International Journal of Medical Informatics 129 (2019) 198–204

Contents lists available at ScienceDirect

International Journal of Medical Informatics

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

A decision support system for predicting the treatment of ectopic pregnancies

Alberto De Ramón Fernández[a], Daniel Ruiz Fernández[b,*], María Teresa Prieto Sánchez[c]

[a] Department of Computer Technology (DTIC), University of Alicante, Carretera San Vicente s/n, 03690 Alicante, Spain
[b] Department of Computer Technology (DTIC), University of Alicante, Carretera San Vicente s/n, 03690 Alicante, Spain
[c] Service of Gynecology and Obstetrics, "Virgen de la Arrixaca" University Clinical Hospital, Institute for Biomedical Research of Murcia (IMIB-Arrixaca), Ctra. Madrid-Cartagena, s/n, 30120 El Palmar, Murcia, Spain

ARTICLE INFO

Keywords:
Aid decision algorithms
Classifier
Ectopics pregnancies
Clinical treatment

ABSTRACT

Background and objective: Ectopic pregnancy is an important cause of morbidity and mortality worldwide. An early diagnosis, as well as the choice of the most suitable treatment for the patient is crucial to avoid possible complications. According to different factors an ectopic pregnancy must be treated from surgery, using a pharmacological treatment or following a conservative treatment. In this paper, a clinical decision support systems based on artificial intelligence algorithms has been developed to help clinicians to choose the initial treatment to be followed by the patient.
Methods: A decision support system based on a three stages classifier has been developed. Each stage acts as a filter and allows re-evaluation of the classification made in the previous stage in order to find diagnostic errors. This classifier has been implemented and tested for four different aid algorithms: Multilayer Perceptron, Deep Learning, Support Vector Machine and Naives Bayes.
Results: The results prove that the evaluated algorithms Support Vector Machine and Multilayer Perceptron can be useful to help gynecologists in their decisions about initial treatment, especially with Support Vector Machine that presents accuracy, sensitivity and specificity outcomes about 96.1%, 96% and 98%, respectively.
Conclusions: According to the results, it is feasible to develop a clinical decision support system using the algorithms that present a higher precision. This system would help gynecologists to take the most accurate decision about the initial treatment, thus avoiding future complications.

1. Introduction

An ectopic pregnancy is a pregnancy outside of the uterine cavity. High rate of ectopic pregnancies is located on the fallopian tube (84%) [1], although other sites are also possible (cervical, cornual, hysterotomy scar, intramural, ovarian or abdominal). Ectopic pregnancy is a potentially life-threatening condition, so early diagnosis and treatment is desirable. With the routine use of transvaginal ultrasonography, location of a pregnancy and diagnosis of ectopic pregnancy can be stablished earlier [2]. Quantitative measurement of the beta subunit of human chorionic gonadotropin (beta-hCG) is also helpful in both diagnosis and follow up of these patients [3].

Treatment need to be started as soon as the diagnosis is confirmed to reduce the risk of rupture of the fallopian tube or another structure and subsequent haemorrhage. The three approaches to the management of ectopic pregnancy are surgery (salpingostomy or salpingectomy), medical treatment or expectant management. While surgical approaches are the gold-standard treatment, the use of medical management of these patients has become an accepted and cost-effectiveness alternative to surgical options [4]. Methotrexate is a folic acid antagonist that interferes with DNA synthesis and is the most widely used agent in the medical treatment of ectopic pregnancies. For appropriately selected patients, methotrexate is a non-invasive option that has comparable efficacy, safety, and fertility outcomes with surgery. The most commonly used protocol consists of a single planned dose of intramuscular methotrexate (50 mg/m^2 per body surface), followed by assessment of beta-hCG levels on days 4, 7 and weekly. Repeat methotrexate dosing is performed at day 7 for beta-hCG drops $< 15\%$. Contraindications for methotrexate therapy include hemodynamically unstable patients, those with abnormal renal or hepatic function, or low white blood count. The overall success rate of medical treatment in properly selected women is nearly 90% [4][5]. In select cases of early

* Corresponding author.
E-mail addresses: aderamon@dtic.ua.es (A. De Ramón Fernández), druiz@dtic.ua.es (D. Ruiz Fernández), mt.prieto@um.es (M.T. Prieto Sánchez).

https://doi.org/10.1016/j.ijmedinf.2019.06.002
Received 26 November 2018; Received in revised form 22 April 2019; Accepted 3 June 2019

ectopic pregnancy in which the risk of tubal rupture is minimal, expectant management might be an option [6]. Surgical indications have to be considered in case of hemodynamically unstable patient, signs or symptoms of current or impending tubal rupture, contraindication of methotrexate therapy or failed medical treatment. The advantages of surgical treatment are less time for resolution of the ectopic pregnancy and avoidance of the need for prolonged monitoring.

Some authors have reported that approximately one-third of women with ectopic pregnancies are candidates for methotrexate treatment [7], while the remaining two-thirds will require surgery. However, if early diagnosis is made, more patients with ectopic pregnancy could be successfully treated with methotrexate.

How to choose the best treatment for each patient in case of ectopic pregnancy continue to be a challenge. Although many scores to predict successful medical treatment has been design, some patients undergone surgical treatment after having been treated with methotrexate. This could be due to treatment failure or to a misclassification of the patients.

The objective of this work is to study the feasibility of developing a clinical decision support system that allows gynecologists to decide with greater reliability the best treatment for a patient with an ectopic pregnancy. To do this, several decision-aid algorithms will be implemented, tested with a database of real cases and the results will be analysed in terms of accuracy, sensitivity and specificity.

This paper has been organised as follows: After presenting related literature in Section 2, the methods used to build up our classifier are described in Section 3. Section 4 presents the results and their discussion. Finally, the conclusions are shown in Section 5.

2. Related literature

In the field of obstetrics, and more specifically in relation to ectopic pregnancies, sometimes it is difficult to obtain an early diagnosis. The high number of conditions that share symptoms with ectopic pregnancy makes this task difficult. Currently, there are different techniques that can help detect an ectopic pregnancy, such as a pelvic exam [8], a blood test to measure different markers (e.g. beta-HCG) [9] or the analysis of the inner tissue of the uterus [10]. Testing with medical diagnostic technology, such as ultrasound and laparoscopy, also allows and simplifies the early detection of this type of pregnancy. However, at present, to the traditional diagnostic techniques, the one that offers the information and communication technologies (ICTs) is being added. Artificial Intelligent (AI) is having a great impact today in any research area, and its use in the clinical field is having a great growth. AI algorithms are computational models that try to solve problems that cannot be solved with statistical methods. Thus, the development of AI algorithms that help predict or classify diseases from a knowledge base has meant a great advance in different areas of application [11], behaving as clinical decision support systems (CDSS).

In recent years, CDSS based on computational techniques for the prediction of pregnancies, risk factors and other information of interest have been developed. Most of these CDSS were implemented using artificial neural networks (ANNs). ANNs are computational algorithms whose structure is organised into layers of interconnected neurons, which are responsible for adjusting the algorithm. In [12], a CSSD developed with a neural network, the multilayer perceptron (MLP), was able to predict whether pregnancy had been desired or not based on five predictors (age of woman, woman's education, husband's education, number of children and use of contraceptives) with an average of sensitivity and specificity of 85.8% and 95.2% respectively.

In the case of [13], two decision support systems based on ANNs and multivariate logistic regression (MLR) were developed and compared to predict possible complications during pregnancy due to hypertension. To do this, it was experimented with women with high-risk of preeclampsia and intrauterine fetal growth retardation as a risk group. The accuracy, sensitivity, and specificity outcomes were 95.2%, 86.2% and 95.4% respectively for the ANN, and 96.2% 79.3% and 97% for the MLR, so they were an adequate and useful tool for the early detection of complications. Related to this pathology, in 2011 [14], a CDSS based on ANN was developed to classify and predict if the woman will suffer hypertension or preeclampsia or on the contrary will maintain a normal tension during pregnancy, from values of heart rate variability, maternal history and blood pressure. The metrics of sensitivity and specificity were around 80% and 90% respectively.

Other example of the use of ANN in obstetrics is the work of Paydar et al. [15], who experimented with two types of neural networks, MLP and radial basis function (RBF) for predicting pregnancy success in women with systemic lupus erythematous. From a sample of 149 women and the evaluation of 16 predictors, they were able to predict the success of pregnancy with an accuracy of 90.9%. Finally, the study of [16] evaluated 16 risk factors (age, bad habits, pregnancy complications, early antibiotic therapy, …) in pregnant women that can influence that children can develop autism. Considering these risk factors before and during pregnancy, the CSSD was able to classify children with or without autism with an accuracy higher than 80%.

However, nowadays, we have not found research works related with decision support systems focused on ectopic pregnancies. In this study, different types of artificial intelligent algorithms have been evaluated in order to develop a model to predict and classify the most suitable treatment to be followed during ectopic pregnancy.

3. Methods

To develop our proposal, we have analysed four different methods based on artificial intelligence and we have tested them with a real database of patients with ectopic pregnancy.

In the following subsections, first, we explain the algorithms implemented and next, the architectures proposed to design the decision support system are described; finally, the database used to train and test the system is also presented.

3.1. Aid decision algorithms

3.1.1. MLP (Auto MLP)

Artificial neural networks (ANNs) are algorithms based on computational models, which allow solving complex regression and classification problems that cannot be solved by traditional statistical methods [17].The use of ANNs provides features such as nonlinearity, high parallelism, robustness, failure tolerance, learning and ability to handle imprecise and fuzzy information [18]. The multilayer perceptron (MLP) is one of the most implemented supervised ANNs in medical decision support systems. Its architecture is based on a network of nodes (neurons) grouped into three or more structural units called layers (see Fig. 1). These layers are classified as:

1. Input layer: In this layer, the nodes represent each one of the input variables (binary or analog) that intervene in the process and have an influence on the output result.
2. Hidden layers: The intermediate layers are responsible for adjusting the intensity of interaction (synaptic weight) between each node of the previous layer (presynaptic) and those of the next layer (postsynaptic). The number of hidden layers and neurons in each of them remain controversial. In general term, one or two hidden layers are sufficient to solve any non-linear problem. However, if a greater precision is required, a third hidden layer can be considered, although this will increase the complexity of the network and the necessary training time.
3. Output layer: In this layer, the nodes represent each one of the output variables (dependent variables). The transfer function more used in the MLP for the neurons in the hidden layers is the sigmoid transfer function. It reduces an infinite input range into a finite output range. Sigmoid function is characterised by the fact that their

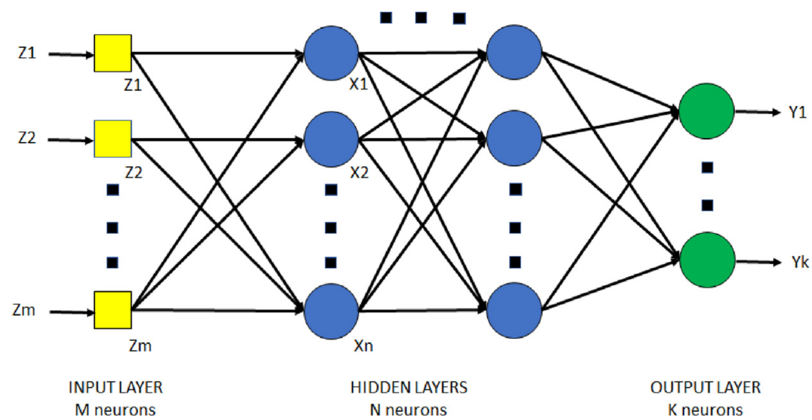

Fig. 1. MLP architecture.

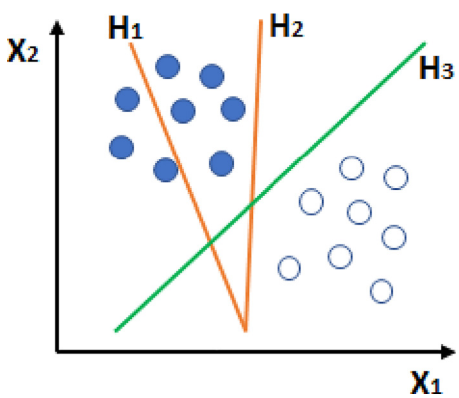

Fig. 2. Classification in a two-dimensional plane by using hyperplanes.

slopes must approach zero [18].

Recently, a variant of this neural network known as Auto MLP has arisen. It is a simple algorithm for both learning rate and size adjustment of neural networks during training. It maintains a small ensemble of networks that are trained in parallel with different rates and different numbers of hidden units. The algorithm combines ideas from genetic algorithms and stochastic optimization. This algorithm automatically finds the optimal parameters of the network configuration (e.g. number of hidden layers, number of neurons in each layer), the most suitable preprocessing method (e.g. normalization) and also optimal training parameters (e.g. number of epochs) [19].

3.1.2. Deep Learning

Deep Learning is a subfield of ANNs that is having a great impact nowadays. Deep learning methods are formed by a set of automatic learning computational algorithms that use multiple nonlinear transformations with the purpose of modeling representations with a high level of abstraction [20]. As described above, ANNs are structured into layers of neurons interconnected with each other. Neurons get activated through weighted connections from previously active neurons. The learning of the network is based on finding those weights that get the network to show the expected behavior. However, depending on the number of layers and intermediate neurons, and how they are connected to each other, this adjustment of weights may require long causal chains of computational stages. In these models with several successive nonlinear layers of neurons, the aggregate activation of the ANN is modified in each stage through multiple nonlinear transformations. Deep Learning is about accurately assigning credit across many such stages.

Deep Learning classifier is based on a multi-layer feed-forward artificial neural network that is trained with stochastic gradient descent using back-propagation. The network can contain a large number of hidden layers consisting of neurons with tanh, rectifier and maxout activation functions. Each compute node trains a copy of the global model parameters on its local data with multi-threading and contributes periodically to the global model via model averaging across the network.

3.1.3. Support Vector Machine (LibSVM)

Support Vector Machines (SVM) are a set of supervised machine learning algorithms developed by Vladimir Vapnik [21]. SVM allow solving classification and regression problems by constructing a hyperplane or set of hyperplanes in a high-or infinite-dimensional space. Given a set of data belonging to two different categories, SVM model represents each of input data as points in space and finds all those hyperplanes capable of separating and classifying them according to their class, as well as it is able to predict whether a point will belong to one category or another (see Fig. 2). SVM seek to solve an optimisation problem, since the optimal solution will be given by the hyperplane that maximizes the distance or margin between the different classes or categories (so-called functional margin), since in general the larger the margin the lower the generalization error of the classifier.

However, not all problems can be represented in a two-dimensional plane and neither perform a correct classification from a vector, straight plane or n-dimensional hyperplane. This may be the case of those problems that have more than two predictor variables, classifications in more than two categories, non-linear separation curves or where the data set cannot be completely separated. This limitation can be overcome through the Kernel functions, which are responsible for projecting the data set to a space with a larger dimension, thus increasing the computational capacity of the SVM (see Fig. 3). For this purpose, Chih-Chung Chang and Chih-Jen Lin [22] developed in 2001 a powerful library for SVM (LibSVM), that supports internal multiclass learning, overcoming the limitation of linear support vector machines.

3.1.4. Naive Bayes classifier

Naives Bayes classifier is a probabilistic classifier based on the

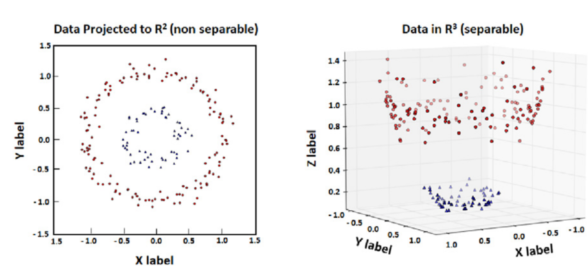

Fig. 3. Representation of a tridimensional classification by using Kernel functions.

theorem enunciated by Thomas Bayes in the theory of probability in 1763, where the Bayes formula or rule is established. It allows calculating the probability of an event A occurring, given a condition B:

$$P(A|B) = \frac{P(A)P(B|A)}{P(B)}$$

However, this classifier takes initial hypotheses of independence between the different predictor variables (attributes) that simplify the problem, contrary to the Bayesian classifiers, which establish dependency representations between variables (Bayesian networks) with probabilistic reasoning. In other words, this classifier assumes that the presence or absence of a certain attribute is not directly related to the presence or absence of any other, and therefore, all the variables that influence the final result are independent of each other [23]. The advantage of the Naive Bayes classifier is that it requires only a small amount of training data to estimate the mean and variances of the variables necessary for classification.

3.2. Decision support systems

Using the algorithms previously explained, we propose two different architectures for the decision support system. The first one is using classifiers individually to decide which treatment will be better; the second one is the combination of the classifiers in one integrated classifier. In the following subsection the architectures are explained and next the design parameters used are presented.

3.2.1. Decision architectures

As it was explained previously, there are three alternatives for the treatment of a patient who has developed an ectopic pregnancy. According to the characteristics of the patient and the variables described above, an ectopic pregnancy can be treated through expectant management, medication (pharmacological treatment) or surgery. While expectant management does not involve any risk for the patient, and surgery is presented as an irreversible and definitive option, the pharmacological treatment, if it is not adequate, may finish leading to a subsequent intervention. These cases constitute the worst possible scenario, since in addition to cause a greater psychological impact on the patient, delaying and making more expensive the clinical process. Initial treatment is decided according to medical criteria. Thus, patients' treatment is assigned with a single consensual decision by the medical team (see Fig. 4). In this study, we have simulated this procedure by using a single classifier (SC) developed for each of the four aid decision algorithms describe above, in order to compare the metrics of right and wrong treatment classifications.

However, there are enough cases where initial treatment must be changed due to the evolution of the patient or misclassification. To solve misclassification cases and improve the ratio of initial right classifications, we have also developed a three-stages classifier (3SC) as a result of the concatenation of three sub-classifiers C1, C2 and C3, as it is shown in Fig. 5. This new classifier pretends to improve the accuracy through the reclassifications in that cases where the surgery is not decided as initial treatment. The first classification (C1) is established between those patients whose case requires a surgical intervention and those who do not. Next, our procedure is focused on patients who do not require surgery, to classify them among those who must follow an expectant management and those who should receive some type of pharmacological treatment (C2). However, as mentioned above, certain cases in which the patient was initially treated through medication may end up leading to surgery, either because the treatment has not taken effect or because it is a wrong initial diagnosis. This is a scenario to be avoid. Therefore, a third classifier (C3) is responsible for reanalysing all those cases proposed for medication and reclassify these cases among those that require surgery or medication. Thus, C3 acts as a corrective classifier of the classifications made by C1 and C2.

3.2.2. Design parameters

The experimental phase consisted in the development of the classifier using the four previously described algorithms. Next, these algorithms were trained and tested using the database provided. With this purpose, we have used a data mining software called RapidMiner Studio©. Table 1 shows some of the design parameters of the different algorithms that we established empirically. Cross validation has been used for the training of the model. This technique consists of using for the validation, the same part of the database used for the training of the model. The training database is divided into a k equal parts or subsets. Each subset is used for validation, while the model is trained with the remaining $k - 1$ subsets, which implies a total of k iterations. The final performance is obtained when calculating the average performance of all iterations. For the training and validation of our models we used 90% of the dataset and a 10-fold cross-validation ($k = 10$). The algorithm is tested with the remaining 10%.

3.3. Clinical database

The ectopic pregnancies database used in this study consists of 406 cases of tubal ectopic pregnancies collected at the Department of Obstetrics and Gynecology of the University Hospital "Virgen de la Arrixaca" in the Murcia Region (Spain) from November of 2010 to September 2015. Women involved were patients attended to the emergency room or to the first-trimester-pathology unit, with ages between 16 and 46 years. Personal and medical variables were obtained from each patient. In addition, a 2-D transvaginal ultrasound were performed using a Voluson E-8 with 4–9 MHz transducer (General Electric Healthcare, USA) by an expertise trained gynaecologist. The original database has a total of 33 attributes as it is described in Table 2. Initial treatment were decided according to medical criteria (clinical status, ultrasound findings and beta-hCG levels).

In the preprocessing of our database, we had to face three common and recurrent problems in the clinical databases: unbalanced classes, data entry mistakes and missing data. The first overcome was solved by using the "Random Over-Sampling technique" [24]. It consists of randomly replicating the records of the minority class until to obtain a greater representation.

There are different methods to work with missing data based on statistical or computational models. However, the simplest and most

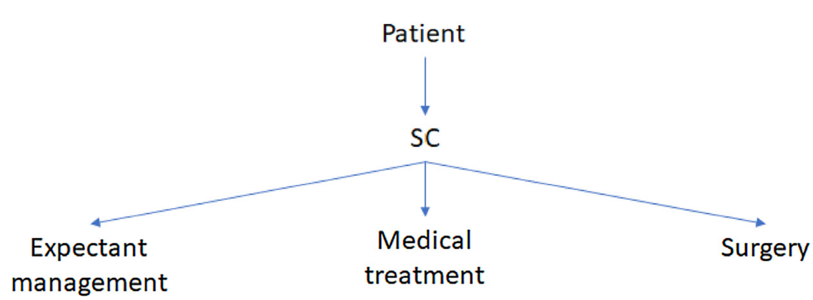

Fig. 4. Single classifier (SC).

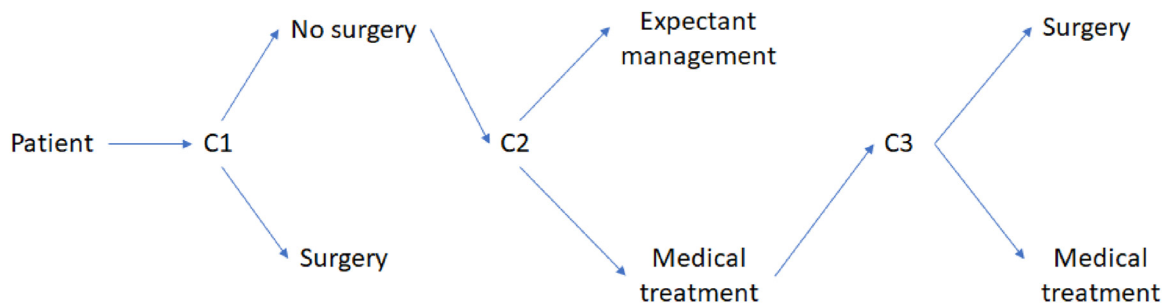

Fig. 5. Three-stages classifier (3SC).

Table 1
Design parameters.

Auto MLP	Value
Training cycles	10
Number of generations	10
Number of esemble MLPs	4
Deep learning	**Value**
Activation function	Rectifier
Number of epochs	10
Hidden layer sizes	50
Epsilon (learning rate)	1.0E−8
Rho (momentum)	0.99
SVM	**Value**
SVM type	C-SVC
Kernel type	RBF
Epsilon (tolerance)	1.0E-3
Naive Bayes	**Value**
Laplacian correction	True

Table 2
Database attributes.

Attribute	Description
Name	–
Date of birth	–
Age	–
Weight (kg)	–
Height (cm)	–
CRN	Clinical record number
Diagnostic date	–
Last menstrual cycle date	–
G	Number of previous gestations
C	Number of previous cesarean deliveries
A	Number of previous abortions
E	Number of previous ectopics pregnancies
VIP	Number of voluntary interruptions of pregnancy
Location	Location of fertilised egg: Fallopian tube, cervical, abdominal, or cornual
Size 1 (mm)	Fertilised egg size 1
Size 2 (mm)	Fertilised egg size 2
Free Liquid	Presence of free liquid: yes, no or scarce
Number of visits	Number of visits to medical consultation
Number of Methotrexate doses	–
Initial beta-HCG	Initial beta-HCG level
beta-HCG 4	beta-HCG level on day 4
beta-HCG 7	beta-HCG level on day 7
beta-HCG 21	beta-HCG level on day 21
beta-HCG 28	beta-HCG level on day 28
beta-HCG 35	beta-HCG level on day 35
beta-HCG 42	beta-HCG level on day 42
beta-HCG 49	beta-HCG level on day 49
Hospitalisation reason	Surgery, observation, pain, MTX dose or control
Entry date	Hospitalisation date
Surgery reason	Accident, live embryo, MTX failure or other reason
Initial treatment	Surgery, medical treatment or expectant management
Final treatment	Surgery, medical treatment or expectant management

realistic approach is to ignore them, as in the case of records with data entry mistakes. By applying these considerations, it was possible to expand the sample and the number of registrations up to 732 records of patients registered correctly. The number of patients whose final treatment was “Surgery” is 240, while 266 received “Medical Treatment” and 226 “Expectant management”. On the other hand, not all the attributes could be used in our study. Some of them (e.g. height, weight, …) did not exist in a large number of records. Others (e.g. beta-HCG 4, beta-HCG 7, …) were registered when the treatment had already been decided, which cannot help to predict it. Finally, some of these variables were discarded because they had no impact on the type of treatment (e.g. diagnostic date, NHC, name, …). In total, 12 variables from the original database were selected (age, G, P, C, A, E, IVE, location, ectopic mm1 at diagnosis, ectopic mm^2 at diagnosis, free liquid and initial beta-HGC) and an additional one calculated as:

$$\text{"Gestional age (weeks)"} = \frac{\text{"Diagnostic date"} - \text{"Last menstrual cycle date"}}{7}$$

4. Results and discussion

During the experimental phase, we established the following objectives to study the implemented models: (1) assessing if our three-stage classifier improves the results of the single classifier for each aid algorithm; (2) analysing which of them was able to classify and predict the most suitable treatment with greater accuracy. For each possible treatment (surgery, medical treatment and expectant management), two parameters have been evaluated from the metrics provided by the confusion matrix: sensitivity (SE) and specificity (SPE). The sensitivity indicates the ability to predict or classify as positive/true those cases that really are. It calculates the percentage of true positive classifications (TP) made over the total amount of real cases, that is, TP and false negatives (FN):

$$\text{Sensitivity}(\%) = \frac{TP}{TP + FN} \times 100\%$$

For its part, the specificity indicates the ability to predict as negative/false those cases that really are, in other words, the percentage of negative true classifications (TN) made over the total amount of real cases, that is, TN and false positives (FP):

$$\text{Specificity}(\%) = \frac{TN}{TN + FP} \times 100\%$$

Moreover, we have calculated an additional global metric, the

Table 3
Single classifier metrics (SC).

Auto MLP	RP	TP	FP	TN	ACC	SE	SPE
Surgery	240	210	33	430	87.4%	87.5%	92.8%
Medical treatment	266	232	29	408		87.2%	93.3%
Expectant management	226	198	30	442		87.6%	93.6%
Average						87.4%	93.2%
Deep learning							
Surgery	240	171	103	250	57.5%	71.2%	70.8%
Medical treatment	266	248	208	173		93.2%	45.4%
Expectant management	226	2	0	419		0.8%	100%
Average						55%	72%
SVM							
Surgery	240	199	24	458	89.7%	82.9%	95%
Medical treatment	266	232	27	425		87.2%	94%
Expectant management	226	226	24	431		100%	94.7%
Average						90%	94.5%
Naive Bayes							
Surgery	240	138	39	361	68,2%	57.5%	90.2%
Medical treatment	266	188	56	311		70.6%	84.7%
Expectant management	226	173	138	326		76.5%	70.2%
Average						68.2%	81.7%

accuracy (ACC), as the percentage of correct predictions:

$$\text{Accuracy}(\%) = \frac{\text{TP} + \text{TN}}{\text{Total cases}} \times 100\%$$

Regarding the first objective, the comparison between Tables 3 and 4 show that our classifier improved the accuracy, sensitivity and specificity for SVM and MLP algorithms. However, Naive Bayes algorithm hardly improved the results obtained with the single classifier and worsened with Deep Learning algorithm in terms of accuracy and sensitivity.

For the second study target, as shown in table 4, the best classification results were obtained with SVM, with values of accuracy, sensitivity and specificity about 96,1%, 96% and 98% respectively. The MLP also presents good results in the classification, although its metrics was lower than with the SVM. On the other hand, the results offered by the classifiers developed with Naive Bayes and Deep Learning models were significantly lower than those obtained with SVM and MLP. Particularly striking is the poor results obtained with Deep Learning. A more detailed analysis of the results of this algorithm showed that the classifier tended to classify as "surgery" all those cases that did not correspond to "medical treatment", offering a very low accuracy in the cases that corresponded to "expectant management".

Table 4
Three stages classifier metrics (3SC).

Auto MLP	RP	TP	FP	TN	ACC	SE	SPE
Surgery	240	224	40	441	90.8%	93.3%	91.6%
Medical treatment	266	237	19	428		89%	95.7%
Expectant management	226	204	8	461		90.2%	98.2%
Average						90.8%	95.2%
Deep learning							
Surgery	240	236	278	172	55.7%	98.3%	38.2%
Medical treatment	266	166	46	242		62.4%	84.7%
Expectant management	226	6	0	402		2.6%	100%
Average						54.4%	81.7%
SVM							
Surgery	240	226	10	478	96.1%	94.1%	97.9%
Medical treatment	266	262	6	442		98.4%	98.6%
Expectant management	226	216	12	488		95.5%	97.6%
Average						96%	98%
Naive Bayes							
Surgery	240	160	64	342	68.5%	66.6%	84.2%
Medical treatment	266	188	54	314		70.6%	85.3%
Expectant management	226	154	112	348		68.1%	75.6%
Average						68.4%	81.7%

Table 5
AUTO MLP confusion matrix.

Auto MLP	true surgery	True no surgery	ACC	SE	SPE
Three stages classifier					
Surgery	224	51	90.8%	93.3%	
No surgery	16	441			89.6%
Single classifier					
Surgery	212	30	92.7%	88.3%	
No surgery	28	462			93.9%

Table 6
Deep Learning confusion matrix.

Deep learning	True surgery	True no surgery	ACC	SE	SPE
Three stages classifier					
Surgery	236	320	55.7%	98.3%	
No surgery	4	172			34.9%
Single classifier					
Surgery	165	52	82.6%	68.7%	
No surgery	75	440			89.4%

At this point, it should be noted that from our database, the initial treatment was correctly assigned by the medical team in the 87.6% of the cases. This accuracy is slightly lower than best results of our single classifier (SVM, 89.7%). However, is greatly improved with two of three-stages classifiers, especially in the case of SVM algorithm.

Finally, Table 5–8 shows the confusion matrices for a two categories classification (surgery, no surgery) of the algorithms implemented. For this type of classification, SVM algorithm was also the one with the highest precision, improving both sensitivity and specificity. In general terms, the rest of the algorithms improved the sensitivity and worsened in terms of accuracy and specificity.

5. Conclusions

Ectopic pregnancies can present important complications if they are not diagnosed in time, such as the rupture of the fallopian tube or other structure, causing internal hemorrhages. That is why the treatment should start as soon as the diagnosis is confirmed. However, the choice of treatment may be controversial. Several factors come into play when deciding which treatment is the most suitable. A wrong initial diagnosis may increase the risk to the patient. In this paper, a clinical decision support system is presented to classify the initial treatment to be followed by a woman suffering an ectopic pregnancy. This clinical decision support system has been evaluated with four different algorithms: MLP, SVM, Deep Learning and Naive Bayes. The decision support system has been designed for a single classifier and a multiple classifier (3 stages). The single stage classifier presented lower accuracy than the obtained by the medical criteria (87.6%), except for the one developed with the SVM that slightly improved the classification accuracy (89.7%). However, the results were significantly better in the case of the three-stage classifier, both in the case of MLP and SVM reaching

Table 7
SVM confusion matrix.

SVM	True surgery	True no surgery	ACC	SE	SPE
Three stages classifier					
Surgery	226	18	96.1%	95.7%	
No surgery	10	478			96.3%
Single classifier					
Surgery	212	30	92%	88.3%	
No surgery	28	462			93.9%

Table 8
Naive Bayes confusion matrix.

Naïve Bayes	True surgery	True no surgery	ACC	SE	SPE
Three stages classifier					
Surgery	160	150	68.5%	66.6%	
No surgery	80	342			69.5%
Single classifier					
Surgery	143	68	77.4%	59.5%	
No surgery	97	424			86.1%

accuracy values in initial diagnosis of 90.8% and 96.1% respectively. The results prove that Deep Learning and Naive Bayes algorithms were ineffective models but the SVM and MLP can be useful to help doctors in their decisions about initial treatment.

Authors contributions

Conception of the study and design of study: Daniel Ruiz Fernández and Alberto de Ramón Fernández Acquisition of data:

María Teresa Sánchez Prieto.

Analysis and/or interpretation of data: Daniel Ruiz-Fernández, Alberto de Ramón-Fernández and María Teresa Sánchez Prieto.

Drafting of the manuscript: Alberto de Ramón-Fernández.

Revising the manuscript critically for important intellectual content: Daniel Ruiz-Fernández and Alberto De Ramón Fernández.

Approval of the version of the manuscript to be published: Daniel Ruiz Fernández, Alberto de Ramón Fernández and María Teresa Sánchez Prieto.

Competing interests statement

The authors declare no competing interests in this paper.

Acknowledgments

This work has been granted by the Ministerio de Economá y Competitividad of the Spanish Government (ref. TIN2014-53067-C3-1-R) and Alberto De Ramón Fernández is supported by grant BES-2015-073611.

References

[1] J. Bouyer, J. Coste, H. Fernandez, J.-L. Pouly, N. Job-Spira, Sites of ectopic pregnancy: a 10 year population-based study of 1800 cases, Hum. Reprod. 17 (12) (2002) 3224–3230.
[2] M. Atri, C. Leduc, P. Gillett, P.M. Bret, C. Reinhold, G. Kintzen, A.E. Aldis, M. Thibodeau, Role of endovaginal sonography in the diagnosis and management of ectopic pregnancy, Radiographics 16 (4) (1996) 755–774.
[3] C. Silva, M.D. Sammel, L. Zhou, C. Gracia, A.C. Hummel, K. Barnhart, Human chorionic gonadotropin profile for women with ectopic pregnancy, Obstet. Gynecol. 107 (3) (2006) 605–610.
[4] R.J. Morlock, J.E. Lafata, D. Eisenstein, Cost-effectiveness of single-dose methotrexate compared with laparoscopic treatment of ectopic pregnancy, Obstet. Gynecol. 95 (3) (2000) 407–412.
[5] C. Farquhar, Ectopic pregnancy, Lancet 366 (9485) (2005) 583–591.
[6] L. Craig, S. Khan, Expectant management of ectopic pregnancy, Clin. Obstet. Gynecol. 55 (2) (2012) 461–470.
[7] S.K. Van Den Eeden, J. Shan, C. Bruce, M. Glasser, Ectopic pregnancy rate and treatment utilization in a large managed care organization, Obstet. Gynecol. 105 (5) (2005) 1052–1057.
[8] F. Kamwendo, L. Forslin, L. Bodin, D. Danielsson, Epidemiology of ectopic pregnancy during a 28 year period and the role of pelvic inflammatory disease, Sex. Transm. Infect. 76 (1) (2000) 28–32.
[9] R.O. Schwartz, D.P. Di, beta-HCG as a diagnostic aid for suspected ectopic pregnancy, Obstet. Gynecol. 56 (2) (1980) 197–203.
[10] J.S. McCollum, Diagnostic curettage in the evaluation of ectopic pregnancy, Am. Fam. Physician 63 (2) (2001) 220–222.
[11] T.J. Bright, A. Wong, R. Dhurjati, E. Bristow, L. Bastian, R.R. Coeytaux, G. Samsa, V. Hasselblad, J.W. Williams, M.D. Musty, et al., Effect of clinical decision-support systems: a systematic review, Ann. Intern. Med. 157 (1) (2012) 29–43.
[12] S.M. Sadat-Hashemi, A. Kazemnejad, C. Lucas, K. Badie, Predicting the type of pregnancy using artificial neural networks and multinomial logistic regression: a comparison study, Neural Comput. Appl. 14 (3) (2005) 198–202.
[13] G. Mello, E. Parretti, A. Ognibene, Prediction of the development of pregnancy-induced hypertensive disorders in high-risk pregnant women by artificial neural networks, Clin. Chem. Lab. Med. 39 (9) (2005) 801–805.
[14] E. Tejera, M. Jose areias, A. Rodrigues, A. Ramoa, J. Manuel Nieto Villar, I. Rebelo, Artificial neural network for normal, hypertensive, and preeclamptic pregnancy classification using maternal heart rate variability indexes, J. Matern.-Fetal Neonat. Med. 24 (9) (2011) 1147–1151.
[15] K. Paydar, S.R.N. Kalhori, M. Akbarian, A. Sheikhtaheri, A clinical decision support system for prediction of pregnancy outcome in pregnant women with systemic lupus erythematosus, Int. J. Med. Inform. 97 (2017) 239–246.
[16] E. Grossi, F. Veggo, A. Narzisi, A. Compare, F. Muratori, Pregnancy risk factors in autism: a pilot study with artificial neural networks, Pediatr. Res. 79 (2) (2016) 339.
[17] K. Hornik, M. Stinchcombe, H. White, Multilayer feedforward networks are universal approximators, Neural Netw. 2 (5) (1989) 359–366.
[18] I.A. Basheer, M. Hajmeer, Artificial neural networks: fundamentals, computing, design, and application, J. Microbiol. Methods 43 (1) (2000) 3–31.
[19] T. Breuel, F. Shafait, Automlp: simple, effective, fully automated learning rate and size adjustment, The Learning Workshop, vol. 4, Utah, 2010, p. 51.
[20] Y. Bengio, A. Courville, P. Vincent, Representation learning: a review and new perspectives, IEEE Trans. Pattern Anal. Mach. Intell. 35 (8) (2013) 1798–1828.
[21] C. Cortes, V. Vapnik, Support-vector networks, Mach. Learn. 20 (3) (1995) 273–297.
[22] C.-C. Chang, C.-J. Lin, Libsvm: a library for support vector machines, ACM Trans. Intell. Syst. Technol. (TIST) 2 (3) (2011) 27.
[23] S. Xu, Bayesian naïve bayes classifiers to text classification, J. Inf. Sci. 44 (1) (2016) 48–59.
[24] R. Dubey, J. Zhou, Y. Wang, P.M. Thompson, J. Ye, A.D.N. Initiative, et al., Analysis of sampling techniques for imbalanced data: an n = 648 ADNI study, NeuroImage 87 (2014) 220–241.