A model based on Bayesian Network for prediction of IVF Success Rate

Samar Ashrafi Kakhki[1], Behnam Malekara[2], Saeed Rahati Quchani[3], Nayereh Khadem[4]

[1]Department of Artificial Intelligence faculty engineering,
IslamicAzad University, Mashhad Branch, Mashhad, Iran
s-ashrafi@mshdiau.ac.ir
[2] Department of Software faculty engineering,
Islamic Azad University, Mashhad Branch, Mashhad, Iran
malekara@gmail.com
[3] Department of Artificial Intelligence faculty engineering,
Islamic Azad University, Mashhad Branch, Mashhad, Iran
rahati@mshdiau.ac.ir
[4]Mashhad University of Medical Sciences, Mashhad, Iran
khademn@mums.ac.ir

Abstract

In vitrofertilization (IVF) is one of assisted reproductive technique that enables infertile couples to achieve successful pregnancy. Accurateand early prediction of the outcome of the IVF is important for both patients and physicians. Given the uncertainty of the treatment, we propose an intelligent decision support system based on Bayesian Networks. We evaluate the effectiveness of four Bayesian network classifiers as potential tools for predictionthe success/failure rate of IVF using realworld database.Results show thatwe can use Bayesian networks as a new practical approach for predict the outcome of IVF treatment

Key words: Bayesian Networks, Prediction, Data mining, In vitro fertilization

1. Introduction

Infertility is a social onus for women in Iran, who are expected to produce children early within marriage. On average, two out of ten couples suffers from the problem of infertility in Iran. Infertility can cause depression, anxiety, social isolation and sexual dysfunction(Fassino, Piero et al. 2002). Due to this frustrating experience many infertile couples would seek medical help and finally will receive assisted reproductive treatment. One of the most widely used treatments is the in vitro fertilization (IVF) procedure. In this technique, firs oocytes and sperm are retrieved separately for fertilization. In order toobtain more oocytes, ovulatory stimulants such as follicle stimulating hormone (FSH), are often given. After this step, sperm and oocytesare fertilized in vitro. Fertilized egg is called embryo. Embryosare then transferredinto thewoman's uterusafter 48-72 hours and after successful implant, pregnancycan be achieved. With so many variables that affect the outcome of IVF, for clinician estimate theaccuratepregnancy rate is difficult. Intelligent decision support systems may enable IVF practitioners to cope with the complexity of the domain during treatment planning and help them discover relationships among data of each patient that can then be used to improve the pregnancy rate. More over since the IVF treatment is a relatively expensive procedure with the cost in the Iran varying from 150–300 million Rial per cycle (using the patient's own gametes and excluding drugs and freezing of spare gametes), such a decision system helps the patient in deciding whether to go ahead with the procedure or not, based on the probability of success.

Others (Kaufmann, Eastaugh et al. 1997, Jurisica, Mylopoulos et al. 1998, Saith, Srinivasan et al. 1998, Trimarchi, Goodside et al. 2003, Uyar, Ciray et al. 2009, Guh, Wu et al. 2011) analyze various factors to predict the success of therapy using different techniques such as logistic regression, decision trees, artificial neural networks, case-based reasoning, support vector machine and Integrating genetic algorithm and decision tree. However, they have not used Bayesian networks which, due to their double nature (visual way and quantitative), can give other insights of IVF. Bayesian networks provide a visual way to recognize the interactions among variables (both dependent vs. independent and independent vs. independent, they provide an explanation in descriptive terms and also in classification terms). In addition, they provide local probability distributions (quantitative part) which can be traced from beginning to end so it is possible to know why and how a BN arrived at a specific conclusion.

2. Bayesian Networks

A Bayesian networks, describes the joint probability distribution over a set of random variables, with defining a series of probability independence and a series of conditional independences(Murphy 2001). In other words a Bayesian networks define a joint probability distribution over a set of random variables(Mitchell 1997). To describe a Bayesian networks we need to provide two items: topology or structure of the graph, and parameters or the conditional probability tables of all variables. Having these two items, one can demonstrate the joint probability distribution implied among variables. Simply we can say that this joint probability distribution is equal to the product of conditional probability of each variable having its immediate parents. The equation is described in Eq. 1.

$$P(y_1, \dots, y_n) = \prod_{i=1}^{n} P(y_i | parents(Y_i) \qquad (1)$$

3.Data and Material

3.1. The Dataset

The real-world database contains 942 records for this study come from clinical files of the IVF treatment in Montaserieh Research & Clinical Center for Infertility in Mashhad, Iran, during 2011-2012. The dataset contain 13 variables which show the features of patient such as age, follicle-stimulating hormone, total motile spermatozoa, endometrial thickness, embryo transfer, cell number, embryo grade, smokingcigarette, follicle number, embryo number, infertility duration, oocyte number andmenstrual cycle and one class variable(outcome); BHCG can take two different values: positive or negative, which show patient pregnancy test. Since the training dataset is imbalance, learning the model is challenging. If we make the model using imbalanced training dataset, features will tend to majority class. Resampling (Stassopoulou and Dikaiakos 2009) is a method in order to solve the imbalanced dataset in training step. We have two approaches for resampling:

1- Random Oversampling: randomly duplicate the cases with minority class, i.e. cases thatget pregnant(positive BHCG pregnancy test) using IVF until the majority and minority class rate get reach to the same level.
2- Random undersampling: randomly eliminate the cases with majority class, i.e. cases thatdo not get pregnant (negative BHCG pregnancy test) using IVF until the majority and minority class rate get reach to the desirable level.

Our training dataset is as follows:

(a) Original dataset without any resampling.

(b) Oversampling to 50%. Cases that get pregnantare randomly duplicated until they amount to 50% of the total number of cases in the trainingset, i.e. until the two cases are equally represented.

(c) Undersampling to 50%.Cases that do not get pregnantare randomly eliminated until they amount to 50% of the total number of sessions in the training.

Table 1 shows the characteristics of our data sets.

Dataset	Training set	Negative BHCGin training set	Positive BHCG in training set	Test dataset
D1	942	549	393	235
D2	1098	549	549	235
D3	786	393	393	235

Table 1. The dataset used for training system.

3.2. BN classifiers

Suppose that each training sample is s vector of attributes $(x_1, x_2, \dots, x_{v-1}, C)$.The goal of classification is predicting the rightvalue of class variable $(c = x_v)$having $(x_1, x_2, \dots, x_{v-1})$.If performance measure is classification accuracy (percentage of correct predictions on test samples),correctprediction for $(x_1, x_2, \dots, x_{v-1})$ is the class that maximizes $P(c|x_1, x_2, \dots, x_{v-1})$. If we have a Bayesian network over $(x_1, x_2, \dots, x_{v-1}, C)$we could compute these probabilities by inference on it. After Structure of Bayesian network is specified, it is important to estimate parameters in a way that the network can provide the best prediction for value of class variable in test samples.

The Bayesian classifiers used in this research are Naïve Bayes classifier, Bayes-N, MP-Bayes, Greedy.

The Naïve Bayes classifier (NB)(Han and Kamber 2006) learns the conditional probability of each node from a training data set, then it uses Bayes' rule to compute the conditional probability of the class given the set of attributes and select the value of the class with the highest posterior probability for new cases.

Bayes-N (Martínez-Morales, Cruz-Ramírez et al. 2004) is an algorithm for learning Bayesian network from data based on local measures of information gain, which we want to predict the class variable on new cases. Bayes-N induces an ancestral ordering of all the variables generating a directed acyclic graph in which the class variable is a sink variable, with a subset of the explanatory variables as its parents. Bayes-N is constraint-based algorithm.

MP-Bayes (Cruz-Ramirez, Nava-Fernandez et al. 2005) is an algorithm which induces Bayesian network structures from data based on entropy measures. One of the main features of this method is its parsimonious nature: it tends to represent the joint probability distribution underlying the data with the least number of arcs. While other methods that build Bayesian networks tend to overfit the data, MP-Bayes create models that seem to have an adequate trade-off between accuracy and complexity.MP-Bayes is constraint-based algorithm.

Greedy (Cruz-Ramirez, Nava-Fernandez et al. 2005) starts at a specific point (an initial structure) in the structure space, considers all nearest neighbors of the current point, and moves to the neighbor that has the highest score; if no neighbors have higher score than the current point (i.e., we have reached a local maximum), the algorithm stops. Greedy is a search and scoring algorithm.

4. Research Methodology

We use Bayesian Networks as supervised learning(Friedman and Goldszmidt 1999) for the task of classification of our dataset into two different classes: positive BHCG and negative BHCG. We chose four types of BN classifiers to check the performance of them. We divided the data set into a train set and a test set by using the 80%/20% train/test ratio as we showed in table 1.All the four BN classifiers use training data to learn model then we use test set to evaluate the performance of the classifiers. We used three databases as we discussed beforein section 3.1. First we use the 942-case original database without resampling method (D1) as input for all the classifiers to learn a classification model. The results of this experiment can be found in table 2.Second, we use the database with oversampling method (D2). The results of this experiment can be found in table 3 and finally we use the database with undersampling (D3). The results of thisare shown in table4.

5. Results and Analysis

In this section we present the result of the experimental methodology. For evaluating the classifiers we use performance measures such as sensitivitywhich is the ability to correctly predictpatients who actually get pregnant, specificity which is the ability to correctly predict patients who do not get pregnant and accuracy. Sensitivity, specificity and accuracy are shown in Eq. 1, 2, 3.

$$\text{Sensitivity:}\frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}} \quad (1)$$

$$\text{Specificity:}\frac{TN}{TN+FP} \quad (2)$$

$$\text{Accuracy:}\frac{TP+TN}{TP+TN+FP+FN} \quad (3)$$

Which TP is the rate ofPeople who has positive pregnancytest and the proposed methodalsoputsthem in thecategory ofpositivepregnancy test.

FPis the rate of people who has positive pregnancytest, but the proposed methodputsthem in thecategory ofnegativepregnancy test.

TNis the rate ofpeople who has negative pregnancytest and the proposed methodalsoputsthem in thecategory ofnegativepregnancy test.

FNis the rate of people who has negative pregnancytest, but the proposed methodputsthem in thecategory ofpositivepregnancy test.

The results of the experiments are shown in table 2, 3, 4.

	Naïve Bayes	Bayes-N	PC	Greedy
Specificity	95	94	97	94
Sensitivity	70	67	58	67
Accuracy	85.05%	83.18%	81.31	83.18%

Table 2. The results of the performance of classifiers using database D1

	Naïve Bayes	Bayes-N	MP-Bayes	Greedy
Specificity	95	94	99	93
Sensitivity	88	88	86	86
Accuracy	92.61%	91.74%	93.91	90%

Table 3. The results of the performance of classifiers using database D2

	Naïve Bayes	Bayes-N	MP-Bayes	Greedy
Specificity	97	96	91	92
Sensitivity	82	73	84	83
Accuracy	92.04%	88.79%	88.79	89.09%

Table 4. The results of the performance of classifiers using database D3

6. Conclusions

The aim of this study is prediction of IVF success rate using Bayesian NetworksIVF using real-world IVF dataset.Accurate and early prediction of the outcome of the IVF is important for both patients and physicians.Bayesian networks provide a visual way to recognize the interactions among variables. In addition, they provide localprobability distributions which can be traced from beginning to end so it is possible to know why and how a BN arrived at a specific conclusion.

As we mentioned in section 3.1, if we use imbalanced training dataset, features will tend to majority class so we use Resampling methods. The results show thatoversampling method gives usbetter result in term of accuracy.

We evaluate the performance of four BN classifiersthe Naïve Bayesian classifier, two constraint-based classifiers (Bayes-N, MP-Bayes), one search and scoring classifier (Greedy), in order to determine their effectiveness to accurately predictsuccess/failure rate ofIVF. Results show that Naïve Bayesclassifier works best among other classifiers using D1 and D3, but MP-Bayes classifier works best among other classifiers using D2.

Results show that we can use Bayesian networks as a new practical approach for prediction the outcome of IVF treatment.

References

Cruz-Ramirez, N., L. Nava-Fernandez, H. G. A. Mesa, E. B. Martinez and J. E. Rojas-Marcial (2005). A parsimonious constraint-based algorithm to induce Bayesian network structures from data. Computer Science, 2005. ENC 2005. Sixth Mexican International Conference on, IEEE.

Fassino, S., A. Piero, S. Boggio, V. Piccioni and L. Garzaro (2002). "Anxiety, depression and anger suppression in infertile couples: a controlled study." Human Reproduction **17**(11): 2986-2994.

Friedman, N. and M. Goldszmidt (1999). Learning Bayesian networks from data, Morgan Kaufmann.

Guh, R. S., T. C. J. Wu and S. P. Weng (2011). "Integrating genetic algorithm and decision tree learning for assistance in predicting in vitro fertilization outcomes." Expert Systems with Applications**38**(4): 4437-4449.

Han, J. and M. Kamber (2006). Data mining: concepts and techniques, Morgan Kaufmann.

Jurisica, I., J. Mylopoulos, J. Glasgow, H. Shapiro and R. F. Casper (1998). "Case-based reasoning in IVF: Prediction and knowledge mining." Artificial Intelligence in Medicine**12**(1): 1-24.

Kaufmann, S., J. Eastaugh, S. Snowden, S. Smye and V. Sharma (1997). "The application of neural networks in predicting the outcome of in-vitro fertilization." Human reproduction **12**(7): 1454-1457.

Martínez-Morales, M., N. Cruz-Ramírez, J. Jiménez-Andrade and R. Garza-Domínguez (2004). "Bayes-N: an algorithm for learning bayesian networks from data using local measures of information gain applied to classification problems." MICAI 2004: Advances in Artificial Intelligence: 527-535.

Mitchell, T. M. (1997). Machine learning. WCB, McGraw-Hill Boston, MA:.

Murphy, K. (2001). "An introduction to graphical models." Rap. tech.

Saith, R., A. Srinivasan, D. Michie and I. Sargent (1998). "Relationships between the developmental potential of human in-vitro fertilization embryos and features describing the embryo, oocyte and follicle." Human Reproduction Update**4**(2): 121-134.

Stassopoulou, A. and M. D. Dikaiakos (2009). "Web robot detection: A probabilistic reasoning approach." Computer Networks**53**(3): 265-278.

Trimarchi, J. R., J. Goodside, L. Passmore, T. Silberstein, L. Hamel and L. Gonzalez (2003). "Comparing data mining and logistic regression for predicting IVF outcome." Fertility and Sterility **80**(3): 100.

Uyar, A., H. N. Ciray, A. Bener and M. Bahceci (2009). "3P: Personalized pregnancy prediction in ivf treatment process." Electronic Healthcare: 58-65.