbe.2016.06.004, URL http://dx.doi.org/10.1016/j.bbe.2016.06.004.

[28] Karsdon J, El Daouk M, Huang WM, Ashmead GG. Electrical pacemaker as a safe and feasible method for decreasing the uterine contractions of human preterm labor. J Perinat Med 2012;40(6):697–700. http://dx.doi.org/10.1515/jpm-2012-0136.

[29] Idowu IO, Fergus P, Hussain A, Dobbins C, Khalaf M, Eslava RV, et al. Artificial intelligence for detecting preterm uterine activity in gynacology and obstertric care. In: Proceedings - 15th IEEE international conference on computer and information technology, CIT 2015, 14th IEEE international conference on ubiquitous computing and communications, IUCC 2015, 13th IEEE international conference on dependable, autonomic and se. IEEE; 2015, p. 215–20. http://dx.doi.org/10.1109/CIT/IUCC/DASC/PICOM.2015.31.

[30] Nicholl J, West J, Goodacre S, Turner J. The relationship between distance to hospital and patient mortality in emergencies: An observational study. Emergency Med J 2007;24(9):665–8. http://dx.doi.org/10.1136/emj.2007.047654.

[31] Bastos LF, van Meurs W, Ayres-de Campos D. A model for educational simulation of the evolution of uterine contractions during labor. Comput Methods Programs Biomed 2012;107(2):242–7. http://dx.doi.org/10.1016/j.cmpb.2011.09.016, https://linkinghub.elsevier.com/retrieve/pii/S016926071100263X.

[32] Vora SA, Montgomery OC, Kurzweg TP. Method for uterine contraction monitoring with passive RFID tags. In: 2017 IEEE EMBS international conference on biomedical and health informatics, BHI 2017. 2017, p. 213–6. http://dx.doi.org/10.1109/BHI.2017.7897243.

[33] Vermeulen-Giovagnoli B, Peters C, van der Hout-van der Jagt MB, Mischi M, van Pul C, Cottaar EJE, et al. The development of an obstetric tele-monitoring system. In: 2015 37th annual international conference of the IEEE engineering in medicine and biology society (EMBC). IEEE; 2015, p. 177–80. http://dx.doi.org/10.1109/EMBC.2015.7318329, URL http://ieeexplore.ieee.org/document/7318329/.

[34] Bajlekov GI, Rabotti C, Oei SG, Mischi M. [7]-Electrohysterographic detection of uterine contractions in term pregnancy. Embc 2015;5851–4.

[35] Altini M, Rossetti E, Rooijakkers MJ, Penders J. Towards non-invasive labour detection: A free-living evaluation. In: Proceedings of the annual international conference of the IEEE engineering in medicine and biology society, EMBS. 2018, p. 2841–4. http://dx.doi.org/10.1109/EMBC.2018.8512964.

[36] Fele-Žorž G, Kavšek G, Novak-Antolič Ž, Jager F. A comparison of various linear and non-linear signal processing techniques to separate uterine EMG records of term and pre-term delivery groups. Med Biol Eng Comput 2008;46(9):911–22. http://dx.doi.org/10.1007/s11517-008-0350-y.

[37] Goldberger AL, Amaral LAN, Glass L, Hausdorff JM, Ivanov PC, Mark RG, et al. Physiobank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. Circulation 2000;101(23):e215–20.

http://dx.doi.org/10.1161/01.CIR.101.23.e215, http://arxiv.org/abs/http://circ.ahajournals.org/content/101/23/e215.full.pdf.

[38] Fergus P, Idowu I, Hussain A, Dobbins C. Advanced artificial neural network classification for detecting preterm births using EHG records. Neurocomputing 2016;188:42–9. http://dx.doi.org/10.1016/j.neucom.2015.01.107.

[39] Acharya UR, Sudarshan VK, Rong SQ, Tan Z, Lim CM, Koh JE, et al. Automated detection of premature delivery using empirical mode and wavelet packet decomposition techniques with uterine electromyogram signals. Comput Biol Med 2017;85(April):33–42. http://dx.doi.org/10.1016/j.compbiomed.2017.04.013, URL http://dx.doi.org/10.1016/j.compbiomed.2017.04.013.

[40] Altini M, Rossetti E, Rooijakkers M, Penders J, Lanssens D, Grieten L, et al. Combining electrohysterography and heart rate data to detect labour. In: 2017 IEEE EMBS international conference on biomedical and health informatics, BHI 2017. IEEE; 2017, p. 149–52. http://dx.doi.org/10.1109/BHI.2017.7897227.

[41] Hoseinzadeh S, Amirani MC. Use of electro hysterogram (EHG) signal to diagnose preterm birth. In: 26th iranian conference on electrical engineering (ICEE2018). IEEE; 2018, p. 1477–81. http://dx.doi.org/10.1109/ICEE.2018.8472416, URL https://ieeexplore.ieee.org/document/8472416/.

[42] Degbedzui DK, Yüksel ME. Accurate diagnosis of term–preterm births by spectral analysis of electrohysterography signals. Comput Biol Med 2020;119(February):103677. http://dx.doi.org/10.1016/j.compbiomed.2020.103677, URL https://doi.org/10.1016/j.compbiomed.2020.103677.

[43] Peng J, Hao D, Yang L, Du M, Song X, Jiang H, et al. Evaluation of electrohysterogram measured from different gestational weeks for recognizing preterm delivery: a preliminary study using random forest. Biocybern Biomed Eng 2020;40(1):352–62. http://dx.doi.org/10.1016/j.bbe.2019.12.003.

[44] Shahbakhti M, Beiramvand M, Bavi MR, Mohammadi Far S. A new efficient algorithm for prediction of preterm labor. In: Proceedings of the annual international conference of the IEEE engineering in medicine and biology society, EMBS. IEEE; 2019, p. 4669–72. http://dx.doi.org/10.1109/EMBC.2019.8857837.

[45] Sheryl Oliver A, Ashwanthika U, Aswitha R. Labor prediction in pregnant woman based on ElectroMyoGram and ElectroHysteroGram using machine learning techniques. In: 2020 fourth international conference on inventive systems and control (ICISC). 2020, p. 77–82. http://dx.doi.org/10.1109/ICISC47916.2020.9171128, URL https://ieeexplore.ieee.org/abstract/document/9171128.

[46] Chen L, Hao Y, Hu X. Detection of preterm birth in electrohysterogram signals based on wavelet transform and stacked sparse autoencoder. PLoS ONE 2019;14(4):1–17. http://dx.doi.org/10.1371/journal.pone.0214712, URL http://dx.doi.org/10.1371/journal.pone.0214712.

[47] Alexandersson A, Steingrimsdottir T, Terrien J, Marque C, Karlsson B. The icelandic 16-electrode electrohysterogram database. Sci Data 2015;2. 150017 EP, URL http://dx.doi.org/10.1038/sdata.2015.17.

[48] Esgalhado F, Batista AG, Mouriño H, Russo S, dos Reis CRP, Serrano F, et al. Automatic contraction detection using uterine electromyography. Appl Sci 2020;10(20). http://dx.doi.org/10.3390/app10207014, URL https://www.mdpi.com/2076-3417/10/20/7014.

[49] Hao D, Peng J, Wang Y, Liu J, Zhou X, Zheng D. Evaluation of convolutional neural network for recognizing uterine contractions with electrohysterogram. Comput Biol Med 2019;113(August):103394. http://dx.doi.org/10.1016/j.compbiomed.2019.103394, URL https://doi.org/10.1016/j.compbiomed.2019.103394.

[50] Garfield RE, Maner WL. Physiology and electrical activity of uterine contractions. Sem Cell Dev Biol 2007;18(3):289–95. http://dx.doi.org/10.1016/j.semcdb.2007.05.004.

[51] Allahem H, Sampalli S. Automated uterine contractions pattern detection framework to monitor pregnant women with a high risk of premature labour. Inform Med Unlocked 2020;20:100404. http://dx.doi.org/10.1016/j.imu.2020.100404, URL https://doi.org/10.1016/j.imu.2020.100404.

[52] Jager F, Libenšek S, Geršak K. Characterization and automatic classification of preterm and term uterine records. PLoS ONE 2018;13(8):1–49. http://dx.doi.org/10.1371/journal.pone.0202125.

[53] Chudáček V, Spilka J, Burša M, Janků P, Hruban L, Huptych M, et al. Open access intrapartum CTG database. BMC Pregnancy Childbirth 2014;14(1). http://dx.doi.org/10.1186/1471-2393-14-16.

[54] Rabotti C, Mischi M, Van Laar J, Oei G, Bergmans J. On the propagation analysis of electrohysterographic signals. In: Conference proceedings of the international conference of IEEE engineering in medicine and biology society. 2008, p. 3868–71. http://dx.doi.org/10.1109/iembs.2008.4650054, URL http://www.ncbi.nlm.nih.gov/pubmed/19163557.

[55] Alberola-Rubio J, Prats-Boluda G, Ye-Lin Y, Valero J, Perales A, Garcia-Casado J. Comparison of non-invasive electrohysterographic recording techniques for monitoring uterine dynamics. Med Eng Phys 2013;35(12):1736–43. http://dx.doi.org/10.1016/j.medengphy.2013.07.008, URL http://dx.doi.org/10.1016/j.medengphy.2013.07.008.

[56] Léman H, Marque C, Gondry J. Use of the electrohysterogram signal for characterization of contractions during pregnancy. IEEE Trans Biomed Eng 1999;46(10):1222–9. http://dx.doi.org/10.1109/10.790499, http://arxiv.org/abs/NIHMS150003.

[57] Euliano TY, Nguyen MT, Darmanjian S, McGorray SP, Euliano N, Onkala A, et al. Monitoring uterine activity during labor: A comparison of 3 methods. Am J Obstet Gynecol 2013;208(1):66.e1–6. http://dx.doi.org/10.1016/j.ajog.2012.10.873, http://arxiv.org/abs/NIHMS150003.

[58] Taralunga DD, Ungureanu M, Hurezeanu B, Gussi I, Strungaru R. Empirical mode decomposition applied for non-invasive electrohysterograhic signals denoising. In: Proceedings of the annual international conference of the IEEE engineering in medicine and biology society, EMBS. IEEE; 2015, p. 4134–7. http://dx.doi.org/10.1109/EMBC.2015.7319304.

[59] Rabotti C, Mischi M, Van Laar J, Oei G, Bergmans J. Electrohysterographic analysis of uterine contraction propagation with labor progression: A preliminary study. In: Annual international conference of the IEEE engineering in medicine and biology - proceedings. 2007, p. 4135–8. http://dx.doi.org/10.1109/IEMBS.2007.4353246.

[60] Jacod BC, Graatsma EM, Van Hagen E, Visser GHA. A validation of electrohysterography for uterine activity monitoring during labour. J Mater-Fetal Neonatal Med 2010;23(1):17–22. http://dx.doi.org/10.3109/14767050903156668.

[61] Phinyomark A, Scheme E. EMG pattern recognition in the era of big data and deep learning. Big Data Cogn Comput 2018;2(3):21. http://dx.doi.org/10.3390/bdcc2030021.

[62] Gavankar SS, Sawarkar SD. Eager decision tree. In: 2017 2nd international conference for convergence in technology, I2CT 2017. 2017, p. 837–40. http://dx.doi.org/10.1109/I2CT.2017.8226246.

[63] Kabiraj S, Raihan M, Alvi N, Afrin M, Akter L, Sohagi SA, et al. Breast cancer risk prediction using xgboost and random forest algorithm. In: 2020 11th international conference on computing, communication and networking technologies, ICCCNT 2020. 2020, p. 1–4. http://dx.doi.org/10.1109/ICCCNT49239.2020.9225451.

[64] Ghosh S, Banerjee C. A predictive analysis model of customer purchase behavior using modified random forest algorithm in cloud environment. In: 2020 IEEE international conference for convergence in engineering, ICCE 2020 - proceedings. 2020, p. 239–44. http://dx.doi.org/10.1109/ICCE50343.2020.9290700.

[65] Ding SF, Liu XL, Zhang LW. Research on ranking support vector machine and prospects. In: Proceedings of the 29th chinese control conference, CCC'10. (9):IEEE; 2010, p. 2829–31.

[66] Braun AC, Weidner U, Hinz S. Support vector machines, import vector machines and relevance vector machines for hyperspectral classification - a comparison. 2, (3):2011, p. 36–9.

[67] Lu S-x, Meng J, Cao G-e. Support vector machine based on a new reduced samples method. In: 2010 international conference on machine learning and cybernetics. Vol. 3, IEEE; 2010, p. 1510–4. http://dx.doi.org/10.1109/ICMLC.2010.5580828, URL http://ieeexplore.ieee.org/document/5580828/.

[68] Balasundaram S, Kapil N. Application of Lagrangian twin support vector machines for classification. In: 2010 second international conference on machine learning and computing. IEEE; 2010, p. 193–7. http://dx.doi.org/10.1109/ICMLC.2010.40, URL http://ieeexplore.ieee.org/document/5460743/.

[69] Jahromi AH, Taheri M. A non-parametric mixture of Gaussian naive Bayes classifiers based on local independent features. In: 19th CSI international symposium on artificial intelligence and signal processing, AISP 2017. (1):2018, p. 209–12. http://dx.doi.org/10.1109/AISP.2017.8324083.

[70] Ma TM, Yamamori K, Thida A. A comparative approach to Naïve Bayes classifier and support vector machine for email spam classification. In: 2020 IEEE 9th global conference on consumer electronics, GCCE 2020. 2020, p. 324–6. http://dx.doi.org/10.1109/GCCE50665.2020.9291921.

[71] Rout AR, Bagal SB. Natural scene classification using deep learning. In: 2017 international conference on computing, communication, control and automation, ICCUBEA 2017. IEEE; 2018, http://dx.doi.org/10.1109/ICCUBEA.2017.8463727.

[72] Chollet F. Deep learning with python. 1st ed. USA: Manning Publications Co.; 2017.

[73] Majumder M. Artificial neural network. 2015, p. 49–54. http://dx.doi.org/10.1007/978-981-4560-73-3_3.

[74] Mesrabadi HA, Faez K. Improving early prostate cancer diagnosis by using artificial neural networks and deep learning. In: Proceedings - 2018 4th iranian conference of signal processing and intelligent systems, ICSPIS 2018. IEEE; 2018, p. 39–42. http://dx.doi.org/10.1109/ICSPIS.2018.8700542.

[75] Sammut C, Webb GI, editors. Encyclopedia of machine learning. Boston, MA: Springer US; 2010, http://dx.doi.org/10.1007/978-0-387-30164-8.

[76] Fawcett T. An introduction to ROC analysis. Pattern Recognit Lett 2006;27(8):861–74. http://dx.doi.org/10.1016/j.patrec.2005.10.010, URL https://www.sciencedirect.com/science/article/pii/S016786550500303X, ROC Analysis in Pattern Recognition.