3P: Personalized Pregnancy Prediction in IVF Treatment Process

Asli Uyar[1], H. Nadir Ciray[2], Ayse Bener[1], and Mustafa Bahceci[2]

[1] Boğaziçi University, Department of Computer Engineering, Istanbul, Turkey
{asli.uyar,bener}@boun.edu.tr
[2] Bahceci Women Health Care Centre, Istanbul, Turkey
{nadirc,mbahceci}@superonline.com

Abstract. We present an intelligent learning system for improving pregnancy success rate of IVF treatment. Our proposed model uses an SVM based classification system for training a model from past data and making predictions on implantation outcome of new embryos. This study employs an embryo-centered approach. Each embryo is represented with a data feature vector including 17 features related to patient characteristics, clinical diagnosis, treatment method and embryo morphological parameters. Our experimental results demonstrate a prediction accuracy of 82.7%. We have obtained the IVF dataset from Bahceci Women Health, Care Centre, in Istanbul, Turkey.

Keywords: In-vitro fertilization, Embryo implantation prediction, Classification, Support vector machines.

1 Introduction

Infertility is defined as couple's biological inability to conceive pregnancy after at least 12 months of regular, well-timed sexual intercourse without contraception. In-vitro fertilization (IVF) [1] is a process during which female germ cells (oocytes) are inseminated by sperm under laboratory conditions. After 1992 the IVF process is combined with intra-cytoplasmic sperm injection (ICSI) method [2] during which a single sperm cell is injected into the cytoplasm of the oocyte. Fertilized oocytes are cultured between 2-6 days under laboratory conditions during which embryonic growth is observed and selected embryos are transferred into the woman's womb. Selection of embryos with highest implantation (i.e. attachment of the embryo to the inner layer of the womb) potential is crucial for achieving a successful pregnancy.

There are various embryo and patient characteristics which may affect the outcome of an IVF cycle. The conventional and most common way of selecting high quality embryos is to inspect their morphologies. However, non-automated analysis of various patient and embryo parameters is a challenge. The main objective of this study is analyzing the underlying factors of embryo implantation and thus to provide a prediction model. To accomplish this, we propose an intelligent system that uses machine learning methods. These methods use available data for learning and establishing a prediction model.

D. Weerasinghe (Ed.): eHealth 2008, LNICST 1, pp. 58–65, 2009.

1.1 Machine Learning Methods in IVF Data Analysis

Existing studies on IVF data analysis heavily focus on statistical relationships between clinical variables and pregnancy results [3][4]. These studies evaluate the most important predictors for the pregnancy outcome or predict implantation rate depending on limited number of embryo related features. Identification of correlations between input IVF data and pregnancy outcome provide a knowledge base, but selection of embryos with highest implantation potential still depends on decision of embryologists, which may vary from one to another. Therefore, intelligent learning systems can be more suitable to assist embryologists in making the right decision. The machine learning methods are used in medicine for providing decision support on diagnosis and treatment process.

IVF treatment is a complex and costly process requiring decision support and future predictions in certain stages. Because of the difficulty faced in manual observation of multiple variables and examination of nonlinear correlations between features, IVF process requires more advanced prediction models. A machine learning system can automatically analyze large IVF databases to train a model and provide future predictions. Such a system would speed up the embryo selection process and possibly improve the number of successful pregnancies.

On the contrary to the importance and emergence of intelligent decision support systems in IVF process, the related literature is limited. As a preliminary study, Jurisica et. al. represent a case-based reasoning system that exploits past experiences to suggest possible modifications to an IVF treatment plan [5]. Later, a decision tree model is applied to express relationships of features characterizing the "take of baby" and "no-take of baby" classes of embryo batches [6]. Decision trees again used for prediction of pregnancy outcome from clinical IVF data with an accuracy of 67.4% [7]. Trimarchi et. al. build decision tree and logistic regression models and reported 75% and 74% accuracy rates respectively [8]. Another research considers automated recognition of embryos suitable for transfer and compare the recognition of experts with that of a machine programme [9]. The most recent study on implantation prediction proposes a Bayesian classification system for embryo selection and reported an accuracy of 71.4% [10].

Existing approaches generally consider transfer of embryo batches including two or three embryos. However, it is not possible exactly to know which embryo of the batch is implanted. Such an ambiguity decreases the reliability.

1.2 Proposed System for Embryo Implantation Prediction

In IVF pregnancy prediction machine learning methods are constructed as cycle based methods. An IVF cycle consists of controlling the follicular stimulation by external administration of hormones, aspirating oocytes from woman's ovaries, inseminating the oocytes with sperm cells in vitro, letting them grow in the laboratory and transferring the embryos into the womb. The number of embryos to be transferred varies. The aim of this prediction model is to identify the embryo to implant according to patient and embryo characteristics. Therefore this model could also be useful to minimize high order pregnancies.

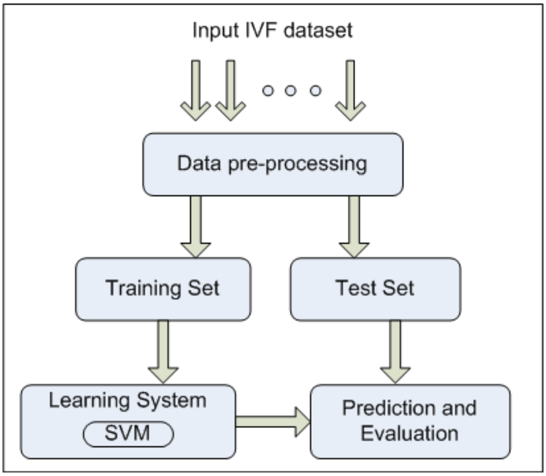

Fig. 1. Schematic representation of proposed learning system for embryo implantation prediction

Cycle based approaches use features of each embryo in a given cycle as a feature vector of all embryos in that cycle. In this case we do not know which embryo got implanted and this causes loss of important information as far as a prediction model is concerned. In this research, we build an embryo based learning system. We classify embryos according to their implantation potentials. In order to avoid the information loss we have only considered clear cases where all of the transferred embryos implanted or not implanted.

The prediction performance of a machine learning based model is based on data and the algorithm to use [14]. In IVF domain, we need to build a model which learns from massive data and comes up with results that can be generalized. Therefore it is important to decide which features to select in order to keep the information content high so that the predictive model returns accurate results. To increase the information content, we chose the most relevant parameters as data features. We used Support Vector Machines (SVM) as the classifier for implantation prediction to learn a model from past IVF data and make predictions on new embryos. The schematic representation of the proposed system can be seen in Fig. 1.

2 Dataset Characteristics

Infertility is a social matter as well as a medical disorder. Because of social and ethical reasons in every country some legislative rules have been defined. Usually, the restrictions apply for donation, embryo manipulation, number of embryos to be transferred in each cycle etc. Besides the legal procedures in countries, every IVF clinic applies different technologies and methodologies in practice. Because of this variety, each clinic has distinctive IVF databases.

In this study, we will analyze the IVF procedure and related database of Bahceci IVF Clinic in Istanbul. The dataset has data on 3000 patients of numerous cycles and embryos collected since 1997. Compared to the ones used in the literature Bahceci dataset uses more numerical values than categorical values and this makes it a more objective dataset to build a prediction model on. We have constructed a dataset from existing database with selected features for certain cases. The proposed classification system is an embryo centered approach and dataset contains a data feature vector for each embryo rather than each cycle. Dataset features and data types are given in Table 1. The features have been selected depending on experiences of senior embryologists in Bahceci IVF Clinic and related studies [11][12][13].

Table 1. Selected dataset features for each embryo feature vector

Dataset Features	**Data Type**
Patient Characteristics	
Woman age	Numerical
Primary or secondary infertility	Categorical
Clinical Diagnosis and Treatment Protocol	
Infertility factor	Categorical
Treatment protocol	Categorical
Duration of stimulation	Numerical
Follicular stimulating hormone dosage	Numerical
Estradiol level	Numerical
Endometrium thickness	Numerical
Sperm quality	Categorical
Embryo Related Data	
Early cleavage morphology	Categorical
Transfer day	Categorical
Number of cells	Categorical
Nucleus characteristic	Categorical
Fragmentation	Categorical
Blastomers	Categorical
Cytoplasm	Categorical
Thickness zona pellucida	Categorical

3 SVM Classification

Support vector machines (SVMs) are a set of related supervised learning methods used for classification and regression. Their common factor is the use of a technique known as the “kernel trick” to apply linear classification techniques to non-linear classification problems.

In classification problem, classifying data as part of a machine-learning process is interesting. When the data points in the set are multidimensional, the classification can be carried out with separating the points by using hyperplanes. This form of classification is linear and the classification is required to be neat with maximum distance to the closest data point from both classes where the distance is the margin [15]. If the hyperplane has this margin property, it is called maximum-margin hyperplane, as are the vectors that are closest to this hyperplane, which are called the support vectors.

Given a set of training data pairs (x_i, y_i), where x_i is the input feature vector and y_i is the class label, the aim of the SVM classifier is to estimate a decision function by constructing the optimal separating hyperplane in the feature space. The key idea of SVM is to map the original input space into a higher dimensional feature space in order to achieve a linear solution. This mapping is done using kernel functions. Final decision function is in the form:

$$f(x) = \left(\sum_i \alpha_i y_i K(x_i \cdot x) + b \right) \tag{1}$$

where $K(x_i \cdot x)$ is the Kernel transformation. The training samples whose Lagrange coefficients α_i are non-zero are called SVs and the decision function is defined by only these vectors.

3.1 Performance Measures

We have used probability of detection (pd) and probability of false alarm (pf) as the performance measures [16]. Formal definitions for these performance criteria are given in Equations 2 and 3 respectively and they are derived from the confusion matrix given in Table 2. *pd* is a measure of accuracy for correctly detecting the embryos that will implant. Therefore, higher *pd*'s are desired. *pf* is a measure for false alarms and desired to have low values.

$$pd = (A)/(A + C) \tag{2}$$

$$pf = (B)/(B + D) \tag{3}$$

$$FN = (C)/(A + C) \tag{4}$$

As a measure of performance false positives (FP) and false negatives (FN) are also used. Since, the *FP* rate is the same as *pf* measure; we additionally consider analysis of *FN* rate in this study. Equation 4 represents the formulation of *FN* rate which is the proportion of implanted embryos that were erroneously reported as not-implanted. *FN* is an error measure for missing the embryos that will implant. So, it is critical to reduce FN rate in prediction results.

Table 2. Confusion Matrix

Actual Case	Predicted	
	Implanted	Not-implanted
Implanted	A	C
Not-implanted	B	D

4 Experiments and Results

The IVF dataset used in this study includes 546 embryo records which have been transferred in day 2 or day 3 and each embryo data vector is represented by 17 feature values (Table 1). There are two classes of embryos, class label 1 indicates implantation and class label 0 indicates no-implantation. The distribution of classes over training and test sets are given in Table 3.

Table 3. Number of implanted and not-implanted embryo samples in training and test sets

Case of classes	Training set	Test set	Total
Implanted	218	55	273
Not-implanted	218	55	273
Total	436	110	546

Experiments have been performed using LIBSVM tool [17] in MATLAB environment. Kernel and model parameter selection is crucial for the performance of SVM classifier. We have tested the classifier with linear, polynomial and Gaussian kernels. Gaussian Kernel has been the choice of kernel because of superior performance with default model parameters. In order to optimize the SVM classifier model with Gaussian kernel, the kernel parameter *gamma* and the model parameter *cost* and are searched in the ranges $\left[2^{-5}, 2^{15}\right]$ and $\left[2^{-15}, 2^{3}\right]$, respectively, using a grid search algorithm. Optimal model parameters and prediction results in terms of defined performance measures are given in Table 4.

Table 4. Optimal model settings for SVM classifier and classification results

Kernel	Kernel parameter	Cost	Accuracy	pd	pf	FN
Gaussian	Gamma = 0.078	32	82.7%	80%	14.5%	20%

Table 5 presents the confusion matrix for classification results. SVM classification for implantation prediction resulted in 82.7% overall accuracy with 80%

Table 5. Confusion matrix for implantation prediction results of SVM classifier

		Predicted Results		
	Case of classes	Implanted	Not-implanted	Error rate (%)
Test Set				
Implanted	55	44	11	20.0
Not-implanted	55	8	47	14.5
Total	110	52	58	17.3

probability of implantation detection and 20% of *FN* rate. The most critical performance measure in embryo selection process is *FN* values. We aim to minimize *FN* rate because we don't want to miss the embryos that will implant.

4.1 Threats to Validity

Compared to similar research in the literature we have used a much larger data set, but, in machine learning standards it may be considered small and hence a threat to validity. However, our previous studies in different domains showed that our models prediction performance is good with as little as 100 samples [16]. The IVF datasets differ in each clinic as explained in Section 2. Hence, direct comparison of results with similar studies is not reliable. This may be a threat to validity as to how reliable and good our empirical results are. However, as to comparing our defect predictors with standard results from the data mining community, in prior work, we have checked the efficacy of data mining on standard machine learning datasets such as UCI repository [18]. On average state of the art data miners perform at (pd,pf) = (81%, 20%). This is close to the results we have obtainedvia data mining on IVF embryo and cycle attributes (pd,pf,FN) = (80%, 14.5%,20%).

5 Conclusions and Future Work

We have defined embryo implantation prediction as a 2-class classification problem and explained its emergence for improving IVF treatment success rate. Our approach is different than the existing ones in the literature that we use embryo based prediction rather than cycle based prediction. Our empirical results show that our proposed model predicts correctly in 82.7% of the time. We have proposed an SVM based learning system and our results showed that a learning based intelligent system can predict pregnancies in a personalized manner at least as good as the most experienced embryologist. Moreover we have seen that our proposed model has *FN* rate which is misclassifying the embryos as no implant is lower than the expert's judgement.

Our future work will be to calibrate our model by using different algorithms to further lower the *FN* rate as well as to enlarge the sample size. We will also work on the selection of features to better understand the correlations among them to better personalize the pregnancy prediction in IVF treatment.

Acknowledgments. This research is supported in part by Bahceci Women Health Care Centre.

References

1. Steptoe, P.C., Edwards, R.G.: Birth After Re-implantation of a Human Embryo. Lancet 2, 366 (1978)
2. Van Steirteghem, A.C., Nagy, Z., Joris, H.: Higher Fertilization and Implantation Rates After Intracytoplasmic Sperm Injection. Human Reproduction 8, 1061–1066 (1993)

3. Shen, S., Khabani, A., Klein, N., Battaglia, D.: Statistical Analysis of Factors Affecting Fertilization Rates and Clinical Outcome Associated with Intracytoplasmic Sperm Injection. American Society for Reproductive Medicine 79 (2003)
4. Holte, J., Berglund, L., Milton, K., Garello, C., Gennarelli, G., Revelli, A., Bergh, T.: Construction of an Evidence Based Integrated Morphology Cleavage Embryo Score for Implantation Potential of Embryo Scored and Transferred on day 2 After Oocyte Retrieval. Human Reproduction 22, 548–557 (2007)
5. Jurisica, I., Mylopoulos, J., Glasgow, J., Shapiro, H., Casper, R.F.: Case-Based Reasoning in IVF: Prediction and Knowledge Mining. Artificial Intelligence in Medicine 12 (1998)
6. Saith, R., Srinivasan, A., Michie, D., Sargent, I.: Relationships Between the Developmental Potential of Human In-Vitro Fertilization Embryos and Features Describing the Embryo, Oocyte and Follicle. Human Reproduction Update 4, 121–134 (1998)
7. Passmore, L., Goodside, J., Hamel, L., Gonzalez, L., Silberstein, T., Trimarchi, J.: Assessing Decision Tree Models for Clinical In-Vitro Fertilization Data. Technical report, Dept. of Computer Science and Statistics University of Rhode Island (2003)
8. Trimarchi, J.R., Goodside, J., Passmore, L., Silberstein, T., Hamel, L., Gonzalez, L.: Comparing Data Mining and Logistic Regression for Predicting IVF Outcome. Fertil. Steril. (2003)
9. Manna, C., Patrizi, G., Rahman, A., Sallam, H.: Experimental Results on the Recognition of Embryos in Human-Assisted Reproduction. Reproductive Bimedicine Online 8, 460–469 (2004)
10. Morales, D.A., Bengoetxea, E., Larranaga, B., Garcia, M., Franco, Y., Fresnada, M., Merino, M.: Bayesian Classification for the Selection of In Vitro Human Embryos Using Morphological and Clinical Data. Computer Methods and Programs in Biomedicine 90, 104–116 (2008)
11. Ciray, H.N., Bener, F., Karagenc, L., Ulug, U., Bahceci, M.: Impact of Assisted Hatching on ART Outcome in Women with Endometriosis. Human Reproduction (2005)
12. Ciray, H.N., Ulug, U., Tosun, S., Erden, H.F., Bahceci, M.: Outcome of 1114 ICSI and Embryo Transfer Cycles of Women 40 Years of Age and Over. Biomed Online 13, 516–522 (2006)
13. Ciray, H.N., Tosun, S., Hacifazlioglu, O., Mesut, A., Bahceci, M.: Prolonged Duration of Transfer Does Not Affect Outcome in Cycles with Good Embryo Quality. Fertil. Steril. (2007)
14. Baskeles, B., Turhan, B., Bener, A.: Software Effort Estimation Using Machine Learning Methods. In: ISCIS 2007, Ankara, Turkey (2007)
15. Alpaydın, E.: Introduction to Machine Learning. MIT Press, Cambridge (2004)
16. Menzies, T., Turhan, B., Bener, A., Gay, G., Cukic, B., Jiang, Y.: Implications of Ceiling Effects in Defect Predictors. In: PROMISE 2008, Leipzig, Germany (2008)
17. Chang, C., Lin, C.: Libsvm: A Library for Support Vector Machines (2001)
18. Blake, C., Merz, C.: UCI Repository of Machine Learning Databases (1998), `http://www.ics.uci.edu/mlearn/MLRepository.html`