ASSESSING DECISION TREE MODELS
FOR CLINICAL IN−VITRO FERTILIZATION DATA

Technical Report TR03−296

Dept. of Computer Science and Statistics

University of Rhode Island

LEAH PASSMORE[1], JULIE GOODSIDE[1], LUTZ HAMEL[1], LILIANA GONZALEZ[1], TALI SILBERSTEIN[2] AND JAMES TRIMARCHI[2]

University of Rhode Island[1] and Women and Infants Hospital of Rhode Island[2]

Summary

Less than one third of in−vitro fertilization (IVF) procedures are successful today. One way to increase the success rate of these procedures is to build predictive models which take into account a patient's physiology as well as results derived from different stages of an IVF treatment. These predictive models can then be used to hone in on the course of treatment that will most likely be successful.

Here we study the viability of such models with an added constraint: we are interested in constructing models based on data mining techniques, in particular decision trees. Decision trees are attractive, since they are fairly straightforward to construct and their transparency allows for easy integration into a medical decision support system.

As the foundation of our study we use the database of IVF cycles developed by the IVF Laboratory at the Women and Infants Hospital in Providence, Rhode Island. The data mining algorithm we chose is the C5.0 decision tree inducer. We consider various feature selection methodologies to reduce the features the data mining algorithm has to take into account. Finally, we compare the performance of the resulting decision trees with a hand constructed statistical model.

Key words: C5.0; Feature Selection; Logistic Regression; Principal Component Analysis; In−Vitro Fertilization

[1]Department of Computer Science and Statistics, University of Rhode Island, Kingston, RI, 02881

[2]In Vitro Fertilization Laboratory, Women and Infants Hospital of Rhode Island, Providence, RI, 02905

1. Introduction

In−Vitro Fertilization (IVF) is an assisted reproductive technology (ART) in which one or more eggs are fertilized outside a female's body. This technology has been successfully applied to human reproduction since 1978. In 1998, 9% (2.3 million) of 25.5 million of married couples in the United States were medically infertile (required assistance with reproductive technology) and an additional 19% (4.9 million) had an impaired ability to conceive. That is, a total of 7.2 million couples in the United States could benefit from assisted reproduction technology.

The IVF cycle begins by stimulating the growth of multiple follicles using hormone medication. In a woman's normal cycle, one to two follicles mature every month. Hormone stimulation allows multiple follicles (15 to 20) to mature. The mature eggs are then harvested from the ovaries and placed in a special fluid medium and are fertilized with sperm. The fertilized eggs are cultured in−vitro (in a dish) and the normally developed embryos are then transferred to the woman's uterus. Typically, multiple embryos are transferred to increase the likelihood of pregnancy.

Today the age of the female patient is used as the primary predictor for IVF success. However, age is not a very good predictor since it ignores many of the other factors influencing the success rate of an IVF procedure. Considering that less than a third of IVF procedures are successful, the goal of this study was to assess the viability of decision tree models to predict a patient's success of becoming pregnant using IVF procedures. Decision trees are a non−parametric multivariate method which expresses relationships as simple rules of features given some target concept, in this case becoming pregnant or not. It is hoped that the existence of such predictive models will significantly improve the success rate of IVF procedures by influencing the way treatments for particular patients are chosen. Having a predictive model also helps to set realistic expectations on the sides of both the patient as well as the medical staff. Decision tree models were chosen due to the fact that they are constructed relatively fast, facilitating experimentation. In addition, their transparent nature is very attractive allowing them to be interpreted in a straightforward manner. The possibility of easily converting decision trees into rule sets is another advantage which allows them to be incorporated into medical decision support systems which is the ultimate goal of this programme of investigation.

The dataset considered for this study comes from a database constructed by Women and Infants Hospital in Providence, Rhode Island. The variables from this database considered in this study are either related to characteristics of the patient or derived from the different stages in the IVF cycle. More precisely, the dataset under consideration had 100 independent variables ranging from body mass index of the patient and the amount of hormone treatment the patient received during the IVF treatment to the measurement of ovary volume. The target variable was a categorical variable with two levels indicating whether the IVF cycle was successful (the patient became pregnant) or not (the patient did not become pregnant). The dataset consisted of 678 observations each representing an individual IVF cycle.

In order to get a better assessment of the performance of our decision tree models, a logistic regression model was constructed. This included a variable selection procedure using a statistical approach based on likelihood ratio statistics, to test for associations between the dependent variable, pregnancy, with each of the independent variables. Only those variables having a significant association with pregnancy were kept for further examination.

We chose to construct C5.0 decision tree models (Quinlan, 1993, 2003) predicting the success or failure of IVF cycles. Given the relatively large number of independent variables it was important to include a feature selection phase as part of the modeling in order to reduce the possibility that highly correlated attributes mislead the decision tree algorithm. A number of different feature selection methodologies were tested, including C5.0's native attribute winnowing (Quinlan, 2003) as well as a wrapper approach (Kohavi and John, 1997). We also considered building decision trees on the variables retained by the statistical feature selection process. The performance of the resulting decision trees was compared to the performance of the decision tree built on the entire set of features.

Our conclusion is that decision trees present a reasonable alternative to hand constructed statistical models. Given this and the fact that decision tree models are straightforward to build and as rule sets are easily incorporated into medical decision support systems, we envision a practical system where we periodically refresh the decision trees in order to provide maximum accuracy of the medical support system.

The remainder of the paper is structured as follows. Section 2 describes the data used used here in some detail. In Section 3 we discuss our statistical model building approach. Section 4 describes our data mining methodology. Related work is discussed in Section 5 and we finally close with some remarks and further research in Section 6.

2. Data

The data for this study came from the IVF Laboratory at Women and Infants Hospital in Providence, Rhode Island. The original data set had 678 IVF records and 100 variables. When patients appeared more than once, only the last cycle for a given patient was considered, decreasing the number of records from 678 to 483.

The selected dependent variable was dichotomous (i.e. 1: successful pregnancy, 0: unsuccessful pregnancy) and the 100 independent variables were a mixture of continuous and categorical variables. These independent variables were measures of patient characteristics, as well as variables taken at every stage of the IVF cycle. Patient characteristics include variables such as cycle number, age, body mass index (BMI = weight / height^2), ovaries' volumes and patient diagnosis. The different stages of the IVF cycle include variables related to follicle stimulation, egg harvesting, fertilization, in−vitro culture of embryos and embryo transfer. Of the 100 independent variables, only 53 were retained after removing variables that were functions of others, variables measured after pregnancy occurred, variables pertaining to dates, and variables that had a very large ratio of missing values in the remaining observations (>10%).

3. Statistical Analysis

3.1 Variable Selection

The first step in the variable selection procedure was to test for associations between the dependent variable, pregnancy, with each of the independent variables using likelihood ratio statistics. Only those variables having a significant association with pregnancy were kept for further examination (24 of the 53 variables). Records with missing values in the 24 variables under consideration were discarded further reducing the number of observations to 402 from 483. Of the 402 patients, 203 became pregnant and 199 did not.

With respect to the type of variables, 11 of the 24 independent variables were continuous and 13 categorical. In an attempt to reduce the dimensionality of the continuous variables, a principal component analysis (PCA) was performed and 77.31% of the variation in the data was accounted for by three components. The first component was formed by eight variables of which all were essentially related to egg counts. This component included the following variables: *number of oocytes*, *number of eggs inseminated*, *number of normal eggs fertilized*, *number of fertilized eggs*, *number cleaved* (number of embryos that have divided), *estradiol level* (estradiol is a form of estrogen), *number of follicles that are between 10 and 14 mm* and *number of mature follicles that are 15 mm or larger*. The second component had only *age* as a contributing variable, and *bmi* contributed to the third component. Thus, the original eleven continuous variables were further reduced to just four; *egg count principal component*, *age*, *bmi* and *final fsh dose* (FSH is a hormone that stimulates the growth of follicles).

Further analysis considered only 17 of the 24 variables that showed a significant association with the likelihood of pregnancy of IVF patients. A correlation analysis was performed on the categorical variables and high dependencies were carefully considered when selecting variables to be included in the final model.

3.2 Modeling: Logistic regression

The success and failure of in−vitro fertilization was modeled using multiple logistic regression with a logit link and binomial error distribution (McCullagh and Nelder, 1989). The Akaike's Information Criterion (AIC) (Akaike, 1974) and Schwarz's Bayesian criterion (SBC) (Schwarz 1978) are used to compare the fit of competing models. The AIC is computed as

$$-2\ln(L) + 2k$$

where L is the likelihood function and k is the number of model parameters. The SBC is computed as

$$-2\ln(L) + \ln(n)\, k$$

where n is the number of observations. The model with the smaller information criteria is said to fit the data better.

The deviance ($=-2\ln(L)$), AIC and SBC of the saturated model (all 17 variables included) are 430.30, 526.30 and 1005.96, respectively. The goal was then to find a more parsimonious model that is an adequate substitute to the saturated model. After careful selection of variables and the use of AIC and SBC, the initial number of independent variables (17) was reduced to just six (*age*, *bmi*, *trauma*, *number of embryos transferred*, *embryo transfer technician* and *final fsh dose*). That is, the final model is of the form

$$\text{logit}\,(?) = ? + ?_1\, \textit{age} + ?_2\, \textit{bmi} + ?_3\, \textit{trauma} + ?_4\, \textit{final fsh dose} + \qquad (1)$$
$$\beta_5\, \textit{number transferred} + ?_6\, \textit{technician},$$

where *age* is the age of the patient; *bmi*, the patient's body mass index; *trauma* is a dichotomous variable, 1 indicating that trauma occurred when transferring embryos and 0 that no trauma occurred; *final fsh dose*, the patient's final follicle stimulating hormone dose; *number transferred*, the number of embryos transferred back into the patient's uterus (five classes: 1, 2, 3, 4, 5 or more) and *technician*, the embryo transfer technician (six technicians label from 1 to 6). This reduced model had a deviance of 469.16, an AIC of 497.16 and a SBC of 637.06. The deviance difference of 38.86 with 32 degrees of freedom indicates that there is no difference in fit between saturated and the more parsimonious reduced model (1) but the AIC and SBC indicate that the reduced model is far superior to the saturated model (overfit).

The estimated coefficients and associated confidence intervals of the reduced model (1) are given in table 1. The coefficients associated with *age* (−0.0929), *bmi* (−0.0439) and *final fsh dose* (−0.00305) all are negative, indicating that the probability of pregnancy using assisted reproductive technology decreases when *age*, *bmi* or *final fsh dose* increases. The estimated coefficient for *trauma* (−0.9766) indicates that the likelihood of pregnancy decreases when trauma occurs in the transfer of embryos. The coefficients associated with the variable *number of transfers* are four since this is a categorical variable with five classes. The coefficients associated each level of the variable indicate the estimated coefficient difference between the respective classes (1, 2, 3 or 4 embryos transferred) and the last class (5 or more embryos transferred). The same applies to the coefficients associated with *technicians*. The estimated values indicate the difference between technicians labeled 1, 2, 3, 4 or 5 with the last technician label 6 respectively.

Parameter	Estimate	95% Confidence Limits	
Intercept	5.1441	2.5548	7.7334
age	−0.0929	−0.1585	−0.0273
bmi	−0.0439	−0.0767	−0.0111
trauma	−0.9766	−1.6100	−0.3433
final fsh dose	−0.0031	−0.0048	−0.0014
number transferred (1)	−1.8758	−2.9156	−0.8360
number transferred (2)	0.1824	−0.3219	0.6866
number transferred (3)	0.6094	0.1419	1.0769
number transferred (4)	0.6497	0.0124	1.2870
technician (1)	0.0075	−0.4921	0.5071
technician (2)	0.0258	−0.9076	0.9591
technician (3)	0.7010	0.2585	1.1436
technician (4)	−0.0419	−0.5028	0.4191
technician (5)	−0.5785	−1.1212	−0.0357

Table 1. Wald Confidence Interval for Model Parameters.

A type 3 analysis (table 2) indicates that when all other variables are included in the logit model (1), the variable that contributes the most is *final fsh dose*, followed by *trauma*, *number of transfers, age*, *bmi* and *embryo transfer technician*.

Effect	DF	Chi−Square	Pr > ChiSq
age	1	7.6970	0.0055
bmi	1	6.8875	0.0087
trauma	1	9.1345	0.0025
final fsh dose	1	12.3199	0.0004
number transferred	4	15.3759	0.0040
technician	5	14.6124	0.0122

Table 2. Type 3 Analysis of Effects.

In order to assess the predictive ability of the models, the percent of concordant results were calculated. That is, all the possible pairs of pregnant/non−pregnant outcomes were formed (40,397=199x203) and the probability of pregnancy was calculated using the model. The number of times that the probability of a pregnant (success) patient was higher than the probability of the non−pregnant (failure) patient divided by the total number of pairs (40,397) is the percent of concordant results. The saturated and reduced models were 79.7% and 75.5% concordant, respectively, and their effective R^2 (R^2/R^2_{max}) was 75% for both models. Table 3 summarizes these results.

Model	*Variables Considered*	*Percent Concordance*
Saturated Model	17	79.7%
Reduced Model	6	75.5%

Table 3. Summary of Statistical Results.

Although these models are not a classification model but compute the probability of success the above analysis provides a good insight into the predictiveness of the IVF dataset and the concordance results give us an idea of what to expect in terms of accuracy for our classification models based on decision trees.

4. Data Mining

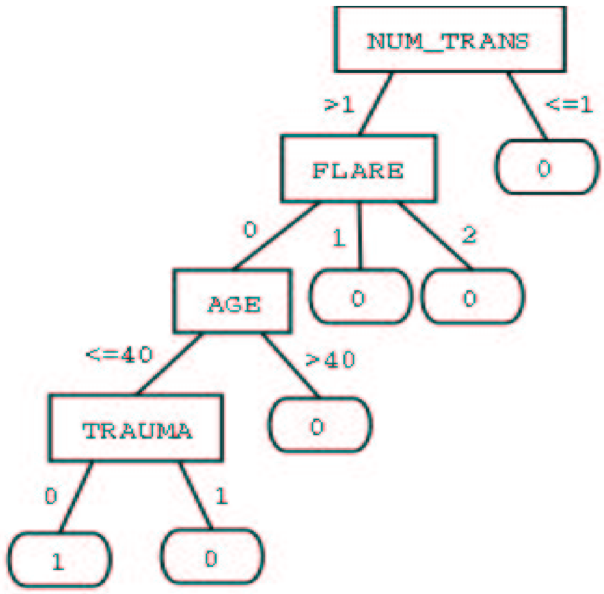
The data mining algorithm we used in this study is Quinlan's C5.0 decision tree inducer (Quinlan, 1993, 2003). Decision trees have a tendency to overfit the training data. In the case of the C5.0 decision tree algorithm the pruning confidence interval and the minimum support in each node allow the user to control this overfitting behavior. In our approach a 10−fold cross validation (Kohavi, 1995) on the dataset was used to estimate the best setting for the pruning and support parameters. Once the optimal values were found, we induced a tree on the full dataset and evaluated its performance against the same dataset. The results computed in this way were immediately comparable to the concordance results in our statistical approach. It is noteworthy that the performance of the decision trees built on the full dataset using the optimal pruning and support parameters did not significantly differ from the average performance of the cross validated trees. For all of the following results it turns out that the best pruning confidence interval was 11% and the best performing trees had a minimum support of 11 records per leaf. It should be mentioned that of the 402 records considered for the statistical modeling 77 still had missing values in various features not included in the statistical model. Since in this study we did not consider replacing missing values, we discarded these 77 records and performed our tree induction on the remaining 325 records.

4.1 Full Featured Data Set

In order to obtain a baseline performance appraisal of the C5.0 algorithm we applied it to the full featured data set with its 53 constituent independent variables. That is, we did not consider any feature selection at all. The resulting decision tree looks as follows:

```
NUM_TRANS <= 1: 0 (22/3)
NUM_TRANS > 1:
:...FLARE = 1: 0 (11/4)
    FLARE = 2: 0 (47/14)
    FLARE = 0:
    :...AGE > 40: 0 (14/2)
        AGE <= 40:
        :...TRAUMA = 0: 1 (193/66)
            TRAUMA = 1: 0 (38/17)
```

A graphic representation of this decision tree is as follows:



This decision tree indicates that successful patients have more than one embryo transferred (NUM_TRANS > 1), have no additional hormonal stimulation (FLARE = 0), are of age less or equal to forty (AGE <= 40), and have no trauma during implantation of the embryos (TRAUMA = 0). The numbers appearing at the leaves of the text based tree representation indicate the support in terms of observations at the leaves. For instance, the first leaf in the tree, NUM_TRANS <= 1: 0, has a support of a total of 25 records of which 22 support the not pregnant classification (0) and 3 do not. The support for a successful IVF treatment appears at the second to last leaf, TRAUMA = 0: 1. This leaf has a total support of 259 observations of which 193 support the classification at this node and 66 do not. The accuracy of this model is 67.4%. Here, accuracy is defined as the number of correct classifications divided by the total number of observations in the dataset. It is interesting to note that there is significant overlap between the variables this decision tree considers and the variables resulting from our statistical analysis.

4.2 C5.0 Attribute Winnowing

Next we considered C5.0's built−in attribute selection utility called *Attribute Winnowing.* Here, the C5.0 algorithm searches through the feature space before model building commences and discards any variables that are considered detrimental to the predictive properties of the model or that have significant overlap with other variables (Quinlan, 2003). Here, attribute winnowing discards virtually all variables except for five: FLARE (hormonal stimulation treatment), TRAUMA (implantation trauma), CLEAVED (in−vitro embryo division) , INSEM_TIME (insemination time of day), R_OVARY_VOL (right ovary volume). The resulting decision tree looks as follows:

```
FLARE = 1: 0 (14/4)
FLARE = 2: 0 (54/14)
FLARE = 0:
:...CLEAVED <= 3: 0 (50/18)
    CLEAVED > 3:
    :...TRAUMA = 1: 0 (34/14)
        TRAUMA = 0:
        :...R_OVARY_VOL <= 52279.24: 1 (154/43)
            R_OVARY_VOL > 52279.24: 0 (19/6)
```

This tree describes a successful patient as having no additional stimulation treatment (FLARE = 0), the number of embryos that divided in−vitro is larger than three (CLEAVED > 3), experienced no trauma during implantation (TRAUMA = 0), and has a right ovary volume of less or equal to 52279.24 mm^3 (R_OVARY_VOL <= 52279.24). The total support for the positive classification is 197 records with 154 supporting the classification. The accuracy of this decision tree is 69.5%.

Although this model is more accurate than the previous model there is some concern of possible overfitting due to attribute selection (Kohavi and John, 1997) due to the inclusion of the right ovary volume and the exclusion of patient age as model variables. It is conceivable that the ovary volume is highly correlated with age but is a better predictor than age. Therefore, the attribute selection did not include age in the model. Further analysis needs to be conducted on this model. One way to check for overfitting is to see if there is a correlation between some combined metric between left and right ovary volumes and the dependent variable and the decision tree algorithm continues to respond to this combined ovary volume metric.

4.3 Attribute Selection with a Wrapper

Next we considered attribute selection via a wrapper (Kohavi and John, 1997). The wrapper search algorithm considered was a best first greedy algorithm with backtracking (Witten and Frank, 2000). This approach returned only a single attribute as significant, namely: FINAL_FSH_DOSE. FSH is a hormone that stimulates the growth of follicles and this attribute captures the final dose of FSH units a patient received. The resulting tree is a trivial one:

```
FINAL_FSH_DOSE <= 300: 1 (195/74)
FINAL_FSH_DOSE > 300: 0 (130/46)
```

The model is simple in the sense that it states that the less stimulation a patient receives the higher the probability of success. The accuracy of this model is 63.1%.

It is interesting to note that this is also the attribute that the statistical feature selection indicated as the most predictive feature. It is surprising that the attribute selection method did not include any of the other highly predictive attributes such as TRAUMA.

4.4 Statistical Feature Set

Finally, we considered the dataset constructed via the statistical feature selection. The main difference being that this dataset included 402 observations vs. 325 observations considered in the above experiments. The resulting tree is as follows:

```
NUM_TRANS <= 1: 0 (28/1)
NUM_TRANS > 1:
:...AGE > 40: 0 (32/6)
    AGE <= 40:
    :...FINAL_FSH_DOSE <= 300: 1 (227/80)
        FINAL_FSH_DOSE > 300: 0 (113/47)
```

It is interesting to note that the decision tree algorithm was not able to completely replicate the statistical model which is probably due to the constraints on the way decision surfaces can be constructed in the data space by the tree inducer. However, the resulting model resembles the statistical model very closely and has an accuracy of 66.5%.

4.5 Data Mining Summary and Discussion

We found that the data mining approach with the decision tree inducer C5.0 works reasonably well, where the best models lag the hand constructed reduced model by only 6−8%. The most intuitive model with an accuracy of 67.4% was perhaps the decision tree induced on the data set with its full feature set in tact. This model has significant overlap with the statistical model. The best performing model was produced by using C5.0 attribute winnowing algorithm before actually building the model. Attribute winnowing reduced the feature set from 53 to 5 features and the performance of the resulting model was 69.5%. However, some concerns in overfitting remain. What is interesting to note is that feature selection did not have a large impact on the performance. Perhaps even more surprising was the fact that the wrapper and statistical feature selection had actually a negative effect on the performance. Table 4 summarizes the performance of the various decision tree models we have considered.

Model	*Variables Considered*	*Accuracy*
C5.0 (Full Feature Set)	53	67.4%
C5.0 (Winnowing)	5	69.5%
C5.0 (Wrapper)	1	63.1%
C5.0 (Statistical Feature set)	6	66.5%

Table 4. Summary of Decision Tree Model Performance.

It can be shown that the female patient's age as a predictor for IVF success has an accuracy of 54%. Therefore, not only do the models obtained from decision tree induction come close to the accuracy of the hand constructed statistical models but they also represent a significant jump in terms of predictive power in today's IVF setting.

5. Related Work

Data mining techniques, in particular decision trees, and the construction of predictive statistical models have been used before in in−vitro fertilization. The paper by Saith *et al* (Saith *et al*, 1998) uses decision trees to investigate the relationship of the features of the embryo, oocyte and follicle to the successful outcome of the embryo transfer. Although 53 features were studied, only 4 had predictive capabilities, embryo grade, cell number, follicle size and follicular fluid volume. This study used 200 IVF records and significantly differs from our study in that it did not consider any clinical data on the female and male patients involved in the procedure.

A paper by Shen *et al* (Shen S. *et al*, 2003) uses statisitical analysis to examine factors involved in IVF procedures. This study, however, only considered fertilizations accomplished with intracytoplasmic sperm injection. Statistical approaches were used to find that sperm motility and ICSI operator were the two most important predictors for the success of an IVF procedure. Sperm motility and ICSI technician were also features considered in our study. Our data set was drastically different because the ICSI method of fertilization was used in only 44% of our records. We did not find these factors predictive in our sample population.

It should also be noted that these studies did not share our goal of ultimately constructing a medical decision support system allowing the attending physician access to the likely outcome of a procedure given certain parameters.

Jurisica and Nixon share our vision of a medical decision support system (Jurisica and Nixon, 1998). However, their approach is very different from ours. Where we are interested in constructing a predictive model over the available IVF they use case−based reasoning in their IVF decision support system.

6. Conclusions and Further Research

Less than one third of IVF procedures are successful today. One way to increase the success rate of these procedures is to build predictive models which take into account a patient's physiology as well as results derived from different stages of an IVF treatment. These predictive models can then be used to develop a course of treatment that will most likely be successful for a particular patient.

Here we studied the viability of such models with an added constraint: we were interested in constructing models based on data mining techniques, in particular decision trees. Decision trees as models are attractive, since they are fairly straightforward to construct and their transparency allows for easy integration into a medical decision support system. The construction of such a medical decision support system is the ultimate goal of our course of investigation.

As the foundation of our study we used the database of IVF cycles developed by the IVF Laboratory at the Women and Infants Hospital in Providence, Rhode Island. The data mining algorithm we chose was the C5.0 decision tree inducer. We used feature selection to reduce the number of features considered by the decision tree inducer. However, it was interesting to observe that this did not have a large impact on the model performance. In fact, the most intuitive and better performing decision tree model was constructed without any feature selection at all. In order to assess the quality of our data mining models we hand constructed a logistic regression model. We found that our decision tree models performed reasonably well lagging between 6−8% in accuracy behind the hand constructed statistical model which had an accuracy of 75.4%. It is also important to note that our decision tree models represent a significant jump in predictive power over solely considering the female patient's age as the predictor of IVF success; a standard practice in today's IVF procedures.

Given that the decision tree inducer cannot replicate the statistical model even when given the identical dataset we are considering other inductive machine learning schemes that allow for more flexible decision surfaces to be constructed in the data space. Currently we are investigating artificial neural networks (Mitchell, 1997) and support vector machines (Kecman, 2001) in order to construct predictive models of the IVF data. Unfortunately we will be sacrificing the transparency of decision trees for more accurate models of the data.

References

Akaike, H. (1974), A New Look at Statistical Model Identification, IEEE Transactions on Automatic Control, 19, 716 −723.

Jurisica, I. and Nixon, B. (1998), Building quality into case−based reasoning systems. Lecture Notes in Computer Science, CAiSE*98, pages 363−380, Springer−Verlag.

Kecman, V. (2001), Learning and Soft Computing: Support Vector Machines, Neural Networks, and Fuzzy Logic Models, MIT Press.

Kohavi, R. (1995). A study of cross−validation and bootstrap for accuracy estimation and model selection. In Proceedings of the 15th International Joint Conference on Artificial Intelligence, pp. 1137−1143.

Kohavi, R., John, G.H. (1997),Wrappers for feature subset selection, Artificial Intelligence 97(1−2), 273−−324.

Mitchell, T. (1997), Machine Learning, McGraw−Hill.

McCullagh, P., Nelder, J.A. (1989). Generalized Linear Models, 2nd edition, Chapman and Hall, Monographs on Statistics and Applied Probability, London.

Quinlan, J.R. (1993), C4.5: Programs for Machine Learning, Morgan Kaufmann.

Quinlan, J.R. (2003). C5.0 Online Tutorial, http://www.rulequest.com.

Saith, R., Srinivasan, A., Michie, D., Sargent, I. (1998). Relationships between the developmental potential of human in−vitro fertilization embryos and features describing the embryo, oocyte and follicle, Human Reproduction Update, Volume 4, Issue 2, March/April.

Schwarz, G. (1978), Estimating the Dimension of a Model, Annals of Statistics, 6, 461 −464.

Shen S., Khabani A., Klein N., Battaglia, D. (2003). Statistical analysis of factors affecting fertilization rates and clinical outcome associated with intracytoplasmic sperm injection, American Society for Reproductive Medicine, Vol. 79, No.2, February.

Wtten, I., Frank, E. (2000). Data Mining: Practical Machine Learning Tools and Techniques in Java, Morgan Kaufmann.