Neurocomputing 151 (2015) 963–974

Contents lists available at ScienceDirect

Neurocomputing

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

Dynamic neural network architecture inspired by the immune algorithm to predict preterm deliveries in pregnant women

A.J. Hussain [a,*], P. Fergus [a], H. Al-Askar [a], D. Al-Jumeily [a], F. Jager [b]

[a] School of Computing and Mathematical Sciences, Liverpool John Moores University, Byrom Street, Liverpool L3 3AF, UK
[b] Faculty of Computer and Information Science, University of Ljubljana, Congress trg 12, 1000 Ljubljana, Slovenia

ARTICLE INFO

Article history:
Received 22 November 2013
Received in revised form
5 March 2014
Accepted 7 March 2014
Communicated by D.-S. Huang
Available online 4 November 2014

Keywords:
Term delivery
Preterm delivery
Machine learning
Classification
Neural networks
Electrohysterography

ABSTRACT

There has been some improvement in the treatment of *preterm* infants, which has helped to increase their chance of survival. However, the rate of premature births is still globally increasing. As a result, this group of infants is most at risk of developing severe medical conditions that can affect the respiratory, gastrointestinal, immune, central nervous, auditory and visual systems. There is a strong body of evidence emerging that suggests the analysis of uterine electrical signals, from the abdominal surface (Electrohysterography – EHG), could provide a viable way of diagnosing true labour and even predict *preterm* deliveries. This paper explores this idea further and presents a new dynamic self-organized network immune algorithm that classifies *term* and *preterm* records, using an open dataset containing 300 records (38 *preterm* and 262 *term*). Using the dataset, oversampling and cross validation techniques are evaluated against other similar studies. The proposed approach shows an improvement on existing studies with 89% *sensitivity*, 91% *specificity*, 90% *positive predicted value*, 90% *negative predicted value*, and an overall accuracy of 90%.

1. Introduction

Preterm birth, also known as premature birth or delivery, is described by the World Health Organisation (WHO) as the delivery of babies who are born, alive, before 37 weeks of gestation [1]. In contrast, *term* births are the live delivery of babies after 37 weeks, and before 42 weeks. According to the WHO, worldwide in 2010, *preterm* deliveries accounted for 1 in 10 births [1]. In 2009, in England and Wales, 7% of live births were also *preterm*.[1] *Preterm* birth has a significant adverse effect on the newborn, including an increased risk of death and health defects. The severity of these effects increases the more premature the delivery is. Approximately, 50% of all perinatal deaths are caused by *preterm* delivery [2], with those surviving often suffering from afflictions, caused by the birth. These include impairments to hearing, vision, the lungs, the cardiovascular system and non-communicable diseases. Up to, 40% of survivors of extreme *preterm* delivery can also develop chronic lung disease [3]. In other cases, survivors suffer with neuro-developmental or behavioural defects, including cerebral palsy, motor, learning and cognitive impairments. In addition, *preterm* births also have a detrimental effect on families, the economy, and society. In 2009, the overall cost to the public sector, in England and Wales, was estimated to be nearly £2.95 billion [4]. However, developing a better understanding of preterm deliveries can help to create preventative strategies and thus positively mitigate, or even eradicate, the effects that preterm deliveries have on babies, families, and society and healthcare services.

Preterm births can occur for three different reasons. According to [2] approximately one-third are medically indicated or induced; delivery is brought forward for the best interest of the mother or baby. Another third occurs because the membranes rupture, prior to labour (*PPROM*). Lastly, spontaneous contractions (termed *preterm* labour or *PTL*) can develop. However, there is still a great deal of uncertainty about the level of risk each factor presents, and whether they are causes or effects. Nevertheless, in [2] some of the causes of *preterm* labour, which may or may not end in *preterm* birth, have been discussed. These include infection, over-distension, burst blood vessels, surgical procedures, illnesses and congenital defects of the mother's uterus and cervical weakness. Further studies have also found other risk factors for *PTL*/*PPROM* [7,8]. These include a previous *preterm* delivery (20%); the last two births have been *preterm* (40%), and multiple births (twin pregnancy carries a 50% risk). Other health and lifestyle factors also include cervical and uterine abnormalities, recurrent antepartum haemorrhage, illnesses and infections, any invasive procedure or surgery, underweight or obese mother,

* Corresponding author.
E-mail address: a.hussain@ljmu.ac.uk (A.J. Hussain).
[1] (Gestation-specific infant mortality in England and Wales, 2009, http://ons.gov.uk).

http://dx.doi.org/10.1016/j.neucom.2014.03.087
0925-2312/

ethnicity, social deprivation, long working hours/late night, alcohol and drug use, and folic acid deficiency.

As well as investigating *preterm* deliveries, several studies have also explored *preterm* labour (the stage that directly precedes the delivery). However, in spite of these studies, there is no internationally agreed definition of *preterm* labour.[2] Nonetheless, in practice, women who experience regular contractions, increased vaginal discharge, pelvic pressure and lower backache tend to show Threatening Preterm Labour (*TPL*). While this is a good measure, Mangham et al., suggest that clinical methods for diagnosing *preterm* labour are insufficient [4]. Following a medical diagnosis of *TPL*, only 50% of all women with *TPL* actually deliver, within seven days [2]. In support of this, McPheeters et al., carried out a similar study that showed 144 out of 234 (61.5%) women diagnosed with *preterm* labour went on to deliver at *term* [5]. This can potentially add significant costs, and unnecessary interventions, to prenatal care. In contrast, false-negative results mean that patients requiring admittance are turned away, but then go on to deliver prematurely [6].

Predicting *preterm* birth and diagnosing *preterm* labour clearly have important consequences, for both health and the economy. However, most efforts have concentrated on mitigating the effects of *preterm* birth. Nevertheless, since this approach remains costly [1], it has been suggested that prevention could yield better results [9]. Effective prediction of *preterm* births could contribute to improving prevention, through appropriate medical and lifestyle interventions. One promising method is the use of Electrohysterography (*EHG*). *EHG* measures electrical activity in the uterus, and is a specific form of electromyography (*EMG*), the measurement of such activity in muscular tissue. Several studies have shown that the *EHG* record may vary from woman to woman, depending on whether she is in true labour or false labour and whether she will deliver *term* or *preterm*. *EHG* provides a strong basis for objective predication and diagnosis of *preterm* birth.

Many research studies have used *EHG* for prediction or detection of true labour. In contrast, this paper focuses on using *EHG* classification to determine whether delivery will be *preterm* or *term*. This is achieved by using a new neural network posited in this paper which is evaluated against several existing machine-learning classifiers using an open dataset, containing 300 records (38 *preterm* and 262 *term*) [10]. A signal filter and pre-selected features that are suited to classifying *term* and *preterm* records are used to produce a feature set from raw signals and is used by all classifiers. The results show that selected classifiers outperform a number of approaches, used in many other studies.

The structure of the remainder, of this paper is as follows. Section 2 describes the underlying principles of Electrohysterography. Section 3 discusses feature extraction from Electrohysterography signals. Section 4 describes machine learning and its use in *term* and *preterm* classification, while Section 5 describes the approach taken in this paper. Section 6 describes the evaluation, whilst Section 7 discusses the results before the paper is concluded in Section 8.

2. Electrohysterography

Since the late 1930s, information on the electrical activity of the uterus has been known [11]. However, it has only been in the last 20 years that formal techniques, for recording this type of activity, have appeared.

In order to retrieve *EHG* signals, bipolar electrodes are adhered to the abdominal surface. These are spaced at a horizontal, or vertical, distance of 2.5–7 cm apart. Most studies, including [10], use four electrodes although one study utilizes two [12]. In a series of other studies, sixteen electrodes were used [13–18], and a high density grid of 64 small electrodes was used in [19]. The results show that *EHG* may vary from women to women. This is dependent on whether she is in true or false labour, and whether she will deliver at *term*, or prematurely.

A raw *EHG* signal results from the propagation of electrical activity, between cells in the myometrium (the muscular wall of the uterus). This signal measures the potential difference between the electrodes, in a time domain. The electrical signals are not propagated by nerve endings; however, the exact propagation mechanism is not clear [20]. Since the late 70s, one theory suggests that gap junctions are the mechanisms responsible. Nevertheless, more recently it has been suggested that interstitial cells, or stretch receptors, may be the cause of propagation [21]. Gap junctions are groups of proteins that provide channels of low electrical resistance between cells. In most pregnancies, the connections between gap junctions are sparse, although gradually increasing, until the last few days before labour. A specific pacemaker site has not been conclusively identified, although, due to obvious physiological reasons, there may be a generalized propagation direction, from the top to the bottom of the uterus [22].

The electrical signals, in the uterus, are 'commands' to contract. During labour, the position of the bursts, in an *EHG* signal, corresponds roughly with the bursts shown in a tocodynamometer or intrauterine pressure catheter (*IUPC*). Clinical practises use these devices to measure contractions. More surprisingly, distinct contraction-related, electrical uterine activity is present early on in pregnancy, even when a woman is not in true labour. Gondry et al. identified spontaneous contractions from *EHG* records as early as 19 weeks of gestation [23]. The level of activity is said to increase, as the time to deliver nears, but shoots up especially so, in the last three to four days, before delivery [24]. As the gestational period increases, the gradual increase in electrical activity is a manifestation of the body's preparation for the final act of labour and parturition. In preparation for full contractions, which are needed to create the force and synchronicity required for a sustained period of true labour, the body gradually increases the number of electrical connections (gap junctions), between cells. In turn, this produces contractions in training.

Before analysis or classification tasks, *EHG* signals in their raw form, need pre-processing. Pre-processing can include filtering, de-noising, wavelet shrinkage or transformation and automatic detection of bursts. Recently, studies have typically focused on filtering the *EHG* signals to allow a bandpass between 0.05 Hz and 16 Hz [25–29]. However, there are some that have taken filtered *EHG* recordings to as high as 50 Hz [20]. Nevertheless, using *EHG* with such a wide range of frequencies is not the recommended method, since more interference affects the signal.

3. Feature extraction from electrohysterography signals

The collection of raw *EHG* signals is always temporal. However, for analysis and feature extraction purposes, translation, into other domains, is possible. These include a frequency representation, via Fourier Transform, [16,29–31] and wavelet transform [25,28,31–34]. The advantage of frequency-related parameters is that they are less susceptible to signal quality variations, due to electrode placement or the physical characteristics of the subjects [27]. In order to calculate these parameters, a transform from the time domain is required, i.e., using a Fourier Transform of the signal. Still, further transformation is often required before the extraction of frequency parameters. In several studies reviewed, in order to obtain frequency parameters, Power Spectral Density (*PSD*) is used. *Peak*

[2] http://bestpractice.bmj.com/best-practice/monograph/1002/basics.html.

frequency is one of the features provided with the Term–Preterm ElectroHysteroGram (TPEHG) dataset. It describes the frequency of the highest peak in the *PSD*. Most studies focus on the *peak frequency* of the burst, and in both human and animal studies, they are said to be one of the most useful parameters for predicting true labour [35]. On the other hand, the study by [10] found *medium frequency* to be more helpful in determining whether delivery was going to be *term* or *preterm*.

Several studies have shown that *peak frequency* increases, as the time to delivery decreases; this usually occurs within 1–7 days of delivery [20,31,25,27,12,36]. In particular, the results in [29] show that there are, statistically, significant differences in the *mean* values of *peak frequency* and the *standard deviations* in *EHG* recordings taken during *term* labour (*TL*) and *term* non-labour (*TN*) and also between *preterm* labour (*PTL*) and *preterm* non-labour (*PTN*).

In comparison to *peak frequency*, the TPEHG study [10] found that *median frequency* displayed a more significant difference between *term* and *preterm* records. When considering all 300 records, the statistical significance was $p=0.012$ and $p=0.013$ for *Channel* 3 on the 0.3–3 Hz and 0.3–4 Hz filter, respectively. Furthermore, this significance ($p=0.03$) was also apparent when only considering early records (before 26 weeks of gestation), with the same 0.3–3 Hz filter on *Channel* 3. The TPEHG study [10] concluded that this might have been due to the enlargement of the uterus, during pregnancy, which would affect the position of electrode placement. The placement of the *Channel* 3 electrode was, approximately, always 3.5 cm below the navel. However, as pregnancy progressed, this would mean that the electrode would move further away from the bottom of the uterus (cervico-isthmic section). If a generalized pacemaker area actually exists, and it is at the cervico-isthmic section, then, as pregnancy progresses, its position would move further and further away from the electrode, resulting in a diminished record of the signal. Whether this explanation is true or not, the results of [10] show that the discriminating capability of *median frequency* is somehow diminished, after the 26th week.

Amplitude-related *EMG* parameters represent the uterine *EMG* signal power, or signal energy. However, a major limitation is that the differences in patients can easily affect these parameters. Patients may differ in the amount of fatty tissue they have, and the conductivity of the skin–electrode interface, which leads to differences in the attenuation of uterine signals [27,6,35]. Examples of amplitude-related parameters include *root mean square*, *peak amplitude* and *median amplitude*. To obtain the *root mean square*, the signal value of every sample, in the recording, is squared, summed together, and then divided by the number of samples before the square root is taken. *Root mean square* is statistically descriptive of the signal's amplitude. The peak amplitude of the *PSD* is, in general, the maximum amplitude of a signal.

Using the *Student's t-test*, [10] found that *root mean square* might be useful in distinguishing between whether the information was recorded early (before 26 weeks of gestation) or late (after 26 weeks). The results obtained are in agreement with [31], [20,37], who found that the amplitude of the power spectrum increased, just prior to delivery. This was despite only analysing the *root mean square* values, per burst, rather than the whole signal. In other studies they found that amplitude-related parameters did not display a significant relationship with gestational age or indicate a transition to delivery (within seven days) [26,24,29]. Some of these discrepancies may be due to the differences in the characteristics of the studies: [10] compared records before and after 26 weeks, whereas [26] only examined records after the 25th week; [30,36] studied rat pregnancy, in contrast to human pregnancy. The frequency band used in [31] and [20] was a much broader band than in other studies (0.3–50 Hz; no bandwidth given for [37]), and also, the studies by [30,36] measured per burst, whilst [26] measured the whole signal.

Meanwhile, the TPEHG study [10] could not find any significant difference in *root mean squares* between *preterm* and *term* records. However, [26] did find that the *root mean squares*, in *preterm* contractions, were higher ($17.5\text{mv} \pm 7.78$), compared to *term* contractions ($12.2\ \text{mV} \pm 6.25$; $p<0.05$). The results, from [26], could not find a correlation between *root mean squares* and the weeks left to delivery. Nevertheless, they do suggest that a greater *root mean square* value was, for the most part, a static symptom that indicated a woman's dispensation to give birth prematurely. They also found that the *root mean square* values, within each pregnancy, did increase within a few days of birth.

Overall, the results suggest that there is no significant difference in the amplitude-related parameters between *term* and *preterm* deliveries, when taken during labour, or close to it. However, there may be considerable differences earlier on in the pregnancy. This suggests that at the time of delivery any differences have equalised themselves.

Sample entropy measures the irregularity of a time series, of finite lengths. This method was introduced by [38] to measure complexity in cardiovascular and biological signals. The more unpredictable the time series is, within a signal recording, the higher its sample entropy is. The process is based on calculating the number of matches of a sequence, which lasts for m points, within a given margin r. The disadvantage of this technique is the requirement to select two parameters, m and r. However, *sample entropy* did show a statistical difference between *term* and *preterm* delivery information, recorded either before or after the 26th week of gestation, when using any of the filters but only using the signal from *Channel* 3 [10].

In this section, numerous studies have been discussed, such as [10,25]. These investigations have used statistical analysis to examine the differences in *EHG* parameters, as well as the potential of such parameters to allow discrimination between different classes. The next section builds on the generally agreed upon idea that features can be used to separate *term* and *preterm* groups; specifically, using *root mean square*, *median frequency*, *peak frequency* and *sample entropy*, in exploring several well-known classification algorithms, and their ability to separate *term* and *preterm* records.

4. Term and preterm classification

Computer algorithms, and visualization techniques, are fundamental in supporting the analysis of datasets. More recently, the medical domain has been using such techniques, extensively. One example of this is the Common Spatial Patterns (*CSP*) algorithm. This was proposed by Woon et al. and has been successfully used to study Alzheimer's [39]. In other studies, Latchoumane et al., examines electroencephalogram (*EEG*) signals using Multi-way Array Decomposition (*MAD*). This is a supervised learning process for evaluating multidimensional and multivariate data like *EEG* [40].

Multi-Layer Perceptrons (*MLP*) and Probabilistic Neural Networks (*PNNs*) have featured widely in research to process and analyse medical datasets. *MLPs* are feed-forward networks that work with back-propagation learning rules. *PNNs* are similar to *MLPs* and consist of three layers, an input layer, radial basis layer, and a competitive layer. This type of feed-forward network operates using the Parzen's Probabilistic Density Function (*PDF*). In terms of overall functionality, *PNN* networks perform slightly better than *PML* networks [41].

The primary goal of such algorithms is to extract meaning from potentially huge amounts of data. Their association with particular data characterizes these features, such as datasets that contain

data about neurodegenerative diseases. This has led to a great deal of work in feature extraction within datasets. One example of this is the Discrete Cosine Transform (*DCT*) algorithm that decreases the number of features and the computation time when processing signals. *DCT* is used to calculate the trapped zone, under the curve, in special bands [42].

New neural network models have also emerged, called Self-organized Network Inspired by the Immune Algorithm (SONIA) [19] which is a single hidden layer neural network that uses a self-organization hidden layer inspired by the immune and back-propagation algorithms for the training of the output layer. The SONIA network was first proposed to improve the generalization and recognition capability that was lacking in back-propagation neural networks. The SONIA network has been used in financial time series prediction [45,46], and the experimental results in this paper will show that the SONIA network can be applied successfully to predict *term* and *preterm* records. Extensions to the SONIA network have been reported in other studies [47,48].

Using *Decision Trees*, *Naïve Bayes* and *Neural Networks*, similar algorithms have been used to predict heart disease. The results show that, using the lift chart for prediction and non-prediction, the *Naïve Bayes* algorithm predicted more heart disease patients than both the *Neural Network* and *Decision Tree* approaches [43]. Using data collected from patients suffering with Alzheimer's, Joshi et al., identified the various stages of Alzheimer's. This was achieved using *neural networks*, *multilayer perceptrons*, including the *coactive neuro-fuzzy inference system* (*CANFIS*) and *Genetic Algorithms* [44]. The results showed that *CANFIS* produced the best classification accuracy result (99.55%) when compared with *C*4.5 (a decision tree algorithm).

Other algorithms, such as dissimilarity based classification techniques, have proven to be very useful for analysing datasets. For example, the classification of seismic signals has been extensively explored using algorithms such as the *k-nearest neighbour* classifier (*k-NN*), and *Linear* and *Quadratic* normal density based classifiers. However, when there are a large number of prototypes, the results have shown that Bayesian (normal density based) classifiers outperform the *k-NN* classifier.

Meanwhile, Artificial Neural Networks (*ANN*) have been used in a large number of studies to classify *term* and *preterm* deliveries [45,46,12]. In other works, they have helped to distinguish between *non-labour* and *labour* events [46,12], irrespective of whether they were *term* or *preterm*. In [15] the focus was on identifying the most important risk factors for *preterm* birth. The global accuracy of these studies varied from between 73% and 97%.

A study by [47] used the TPEHG database [10] to evaluate classification accuracy. This occurred via *sample entropy*, against thirty cepstral coefficients and then against three. One of the feature sets used for classification consisted of calculating 30 cepstral coefficients, from each signal recording. The second feature set contained three cepstral coefficients. The selection of these values occurred by sequential forward selection and Fisher's discriminant. A multi-layer perceptron neural network classified the records, into *term* and *preterm* records. The results indicate that the reduced feature set, of three cepstral coefficients, gave the best classification accuracy of 72.73% ($\pm$ 13.5). This was in contrast to the entire thirty coefficients, whose accuracy was 53.11% ($\pm$ 10.5), and *sample entropy*, which was51.67% ($\pm$ 14.6). In addition, Support Vector Machine (*SVM*) classification has been used in [13–15]. They classify contractions into labour or non-labour using different locations on the abdomen. The classifications were also combined with decision fusion rules – majority voting, weighted majority voting (*WMV*) and kernel function was the Gaussian radial basis function (*RBF*). The features that were used include the *power* of the *EMG* signal, and the *median frequency*. The highest accuracy for a single *SVM* classifier, at one location, was 78.4% [13,14], whilst the overall classification accuracy, for the combined *SVM*, was 88.4% [15]. Finding the coefficients, for the decision boundary, occurs by solving a quadratic optimization problem.

In contrast, [48] have utilized *k-NN* for classification. However, the emphasis of this work was much more on the Autoregressive (*AR*) modelling and wavelet transform pre-processing techniques. The study aimed to classify contractions, from 16 women, into three types. G1, were women who had their contractions recorded at 29 weeks, and then delivered at 33 weeks; G2 were also recorded at 29 weeks, but delivered at 31 weeks, and G3 were recorded at 27 weeks and delivered at 31 weeks. Classification occurred against G1 and G2 and against G2 and G3. The comparison of two different classification techniques then occurred. (1) *k-NN* combined with the pre-processing method of *AR*, and (2) an Unsupervised Statistical Classification Method (*USCM*) combined with the pre-processing method of Wavelet Transform. The basis of *USCM* is the *Fisher Test and k-Means* methods and it would appear the authors have designed this method. The wavelet transform, combined with *USCM*, could provide a classification error of 9.5% when discerning G1 against G2, and 13.8% when classifying G2 against G3. On the other hand, using *AR*, the *k-NN* was able to provide a classification error of 2.4% for G1 against G2 and 8.3% for G2 against G3. In other words, in both classifications, the *AR* and *k-NN* methods performed better than the *USCM*. In addition, the classification accuracy of G1 against G2 was always lower than the equivalent G2 to G3 classifications. This suggests that it is easier to distinguish between pregnancies recorded at different stages of gestation than it is to distinguish between the time of delivery.

5. Prediction of preterm deliveries

Despite the advances, within the last twenty years in the *EHG* diagnosis and prediction field, knowledge of the uterus, and its mechanisms remains particularly poor. This is especially evident when compared to other organs such as the heart, and to a lesser extent, the gastro-intestinal system [21]. Given this inadequate knowledge, it may be easier to utilize an empirical backward looking, 'data mining' or 'brute force' approach. This is opposed to a forward-looking conceptual model approach, in order to find features that best describe pregnancy.

The aim of most studies, in *EHG* prediction or detection, has been to detect true labour, rather than predicting, in advance, whether delivery will be *preterm* or *term*. Furthermore, many of the studies concentrated on a late state in gestation. Even if earlier stages were incorporated, they always only included those with threatened *preterm* labour. However, the TPEHG dataset is different, as it involves the general population of pregnant women. Therefore, this collection includes fewer records for women who delivered *preterm* than *term*.

For *term* deliveries, true labour only starts within 24 h. For *preterm* deliveries, it may start within 7–10 days. The change in *EHG* activity, from non-labour to labour, is dramatic; throughout the rest of pregnancy, any change in *EHG* is more gradual. Therefore, it is expected that, classification of the records, into *preterm* and *term*, is particularly challenging. For this reason, and due to the configuration of the dataset, the study attempts to classify records from an earlier stage, according to whether they will eventually result in *term* or *preterm* deliveries.

5.1. Self-organized multilayer network inspired by immune algorithm

The self-organized network inspired by the immune algorithm is developed to improve recognition and generalization capability of the backpropagation neural networks.

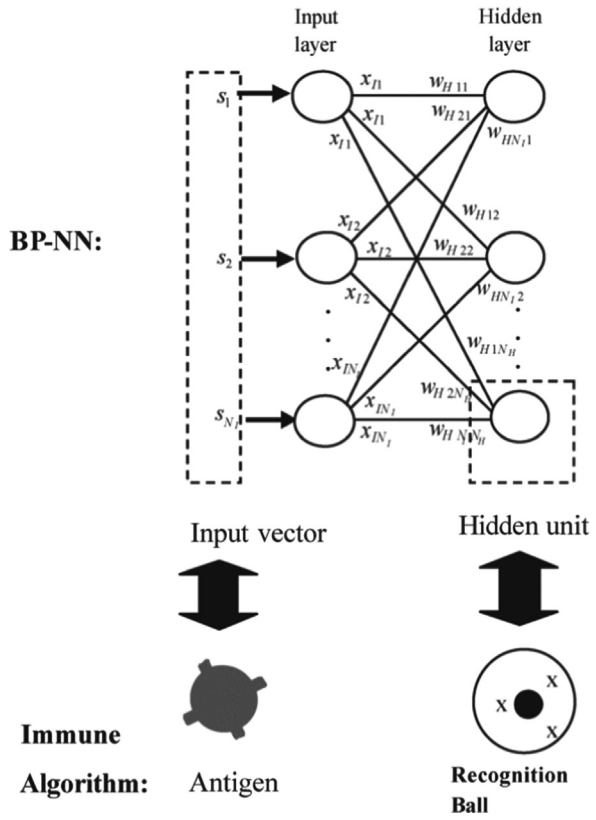

Fig. 1. Input vector and hidden units of Backpropogation-NN are considered as antigen and recognition ball of immune algorithm [58].

The concept of immune algorithm was initially discussed in [52]. Then in 2005, Widyanto et al. [58] suggested the self-organized inspired by immune algorithm as a method to improve recognition and generalization capability of the back-propagation neural networks which is called Self-organized neural network inspired by the immune algorithm (SONIA).

The idea is based on the relation between antigen and cell (Recognition Ball). Biologically, it is known that when the cell matched with antigen this antigen stimulates the cell to duplicate itself. Then a mutated cell is created to fight unknown viruses that are attacked via the antigen.

In neural network, the input vector represents an antigen while the hidden layer of the network is considered as a recognition ball as shown by Fig. 1.

In immune system, there is a recognition ball and an antigen. The recognition ball includes the B-cell, a single epitope and many paratopes, the epitope is attached to B-cell and paratopes are attached to antigen. Single B-cell can represent various antigens. It should be noted that the antigen which has the most characteristic with a particular type of recognition ball, simulates it to create mutated B cells allowing it to fight unknown viruses. For the SONIA network, the hidden unit is created to deal with the testing data that has similar characteristics to the trained data. The mutated hidden unit is generated in the region where no hidden unit currently exists. The idea of adding a hidden unit similar to adding artificial data to this region, this will allow the training data to be distributed in the input space better and hence improves the generalization capability.

In neural networks, to overcome the overfitting problem, a hidden unit represents the recognition ball. This hidden unit has a centroid, which can represent several input vector, therefore the centroids are chosen to determine the value of the connection weights between input nodes and the corresponding hidden unit. This will provide the neural network with the ability to prevent memorizing the training data.

The self-organized neural networks inspired by immune algorithm have a mutated hidden unit that is generated to deal with the testing data. The generated units of the hidden layer can be realized as adding artificial data to that region; therefore, the distribution of training data in the input space will become more acceptable and lead to improved generalization ability.

5.2. Dynamic Self-organised Network Inspired by the Immune Algorithm (DSIA)

In this paper, we build on previous research and propose a new dynamic neural network architecture that incorporates recurrent links within the structure to create a self-organising layer; inspired by Artificial Immune System theory [52]. As shown previously, Widyanto et al. [58] introduced a method to improve the recognition as well as the generalization capabilities of the backpropagation algorithm. Our proposed neural network architecture extends the SONIA network by introducing recurrent links.

The proposed network has three layers, the input layer the hidden layer and the output layer. These include the dynamic self-organization of hidden-layer units and hidden-based feedback to the input layer. This represents a major improvement compared to feed-forward networks, which can only implement a static mapping of the input vectors. In order to model dynamic functions, it is essential to exploit a system capable of storing internal states and implementing complex dynamics. Neural networks with recurrent connections are dynamic systems with temporal/state representations, which, because of their dynamic structure, have been successfully used for solving a variety of problems. This work is motivated by the potential of recurrent dynamic systems in solving complex real-world problems. Furthermore, in the last couple of years, various medical applications based on Recurrent Neural Networks (RNN) have been employed, which use the recurrent Elman neural network trained with the Levenberg–Marquardt algorithm to classify arterial disease. In this study, the trained Elman network obtained a high classification accuracy of 97% [56].

This section provides an overview of the Dynamic Self-organising Multilayer network inspired by the Immune Algorithm (DSIA) as shown in Fig. 2.

The input vector and the hidden layer of the proposed DSIA network represent the antigen and recognition ball, respectively. The recognition ball, which is the generation of the immune system, is used for hidden unit creation.

For the DSIA network, each hidden unit has a centre that represents the number of input vectors that are attached to it. To avoid the overfitting problem each centre has a value that represents the strength of the connections between input units and their corresponding hidden units. Furthermore, recurrent connections are established between the hidden layer and the input layer.

In Self-organised Kohonen networks (SOM), each unit j of a map ($1 \le j \le nh$), where nh is the number of hidden units, is compared with the weight vector w_j and an input $x(t)$ and $t=1,\ldots,ni$, and ni is the number of input units and the output. The Euclidean distance function is used for the comparison between w_j of the hidden map and the $x(t)$ input.

$$E_j = \sqrt{\sum_{i}^{ni}(x(t)_i - w_j)} \tag{1}$$

For an input vector, the best matching unit is the unit that minimizes the error function:

$$E = ||x(t) - w_j|| \tag{2}$$

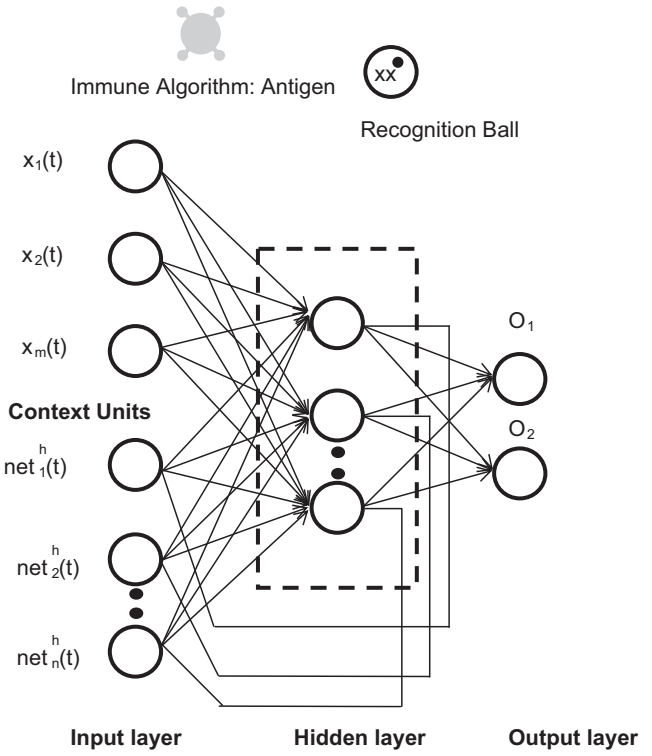

Fig. 2. Dynamic Self-organized Neural Network Inspired by Immune Algorithm (DSIA)

The learning rule is based on updating the weights of neurons that are related to a neighbourhood of the best matching unit:

$$\Delta w_j = \gamma h_k(x(t) - w_j) \tag{3}$$

where γ is the learning rate, k is the index of the best matching unit and h is the neighbourhood function, which decreases the distances between units j and k on the map.

In the case of the proposed DSIA, each unit j on the map has two weights, w_{ij}^x and w_{jj}^h, where wx is the weight matrix of the map with the input and wh is the weight matrix of the context unit and the output of the hidden layer at the previous time step $\text{net}_j^h(t-1)$ where

$$\text{net}_j^h(t) = \sigma(\alpha||x(t) - w_{ij}^x|| + \beta||\text{net}_j^h(t-1) - w_{jj}^h||) \tag{4}$$

with $\sigma > 0$ and $\beta > 0$, $\|.\|$ denotes the standard Euclidean distance of the vectors. The best matching unit is defined as the unit that minimized net where

$$D = \text{argmin}\{\text{net}_j^h(t)\} \tag{5}$$

Finally, the learning rule applied to update the weights of input units and context units are

$$\Delta w_{ij}^x = \gamma h(x(t) - w_{ij}^x) \tag{6}$$

$$\Delta w_j^h = \gamma h(\text{net}_j^h(t-1) - w_{jj}^h) \tag{7}$$

The purpose of hidden unit creation is to form clusters from input data and to determine the centroid of each cluster formed. These centroids are used to extract local characteristic of the training data and to enable the DSIA network to memorize the characteristics of training data. The use of Euclidean distance to measure the distance of input data and these centroids, enables the network to exploit local information of the input data, while the recurrent links enable the proposed network to remember past behaviours.

The weights in the hidden layers are updated using B cell creation, where the hidden unit is considered as recognition ball in the immune algorithm. In the initialization procedure, the first hidden unit (t_1, w_{h1}) is generated with $t_1 = 0$ and w_{h1} is taken arbitrarily from the input vector.

The procedure of the immune algorithm is used to create the hidden unit [52]. This procedure will be repeated until all inputs have found their corresponding hidden unit as follows:

For $m = 1$ to N_i which is the total number of input, repeat the following procedure:

1. Calculate the Euclidean distance between mth input and the centroid of the jth hidden unit $j = \{1,\ldots,N_h\}$ by

$$D_{hj}(t) = \sqrt{\sum_{i=1}^{N_i} (u_i(t) - W_{hji})^2} \tag{8}$$

where $u_i(t)$ is the ith input unit of the input vector and the weight W_{hji}.

2. Find the short distance D_c

$$D_c(t) = \arg\min D_{hj}(t) \tag{9}$$

3. Check the distance D_c if it is below stimulation level s_1, where $s_1 \in [0\ 1]$, then the input has found its corresponding hidden unit.
4. If the shortest is bigger than the stimulation level s_1, adjust the number of the hidden neurons N_h as follows:

$$N_h = N_h + 1 \tag{10}$$

i.e. generate a new hidden unit, and go to step 1.

5.3. Methodology

Fele-Zorz et al. conducted a comprehensive study that compared linear and non-linear signal processing techniques to separate uterine *EMG* records of *term* and *preterm* delivery groups [10]. The study was based on *EHG* records that were collected from a general population of pregnant patients at the Department of Obstetrics and Gynaecology Medical Centre in Ljubljana, between 1997 and 2006. These records are publically available, via the TPEHG dataset, in Physionet[3].

The aim of this paper was to evaluate the effectiveness, and efficiency, of using machine-learning algorithms, on the TPEHG dataset, to classify *preterm* and *term* delivery records. In the TPEHG dataset, there are 300 records/recordings (one record per pregnancy). Each recording was approximately 30 min long, had a sample frequency of 20 Hz, and had a 16-bit resolution, with an amplitude range ± 2.5 mV. Prior to sampling, the signals were passed through an analogue three-pole Butterworth filter, in the range of 1–5 Hz. Four electrodes were attached to the abdominal surface, with the navel at the symmetrical centre. Three signals were actually obtained simultaneously per 'record' by recording through three different channels – *Channel* 1, *Channel* 2, and *Channel* 3.

Although 1211 records were collected, only 300 of these were used. Rejection occurred on those records which had excessive noise or no discernible electrical activity, or that ended in Caesarean sections or induced delivery. Records were either recorded early, < 26 weeks (at around 23 weeks of gestation) or later, ≥ 26 weeks (at around 31 weeks).

The 0.3–3 Hz filtered signals on *Channel* 3 was chosen, since it is the best filter for discriminating between *preterm* and *term* delivery records, as reported in [49]. The dataset records were

[3] http://www.physionet.org

generated using four features – *root mean squares*, *peak frequency*, *median frequency*, and *sample entropy. Mean frequency* and *sample entropy* have the most potential to discriminate between *term* and *preterm* records, while *root mean squares* and *peak frequencies* have had conflicting results, in the classification of *term* versus *preterm* records, but, nonetheless, have shown potential. In this paper, all these features are considered.

The TPEHG data has pre-labelled classes. Therefore, supervised learning has been chosen as the learning technique. There is no such thing as one best classifier for all data domains; the choice of classifier depends on the dataset to some extent. The selection of an appropriate classifier, still generally involves a trial-and-error process, although statistical validation can be used to guide the process [50].

Artificial Neural Networks (ANN) appeared to be the classifier of choice, as they are featured in many studies [45,46]. In this paper, our new DSIA model will be used and compared against other common neural networking models (MLP, SONIA and Fuzzy SONIA).

Following an analysis of the literature simple yet powerful algorithms, which give good results, will also be considered in our experiments. These include the k-nearest neighbour, decision tree classifier and the support vector classifier. The support vector and k-nearest neighbour classifiers are nonlinear classifiers. Nonlinear classifiers compute the optimum smoothing parameter between classes in the datasets. Using smoothing parameters without any learning process produces discrimination. Smoothing parameters may be a scalar, a vector or a matrix with objects and their features. The decision tree classifier uses binary splitting and classes are decided upon the basis of a sequence of decision rules.

Sensitivity, *specificity*, *positive predicted values*, *negative predicted values*, overall accuracy of the classifier and ROC curves are used as the performance evaluation techniques. These have been chosen since they are suitable evaluation methods for classifiers, which produce binary output [51].

5.3.1. Data pre-processing

By initially reviewing the features, in the dataset, using quantile–quantile plots (Q–Q Plots), for each of the features, normal distributions were not evident, as illustrated in Fig. 3.

The plots show that there are likely outliers in the data. This is particularly the case with *root mean squares*, *median frequency* and *peak frequency*, as there are significant departures from the reference line for several observations. The likely cause for these outliers could either be from the sensor equipment itself; movement by the mother, or child, during data capture; or interference from other equipment, in the hospital room or ward. From the

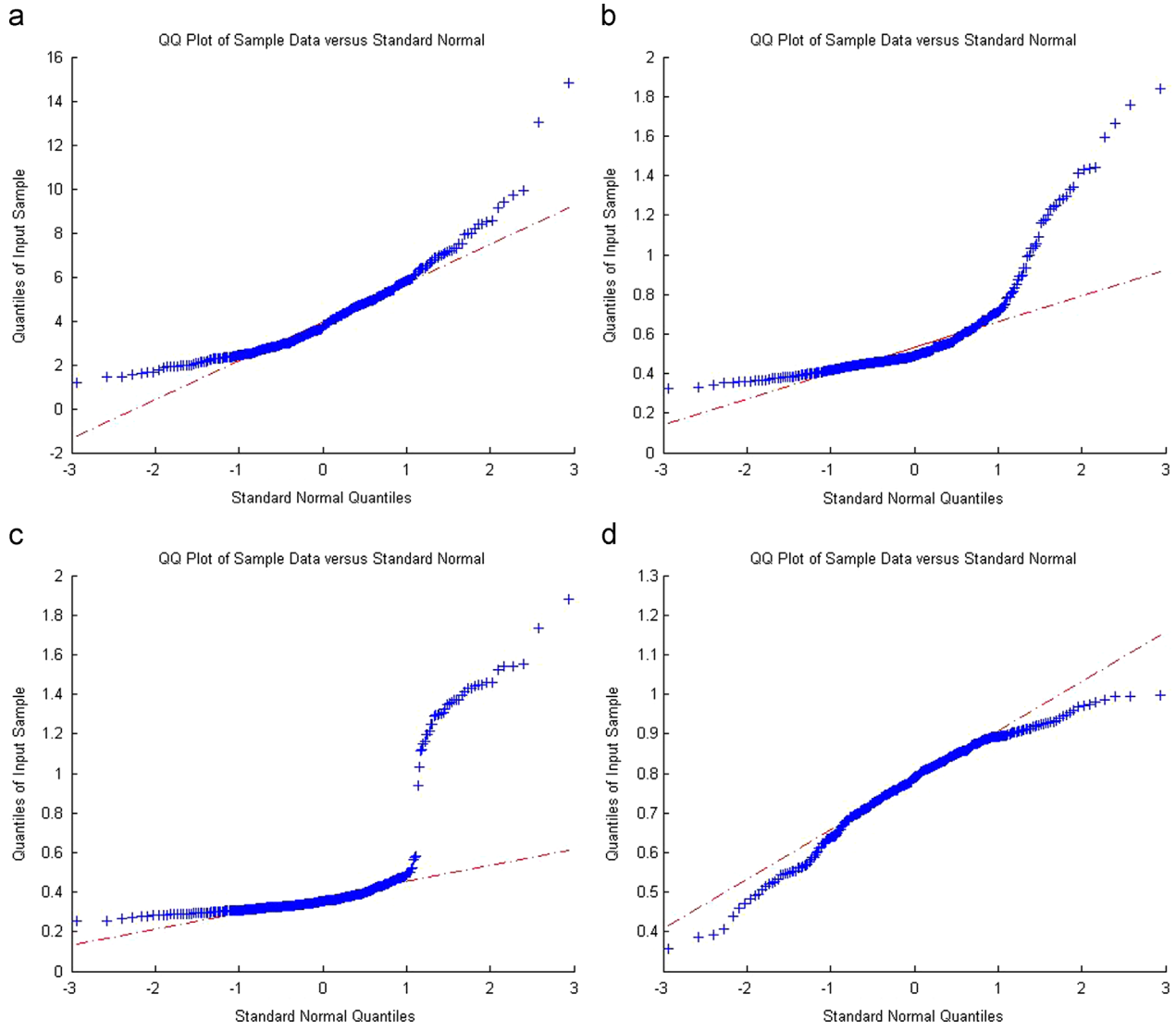

Fig. 3. Outliers in TPEHG data. (a) RMS QQPlot. (b) median frequency QQPlot, (c) peak frequency QQPlot and (d) sample entropy QQPlot.

dataset, and information provided by the original producers of TPEHG, this information was not available.

The removal of all outliers occurred by looking at the upper and lower limits, for each of the features. This transpired across all records that reside outside the body of records, in the dataset. For example, looking at the *root mean squares* feature, most of the records reside within 1.5 and 7. All records with *root mean square* values bigger than 7, and less than 1.5, have been removed. This process was repeated for *median frequency* (values bigger than 0.7 and less than 0.3 have been removed), *peak frequency* (values bigger than 0.5 and less than 0.25 have been removed) and *sample entropy* (values bigger than 1.0 and less than 0.5 have been removed). This ensures the removal of values, which are furthest from the sample *mean*. The results from a *lilliefors* test, on each of the features, still conclude that the data is not normally distributed. However, the Q–Q Plots illustrate that several of the features are close to being normally distributed, as shown in Figs. 4. Fig. 5 shows the proposed system for term/pre-term classifications.

6. Evaluation

In this paper, the proposed DSIA network was benchmarked against the SONIA network, the multilayer MLP neural network, Fuzzy-SONIA, K-Nearest Neighbour Classifier (KNN), Decision tree classifier (TREEC) and Support Vector Classifier (SVC). Their performances have been evaluated using the sensitivity, specificity, positive, and negative predicted values that each classification algorithm produced to separate term and preterm signals. The data has been split up as follows – 60% of the data is randomly selected for training data, 25% for validation and 15% for testing. The experiments have been run thirty times to generate an average of the results obtained. The formulas used to measure sensitivity; specificity and accuracy are defined as follows:

$$\text{Accuracy} = ((TP+TN)/(TP+TN+FP+FN))100 \tag{11}$$

$$\text{Sensitivity} = TP/(TP+FN) \tag{12}$$

$$\text{Specificity} = TN/(FP+TN) \tag{13}$$

This section presents the classification results for *term* and *preterm* delivery records using the TPEHG dataset. The 0.3–3 Hz filter on Channel 3 is used.

6.1. Results for 0.3–3 Hz TPEHG filter on Channel 3 with RMS, FMean, FPeak, and sample entropy with oversampling

This evaluation uses the 0.3–3 Hz filtered signals on Channel 3 with seven classifiers. The performance for each classifier is evaluated using Sensitivity, Specificity, Negative and Positive predicted values with 30 simulations and randomly selected training and testing sets for each simulation.

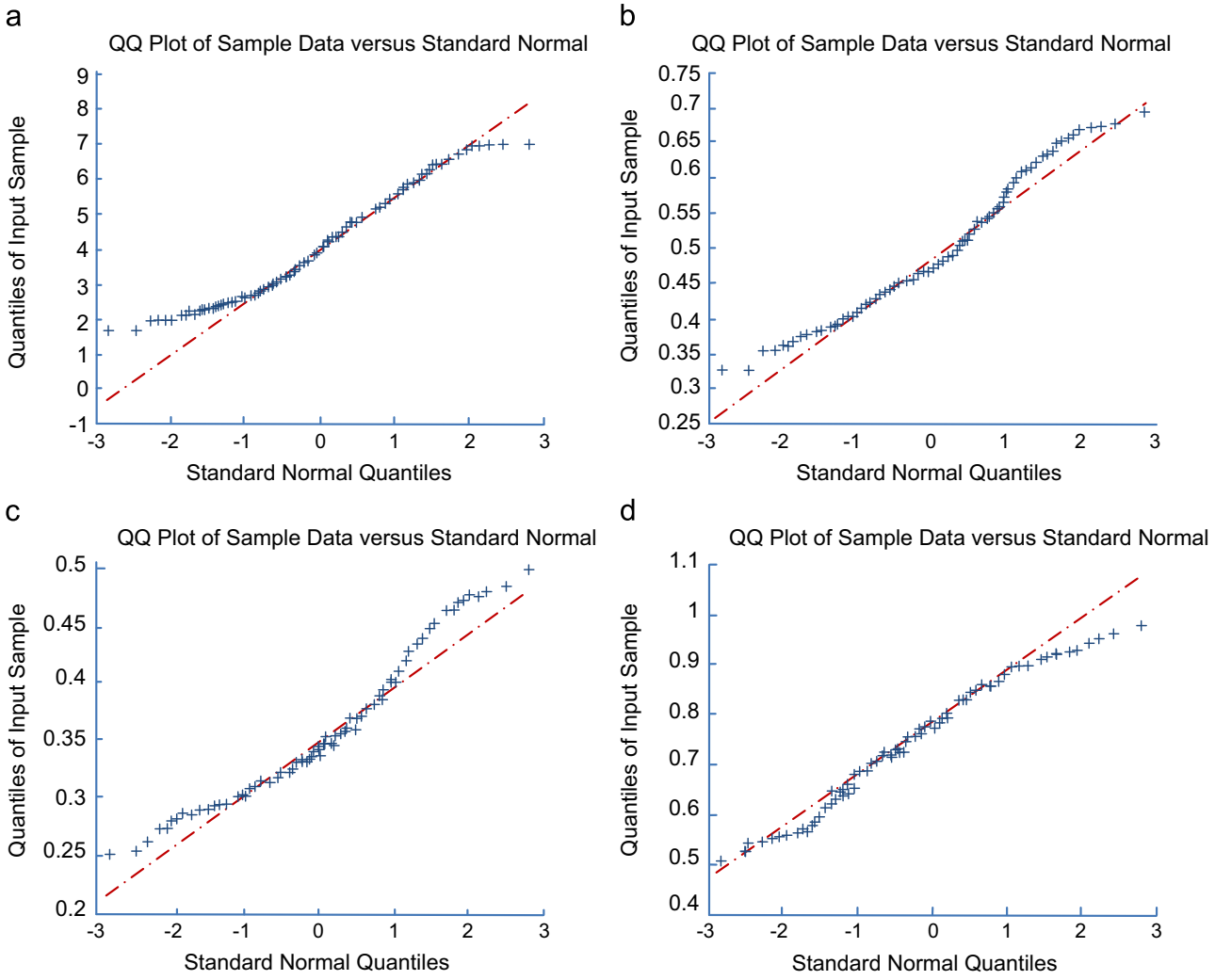

Fig. 4. TPEHG data with Outliers removed. (a) RMS QQPlot, (b) median frequency QQPlot, (c) peak frequency QQPlot and (d) sample entropy QQPlot.

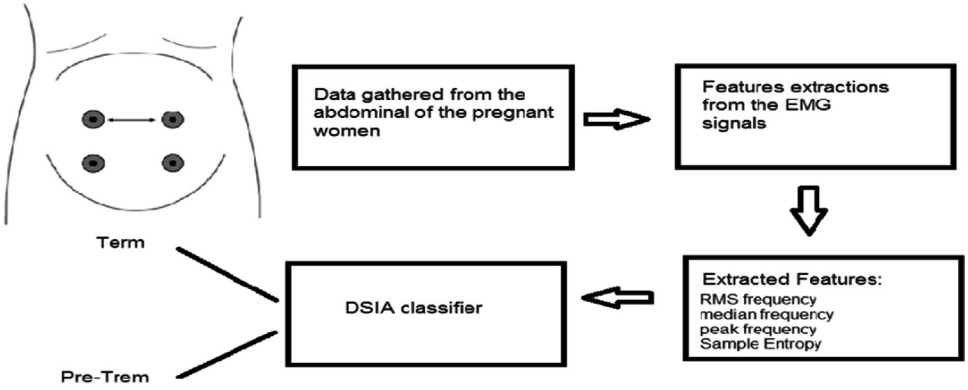

Fig. 5. The complete term/pre-term data classification system

Table 1
Classifier performance results for the 0.3–3 Hz filter.

	Sensitivity	Specificity	True negative	True positive
MLP	0.6481	0.5691	0.6261	0.6205
SONIA	0.6316	0.6920	0.7959	0.6073
KNNC	0.6944	0.6388	0.6764	0.6578
TREEC	0.5833	0.6111	0.5945	0.6000
SVC	0.6660	0.6388	0.6571	0.6486
DSIA	0.8269	0.3229	0.6676	0.5314
Fussy_SONIA	0.8642	0.4566	0.7647	0.6227

Table 2
Mean error, standard deviation and classifier accuracy.

	Mean error	Std error	Accuracy (%)
MLP	0.3073	0.0369	61.60
SONIA	0.2244	0.0031	70.34
KNNC	0.4000	0.0424	66
TREEC	0.3907	0.0565	59
SVC	0.4292	0.0476	65
DSIA	0.2410	7.8507e-05	56
Fuzzy_SONIA	0.2287	0.0015	66.41

6.1.1. Classifier performance

The first evaluation uses the original TPEHG dataset (38 *preterm* and 262 *term*) – the preterm are oversampled using min and max to produce 262 *preterm* records. Tables 1 and 2 show the average performance for all the classifiers used in this experiment. The simulation results as shown in Table 1 indicated that the sensitivity for preterm records obtained with the SONIA network is slightly better than the MLP; the specificity is also higher than the MLP. This means that the SONIA network has the ability to predict the true positive value of the preterm class; it can also predict the true negative value, which is the term class. The DSIA shows the highest values for *Sensitivity* and *True Positives* with slightly lower values for *Specificity* and *True Negatives* as can be seen in Table 2.

6.1.2. ROC analysis

The receiver operator characteristic (ROC) curve shows the cut-off values for the true negatives and false positives. The simulation results indicated that the proposed DSIA network showed an average sensitivity of 0.8269, while the SONIA network demonstrated an average sensitivity value of 0.6316. The ROC in Figs. 6 and 7 illustrate the trade-off between a classifier's true positives rate (*sensitivity*) versus its false positives rate (*specificity*).

Table 3
Classifier performance results for the 0.3–3 Hz filter.

	Sensitivity	Specificity	True negative	True positive
MLP	0.8070	0.8627	0.8000	0.8803
SONIA	0.9123	0.9451	0.9060	0.9490
KNNC	0.8076	0.9047	0.7916	0.9130
TREEC	0.8846	0.7619	0.8421	0.8214
SVC	0.8076	0.8571	0.7826	0.8750
DSIA	0.9123	0.8824	0.9000	0.8966
Fuzzy-SONIA	0.8401	0.7881	0.8561	0.7673

Fig. 6 shows that the SONIA performance is lower than the DSIA, which indicates that the DSIA curve is close to the upper left corner and its area is greater than the SONIA curve; this confirms that the DSIA has greater power for classification than SONIA does.

6.2. Results for 0.3–3 Hz TPEHG filter on Channel 3 with RMS, FMean, FPeak, and sample entropy with clinical data and over-sampling

Each TGEHG signal record contains clinical information relating to each of the patients that includes the pregnancy duration, gestation duration at the time of recording; maternal age, number of previous deliveries (parity); previous abortions, weight at the time of recording, hypertension, diabetes, placental position, bleeding first trimester, bleeding second trimester, funnelling, and whether they are a smoker. These 11 items of clinical information are added to the original TPEHG feature set (RMS, FMean, FPeak and SampEmp). Some information was missing for some of the patients, which led to unknown features on some recorders. Hence, the records with unknown information are removed for the next experiment, reducing the number of samples in the dataset to 19 *preterm* data samples and 108 *term* data samples. As before, the re-sampling method was applied to generate the 150 *preterm* data items.

6.2.1. Classifier performance

The 19 preterm records are oversampled using a min/max technique. This technique allows a new dataset to be constructed that provides an even balance between *term* and *preterm* records. The evaluated results for the proposed DSIA network are illustrated in Tables 3 and 4. The SONIA model scored the highest in all evaluation parameters, followed closely by the proposed DSIA model.

The results in Table 4 show that the highest values obtained are by the SONIA network in terms of mean error, standard deviation and classification accuracy followed by the DSIA network.

Table 4
Mean error, standard deviation and classifier accuracy.

	Mean error	Std error	Accuracy (%)
MLP	0.1681	0.0491	84.87
SONIA	0.0741	0.0011	92.77
KNNC	0.2260	0.0505	65.00
TREEC	0.2433	0.569	82.00
SVC	0.1761	0.0549	83.00
DSIA	0.0870	3.3486e-04	89.8148
Fussy-SONIA	0.1535	0.0025	81.17

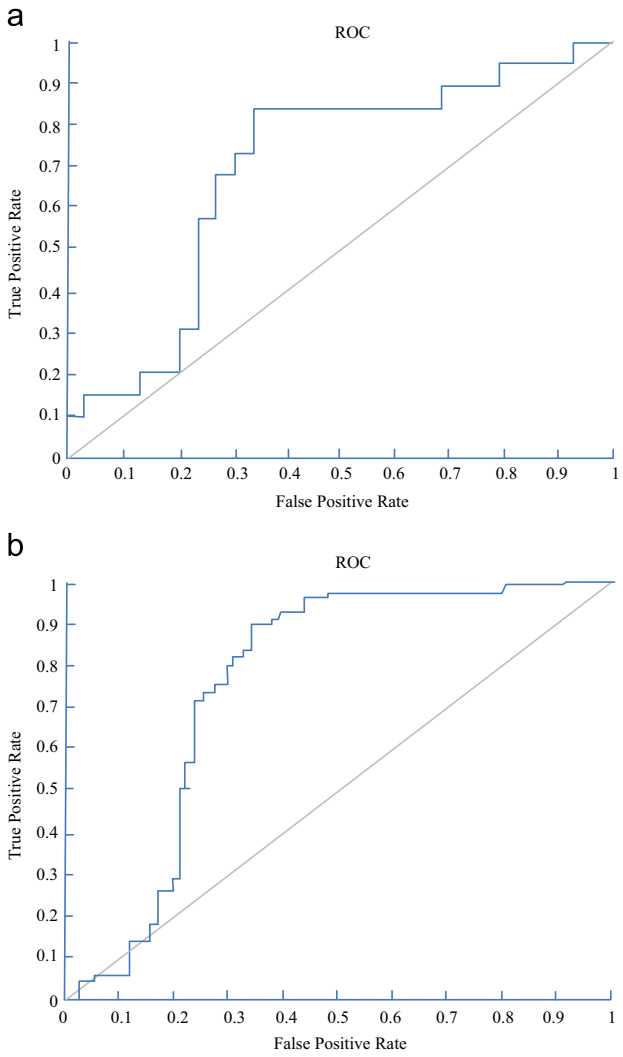

Fig. 6. ROC curve for two best performing classifiers. (a) ROC curve for SONIA and (b) ROC curve for DSIA.

6.2.2. *ROC analysis*

The ROC curves in Fig. 7 show an improvement compared to the ROC curve illustrated in Fig. 6. The area under the curve for DSIA and SONIA are 0.90 and 0.93, respectively. Extending the number of features to 15 has significantly improved the classifiers' performance. These features have provided classification methods with enough information from each record to allow them to obtain better values in all of the evaluation functions.

The results illustrate that using the proposed DSIA are encouraging. Within a wider context this approach might be able to utilize real-life data to predict, with high confidence, whether an expectant mother is likely to have a premature birth or proceed to full term.

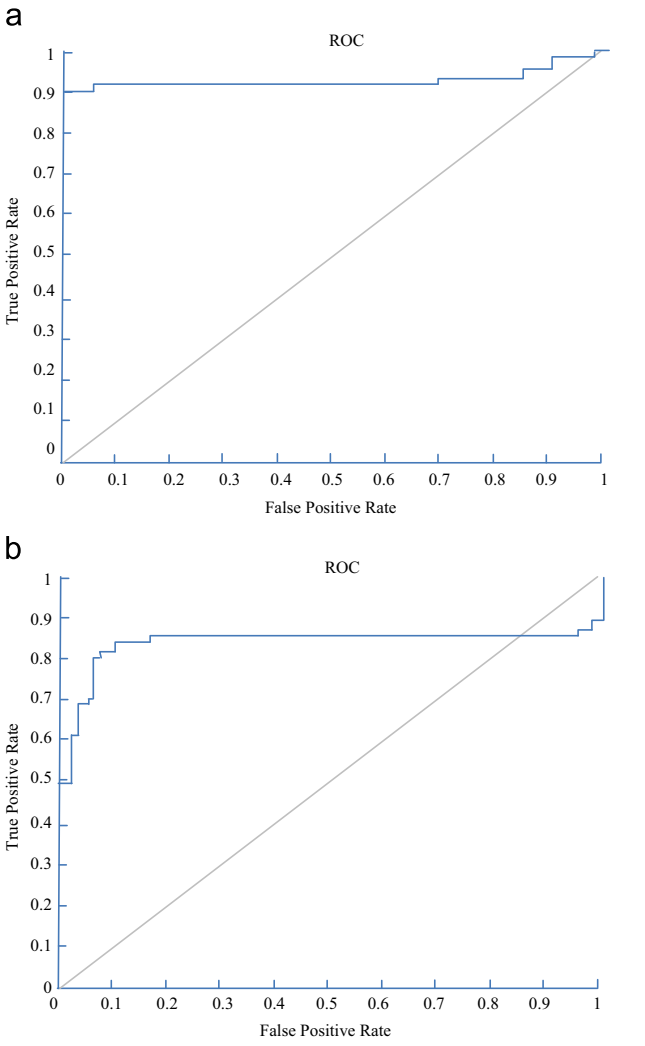

Fig. 7. ROC curve for two best performing classifiers. (a) ROC curve for SONIA and (b) ROC curve for DSIA

7. Discussion

Most of the uterine EHG signal studies have concentrated on predicting true labour, which is based on the last stage of the pregnancy duration. This paper has evaluated the use of a machine learning approach, using records from earlier stages of gestation, to predict term and preterm deliveries.

The method of classification used in this paper compares several existing classification algorithms and our newly proposed DSIA neural network. The evaluation of classifier performance has been measured using sensitivity and specificity, which are suitable evaluation measures for binary outputs (term/preterm). In addition, the capabilities of classifiers have been visually compared using the ROC curve, which is a technique commonly used in decision-making. It is a useful method for visualizing classifier performance.

From the results obtained from the oversampling data, the results show that the self-organized hidden layer immune systems and dynamic links improve the predictive capabilities of classifiers. More importantly, the proposed DSIA model shows promise where the results indicate that it outperformed several classification algorithms. This improvement can be associated with the novel combination of supervised and unsupervised learning techniques used in the DSIA model and neural networks in general [50]. This has helped to overcome the limitations often found in back propagation learning.

We conclude that the DSIA network has performed well in the classification of uterine signal because it has used SOM unsupervised methods in the hidden layer and recurrent links. The hidden layer can cluster the input nodes to the centroids of hidden units, which gives the local network pattern of the input data. The Euclidean distance was utilized to compute the distance between the input units and the centroids of hidden units. Thus, DSIA is able to exploit locality characteristics of the data [19].

8. Conclusions and future work

This research work underlines an important contribution of a new recurrent self-organising multilayer neural network inspired by artificial immune systems. In this case, recurrent links from the hidden layer to the input layer was proposed, which allows the network to have memory. It is applied to and evaluated against the classification of term and preterm records for pregnant women. The simulation results indicated that the proposed network achieved improved results when compared to a number of machine learning algorithms.

The focus of the paper has been to improve sensitivity rates, as it is more important to predict preterm delivery, as opposed to misclassifying a term pregnancy. The results indicated that using the original TPEHG dataset, the number of preterm records (minority class) was considerably lower than the number of term records (majority class). Since the classifiers do not have enough preterm records to learn from, the results were not significant; however when using the oversampling technique for the minority class, this enables the distribution between the two classes (term and preterm) to be more balanced.

The simulation results indicated that the DSIA network has performed well in the classification of uterine signal due to the SOM unsupervised methods in the hidden layer and recurrent links. The simulation results indicated that the proposed technique achieved 56% accuracy and 0.8269 sensitivity when using the oversampled TPEHG dataset. Furthermore, when using additional clinical data the accuracy had improved to 89.8148% while the sensitivity was 0.9123.

Future work will investigate and assess an improved, regularised-DSIA scheme for the proposed DSIA network. While weight-decay is not without its performance-related problems, work by others has shown that it can help to avoid over-fitting the network to training data, as such, improving the network. Another direction of research is to investigate the best choice of network architecture [53–55], [57], which includes the number of inputs and the use of higher order terms in the input units similar to the functional link neural network. This may improve the performance of the proposed network since this can extend the input space into higher dimensional spaces where linear separability is possible.

Acknolwedgements

The authors would like to thanks Prof. Paulo Lisboa from the School of Computing and Mathematical Sciences from Liverpool John Moores University for his valuable comments and supports for this research.

References

[1] WHO, Born Too Soon: The Global Action Report on Preterm Birth, 2012.
[2] P.N. Baker, L. Kenny, Obstetrics by Ten Teachers, 19th ed., Hodder Arnold (2011) 436.
[3] A. Greenough, Long term respiratory outcomes of very premature birth (< 32 weeks), Semin. Fetal Neonatal Med. 17 (2) (2012) 73–76.
[4] L.J. Mangham, S. Petrou, L.W. Doyle, E.S. Draper, N. Marlow, The cost of preterm birth throughout childhood in England and Wales, Pediatrics 123 (2) (2009) 312–327.
[5] M. McPheeters, W.C. Miller, K.E. Hartmann, D.A. Savitz, J.S. Kaufman, J.M. Garrett, M.J. Thorp, The epidemiology of threatened premature labor: a prospective cohort study, Am. J. Obstetr. Gynecol. 192 (4) (2005) 1325–1329.
[6] M. Lucovnik, R.J. Kuon, L.R. Chambliss, W.L. Maner, S.-Q. Shi, L. Shi, J. Balducci, R.E. Garfield, Use of uterine electromyography to diagnose term and preterm labor, Acta Obstetr. Gynecol. Scand. 90 (2) (2011) 150–157.
[7] R. Rattihalli, L. Smith, D. Field, Prevention of preterm births: are we looking in the wrong place?,, Arch. Dis. Child. Fetal Neonatal Ed. 97 (3) (2012) 160–161.
[8] R.L. Goldenberg, J.F. Culhane, J.D. Iams, R. Romero, Epidemiology and causes of preterm birth, Lancet v371 (9606) (2008) 75–84.
[9] L.J. Muglia, M. Katz, The enigma of spontaneous preterm birth, N. Engl. J. Med. 362 (6) (2010) 529–535.
[10] G. Fele-Žorž, G. Kavšek, Z. Novak-Antolič, F. Jager, 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. 46 (9) (2008) 911–922.
[11] E. Bozler, Electrical stimulation and conduction in smooth muscle, Am. J. Physiol. 122 (3) (1938) 614–623.
[12] M. Doret, Uterine electromyograpy characteristics for early diagnosis of mifepristone-induced preterm labour, Obstetr. Gynecol. 105 (4) (2005) 822–830.
[13] B. Moslem, M. Khalil, M.O. Diab, A. Chkeir, C. Marque, A Multisensor Data Fusion Approach for Improving the Classification Accuracy of Uterine EMG Signals, in: Proceedings of the 18th IEEE International Conference on Electronics, Circuits and Systems (ICECS), 2011, no. Mv, 11–14th December 2011, pp. 93–96.
[14] B. Moslem, M. Khalil, M.O. Diab, C. Marque, Classification of multichannel uterine EMG signals by using a weighted majority voting decision fusion rule, in: Proceedings of the 2012 16th IEEE Mediterranean Electrotechnical Conference, 25–28th March 2012, pp. 331–334.
[15] B. Moslem, M. Khalil, M. Diab, Combining multiple support vector machines for boosting the classification accuracy of uterine EMG signals, in: Proceedings of the 2011 18th IEEE International Conference on Electronics, Circuits and Systems (ICECS), no. Mv, 11–14th December 2011, pp. 631–634.
[16] B. Moslem, B. Karlsson, M.O. Diab, M. Khalil, C. Marque, Classification performance of the frequency-related parameters derived from uterine EMG signals, in: Proceedings of the International Conference of the IEEE Engineering in Medicine and Biology Society, 2011, pp. 3371–4.
[17] B. Moslem, M.O. Diab, M. Khalil, C. Marque, Classification of multichannel uterine EMG signals by using unsupervised competitive learning, in: IEEE Workshop on Signal Processing Systems, 2011, pp. 267–272.
[18] B. Moslem, M.O. Diab, C. Marque, M. Khalil, Classification of multichannel Uterine EMG Signals, in: Proceedings of the IEEE International Conference on Engineering in Medicine and Biology Society, 2011, pp. 2602–2605.
[19] C. Rabotti, M. Mischi, S.G. Oei, J.W.M. Bergmans, Noninvasive estimation of the electrohysterographic action-potential conduction velocity, IEEE Trans. Biomed. Eng. 57 (9) (2010) 2178–2187.
[20] C. Buhimschi, M.B. Boyle, R.E. Garfield, Electrical activity of the human uterus during pregnancy as recorded from the abdominal surface, Obstetr. Gynecol. 90 (1) (1997) 102–111.
[21] W.J. Lammers, The electrical activities of the uterus during pregnancy, Reprod. Sci. 20 (2) (2013) 182–189.
[22] R.E. Garfield, W.L. Maner, Physiology and electrical activity of uterine contractions, Semin. Cell Dev. Biol. 18 (3) (2007) 289–295.
[23] J. Gondry, C. Marque, J. Duchene, D. Cabrol, Electrohysterography during pregnancy: preliminary report, Biomed. Instrum. Technol./Assoc. Adv. Med. Instrum. 27 (4) (1993) 318–324.
[24] M. Lucovnik, W.L. Maner, L.R. Chambliss, R. Blumrick, J. Balducci, Z. Novak-Antolic, R.E. Garfield, Noninvasive uterine electromyography for prediction of preterm delivery, American journal of obstetrics and gynecology 204 (3) (2011) 228.e1–228.e10.
[25] H. Leman, C. Marque, J. Gondry, Use of the electrohysterogram signal for characterization of contractions during pregnancy, IEEE Trans. Bio-med. Eng. 46 (10) (1999) 1222–1229.
[26] I. Verdenik, M. Pajntar, B. Leskosek, Uterine electrical activity as predictor of preterm birth in women with preterm contractions, Eur. J. Obstetr., Gynecol., Reprod. Biol. 95 (2) (2001) 149–153.
[27] W.L. Maner, R.E. Garfield, H. Maul, G. Olson, G. Saade, Predicting term and preterm delivery with transabdominal uterine electromyography, Obstetr. Gynecol. 101 (6) (2003) 1254–1260.
[28] C.K. Marque, J. Terrien, S. Rihana, G. Germain, Preterm labour detection by use of a biophysical marker: the uterine electrical activity, BMC Pregnancy Childbirth 7 (Suppl. 1) (2007) S5.
[29] W.L. Maner, R.E. Garfield, Identification of human term and preterm labor using artificial neural networks on uterine electromyography data, Ann. Biomed. Eng. 35 (3) (2007) 465–473.
[30] M. Hassan, J. Terrien, C. Marque, B. Karlsson, Comparison between approximate entropy, correntropy and time reversibility: application to uterine electromyogram signals, Med. Eng. Phys. 33 (8) (2011) 980–986.
[31] C. Buhimschi, M.B. Boyle, G.R. Saade, R.E. Garfield, Uterine activity during pregnancy and labor assessed by simultaneous recordings from the myometrium and abdominal surface in the rat, Am. J. Obstetr. Gynecol. 178 (4) (1998) 811–822.

[32] M.O. Diab, A. El-Merhie, N. El-Halabi, L. Khoder, Classification of Uterine EMG signals using supervised classification method, Biomed. Sci. Eng. 3 (9) (2010) 837–842.
[33] P. Carre, H. Leman, C. Fernandez, C. Marque, Denoising of the uterine EHG by an undecimated wavelet transform, IEEE transactions on Bio-med. Eng. 45 (9) (1998) 1104–1113.
[34] W.L. Maner, L.B. MacKay, G.R. Saade, R.E. Garfield, Characterization of abdominally acquired uterine electrical signals in humans, using a non-linear analytic method, Med. Biol. Eng. Comput. 44 (1–2) (2006) 117–123.
[35] M.P. Vinken, C. Rabotti, M. Mischi, S.G. Oei, Accuracy of frequency-related parameters of the electrohysterogram for predicting preterm delivery, Obstetr. Gynecol. Surv. 64 (8) (2009) 529.
[36] R.E. Garfiled, W.L. Maner, H. Maul, G.R. Saade, Use of uterine EMG and cerical LIF in monitoring pregnant patients, BJOG: Int. J. Obstetr. Gynaecol. 112 (2005) 103–108.
[37] C. Buhimschi, R.E. Garfield, Uterine contractility as assessed by abdominal surface recording of electromyographic activity in rats during pregnancy, Am. J. Obstetr. Gynecol. 174 (2) (1996) 744–753.
[38] J.S. Richman, J.R. Moorman, Physiological time-series analysis using approximate entropy and sample entropy, Am. J. Physiol. – Heart Circ. Physiol. 278 (6) (2000) H2039.
[39] W.L. Woon, Techniques for Early Detection of Alzheimer's Disease using Spontaneous EEG Records, vol. 28, IOP Publishing (2007) 335–347.
[40] C.F.V. Latchoumane, Multiway array decomposition analysis of EEGs in Alzheimer's disease, J. Neurosci. Methods 207 (1) (2012) 41–50.
[41] D.I. Ispawi, N.F. Ibrahim, N.M. Tahir, Classification of Parkinson's Disease based on Multilayer Perceptrons (MLP) Neural Networks and ANOVA as a Feature Extraction, in: IEEE Conference on Signal Processing and its Applications, 2012, pp. 63–67.
[42] J.L. Whitwell, Neroimaging correlates of pathologically defined subtypes of Alzheimer's disease: a case-control study, Lancet Neurol. 11 (10) (2012) 868–877.
[43] S. Palaniappan, R. Awang, Intelligent heart disease prediction system using data mining techaniques, in: IEEE Computer Systems and Applications, 2008, pp. 108–115.
[44] S. Joshi, V. Simha, D. Shenoy, Classification and treatment of different stages of Alzheimer's disease using various machine learning methods, Int. J. Bioinf. Res. 2 (1) (2010) 44–52.
[45] A. Cawsey, The Essence of Artificial Intelligence, Prentice Hall (1998) 200.
[46] E. Charniak, Bayesian networks without tears, AI Mag. 12 (4) (1991) 50–63.
[47] S. Baghamoradi, M. Naji, H. Aryadoost, Evaluation of cepstral analysis of EHG signals to prediction of preterm labor, in: Proceedings of the 18th Iranian Conference on Biomedical Engineering, no. December 2011, pp. 1–3.
[48] M.O. Diab, C. Marque, M.A. Khalil, Classification for uterine EMG Signals: comparison between AR model and statistical classification method, Int. J. Comput. Cogn. 5 (1) (2007) 8–14.
[49] G. Fele-Zorz, G. Kavsek, Z. Novak-Antolic, F. Jager, 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. 46 (9) (2008) 911–922.
[50] C.V.D. Walt, E. Barnard, Data characteristics that determine classifier performance, in: Proceedings of the 17th annual symposium of the pattern recognition association of South Aftrica, 2006, pp. 1–6.
[51] T.A. Lasko, J.G. Bhagwat, K.H. Zou, L. Ohno-Machada, The use of receiver operating characteristic curves in biomedical informatics, J. Biomed. Inform. 38 (5) (2005) 404–415.
[52] J.I. Timmis, Artificial immune systems: a novel data analysis technique inspired by the immune network theory, University of Wales, Aberystwyth, 2001 (Ph.D. dissertation).
[53] D.S. Huang, Ji-Xiang Du, A constructive hybrid structure optimization methodology for radial basis probabilistic neural networks, IEEE Trans. Neural Netw. 19 (12) (2008) 2099–2115.
[54] D.S. Huang, The united adaptive learning algorithm for the link weights and the shape parameters in RBFN for pattern recognition, Int. J. Pattern Recognit. Artif. Intell. 11 (6) (1997) 873–888.
[55] D.S. Huang, The local minima free condition of feedforward neural networks for outer-supervised learning, IEEE Trans. Syst., Man Cybern. 28B (3) (1998) 477–480.
[56] N. Guler, E. Ubeyli, I. Guler, Recurrent neural networks employing Lyapunov exponents for EEG signals classification, Expert Syst. Appl. 29 (3) (2005) 506–514.
[57] D.S. Huang, Radial basis probabilistic neural networks: model and application, Int. J. Pattern Recognit. Artif. Intell. 13 (7) (1999) 1083–1101.
[58] M.R. Widyanto, H. Nobuhara, K. Kawamoto, K. Hirota, B. Kusumoputro, Improving recognition and generalization capability of back-propagation NN using Self-Organized Network Inspired by Immune Algorithm (SONIA), Appl. Soft Comput. 6 (2005) 72–84.

Dr. Abir Jaafar Hussain is a Reader at the School of Computing and Mathematical Sciences at Liverpool John Moores University, UK. She completed her PhD study at The University of Manchester, UK in 2000 with a thesis title Polynomial Neural Networks for Image and Signal Processing. She has published numerous referred research papers in conferences and Journal in the research areas of Neural Networks, Signal Prediction, Telecommunication Fraud Detection and Image Compression. She is a PhD supervisor and an external examiner for research degrees including PhD and MPhil.

Dr. Paul Fergus received a BSc (Hons) in Artificial Intelligence from Middlesex University in 1997, a Postgraduate Diploma and an MSc in Computing for Commerce and Industry from the Open University in 1999 and 2001, respectively and a PhD from Liverpool John Moores University in 2006.

Paul also became a member of the British Computer Society in 2004. Prior to his PhD Paul worked in industry for several years as a Software Engineer and has been directly involved in the development of several national and international projects for Sharp, Bardon Aggrigates, HM Prison Service, NMC, Liverpool Underground, and Ericsson.

Mrs. Haya Al-Askar is a full time research student working at the School of Computing and Mathematical Sciences at Liverpool John Moores University.

During her PhD study, she has published in a number of refereed journal and conference papers for the prediction of financial time series and the classification of medical data.

Dr. Dhiya Al-Jumeily is a principle lecturer in eSystems Engineering and leads the Applied Computing Research Group at the faculty of Technology and Environment. He has already developed fully online MSC and BSc programmes and currently heads the enterprise activities at the School of Computing and Mathematical Sciences. Dr. Al-Jumeily has published numerous referred research papers in multidisciplinary research areas including Technology Enhanced Learning, Applied Artificial Intelligence, Neural Networks, Signal Prediction, Telecommunication Fraud Detection and Image Compression. He is a PhD supervisor and an external examiner for the degree of PhD. He has been actively involved as a member of editorial board and review committee for a number peer reviewed international journals, and is on program committee or as a general chair for a number of international conferences.

Prof. Franc Jage is the head of the Laboratory for Computer Graphics and Multimedia at the University of Ljubljana. He has a large number of publications and research related to cardiovascular reflex tests and spectral analysis of heart rate variabili and the development of predictive models to assess the risk of preterm delivery.