Pak. J. Biotechnol. Vol. 15 (Special Issue 1) Pp. 36-39 (2018) Durairaj and Ramasamy
www.pjbt.org, PISSN: 1812-1837, EISSN: 2312-7791

FEATURE DIMINUTION BY HYBRID ALGORITHM FOR IMPROVING THE SUCCESS RATE FOR IVF TREATMENT

Durairaj M[1] and N. Ramasamy[2]

School of Computer Science, Engineering and Applications, Bharathidasan University, Trichy,Tamilnadu, India.
[1]durairaj.csbdu@gmail.com; [2]nandhakumar.897@gmail.com

ABSTRACT

Infertility is the most common problem faced by today's generation. The factors like environment, genetic or personal characteristics are responsible for these problems. Different infertility treatments like IVF, IUI etc are used to treat those infertile people. But the cost and emotions beyond each and every cycle of IVF treatment is very high and also the success rate differs from person to person. So, there is a need to find a system which would predict the outcome of IVF to motivate the people both in psychologically and financially. Many Data Mining techniques are applied to predict the outcome of the IVF treatment. Reducing the unwanted features which affects the quality of result is one of the significant tasks in Data Mining. This paper proposes a hybrid algorithm which combines the core features of Ant Colony Optimization Algorithm and Relative Reduct Theory for Feature Reduction. In this work, the proposed Algorithm is compared with the existing related algorithms. It is evident from the results that the proposed algorithm achieved its target of reducing the features to minimum numbers without compromising the core knowledge of the system to estimate the success rate.

INDEX TERMS:ANT COLONY, RELATIVE REDUCT ALGORITHM, SUCCESS RATE, FEATURE REDUCTION, ACCURACY.

I. INTRODUCTION

Infertility is distressing a growing number of married couples around the planet. Assisted Reproductive Technologies (ART) and In Vitro Fertilization (IVF) are the effective ways to address the problem of infertility. In IVF procedure, several eggs are collected from the woman's ovaries and fertilized with the donor's sperm to produce embryos. The best one among the embryos will be transferred to the woman's uterus, where as the biggest problem arises here since the best embryo is chosen based on recorded features, characterizing the morphology, oocytes, follicle and the sperm sample. The success rate achieved by this treatment has been increased recently up to 10%. But it still fits only to the 40% of the people (Milewski et al., 2013). It is a complicated task for an embryologist to analyze and correlate the features, since it involves several features, which are sufficiently large. Even though there are processes customized for each case to improve the success rate in fertility and there is still some technology lagging behind to achieve it. Hence in such a case, there arises a compelling need to appeal for more advanced methods like Data Mining and Artificial Intelligence. This field is falling sort of an automated tool which can intelligently analyze the stage of IVF treatment, patient's demography and other parameters. The practitioners at IVF centres have begun to feel the absence of an expert system as a data processing tool to help them. Feature Reduction is a significant task in Data Mining since it removes the irrelevant or redundant features without loss of much information. Furthermore, a Feature Reduction technique will reduce the amount of time taken for number of tests and the predicted success rate of the treatment can be used for the patients to psychologically strengthen them knowing that their success rate is going to be positive. To achieve this objective, this paper develops a hybrid model which will reduce the maximum features to minimum number to increase the accuracy of the success rate estimation in the held at IVF clinics.

The paper is organized as follows: Section II discusses some of the existing work carried out in predicting the success rate of IVF treatment. Section III briefs the data set used for the experimentation. Section IV describes the proposed Algorithm with its framework. Section V discusses the results obtained and the paper is concluded in Section VI.

II. LITERATURE REVIEW

Kaufmann et al., (1997) applied Neural Network to predict the outcome of the IVF treatment. A total of 8 different types of Neural Network is applied on the same dataset. A sensitivity of 0.55%, Specificity 0.68% and accuracy of 59% was obtained by applying Neural Network. Al-Shawi (2017) applied Naïve Bayes Classifier to the dataset to classify the embryos. In order to overcome the problem arising due to the imbalanced dataset, the author analyzed the effects of oversampling, under sampling and change in threshold. The value of 0.3 is found the perfect threshold value for the correct classification of embryos. True Positive Rate 64.4% and False Positive Rate 30.6% are obtained.

Asli et al., (2010) evaluated six dissimilar methods like Naïve Bayes, K- Nearest Neighbor, Decision Tree, Support Vector Machine, Multi layer Perceptron and Radial Basis Function Network for envisaging the results for embryo implantation. Among all the six methods, Naïve Bayes and Radial Basis Function Network were observed to function better. Durairaj and Meena (2011) implemented a hybrid system for predicting the infertility treatment based on Rough Set and Artificial Neural Network. Rough Set Theory was used to find the reduct set. It was used as a pre-processing tool to reduce the number of variables which are used as input for the Neural Network. This system worked in a better way for large and medium size of medical data.

Uzair Ahmed (2017) compared three techniques, Decision Tree, Multi Layer Perceptron and Support Vector Machine to evaluate the male partner. Clinicians obtained the data from semen analysis and compared it with the corresponding reference value. An accuracy of 86% was obtained from Multilayer Perceptron Network and Support Vector Machine. Durairaj and Nandhakumar (2013) illustrates the process of app-

lying data mining techniques for identifying influential tests for infertility couples to determine the success rate of IVF treatment. The data set are pre-processed to select only most influential parameters using attribute selection algorithm, which filters the noisy data and selected only the para-meters with high impact factors. The experimental results show that the filter and classifier tool using data mining techniques employed to evaluate and produce the minimum set of data which have most influence on estimating the success rate of IVF treatment. Artificial Neural Network was used for predicting the fertility success rate based on the IVF data. An accuracy of 73% was obtained (Durairaj and Thamilselvan 2013). Claudio Manna et al., (2013) applied Artificial Intelligence for classifying the embryo and oocytes. An integrated method based on Artificial Neural Network and Rough Set. Theory was adopted for analyzing the IVF data (Durairaj and Nandhakumar 2014). The Rosetta tool is used for analyzing the data. An accuracy of 90% was obtained by using the integrated method (Durairaj and Ramasamy 2015 & 2016).

III. DATA SET

The data set used for the experimentation is collected from various Fertility clinics, Hospitals and Research centres in Tamil Nadu. This data set has 42 attributes. Among all the 42 attributes, 34 attributes are taken for the experiments based on the doctor's suggestion.

TABLE I. Attributes used for this work

Attributes used for this work				
Name	Previous Surgery	Endometriosis	Liquefaction Time	Male Factor Only
Unknown Factor	Pre-Existing Symptoms Of Depression	Tubal Infertility	Sperm Concentration	Severe Male Factor
Place	Fear And Negative Treatment Attitude	Ovulatory Factor	Sperm Motility	Female Factor Only
IVF Treatment	Psychological And Emotional Factors	Hormonal Factor	Sperm Vitality	Combined Factor
Miscarriage	Difficulty In Tolerating Negative Emotions For Extended Time	Cervical Factor	Sperm Morphology	Unknown Factor
Miscarriage Causes	Uncertainty	Unexplained Factor	No.of Oocytes Retrieved	Place
Medical Disorders	Strain Of Repeated Treatment	Semen Ejaculate Volume	No.of Embryos Transferred	IVF Treatment

The list of attributes given in Table 2 is taken for reduction process based on doctor's suggestion.

TABLE II. List of Attributes chosen for experimentation

Age
Endometriosis
Ovulatory Factor
Hormonal Factor
Crevical Factor
Unexplained Factor
Semen Ejaculate Volume
Liquefaction Time
Sperm Concentration
Sperm motility
Sperm vitality
Sperm morphology
No. of oocytes retrieved
No. of embryos transferred
Male factor only
Severe male factor
Female factor only
Combined factor
IVF Treatment

Ant Colonized Relative Reduct Algorithm (ACRRA) At the initial stage the Pheromone and R values are initialized. The ants are created using the attributes. A solution is constructed for each Ant. A Feature Subset with Conditional Features C is selected. Then the Conditional Features are stored in R. After storing the Conditional Features, the dependency of each attribute is checked. If the dependency is equal to one, that attribute is eliminated. The remaining features are taken and stored in R. If the condition is not satisfied, then the process is repeated until the reduct set is obtained. Instead of indescernabiliity matrix in Rough Set Theory, this dependency measure is taken as a new technique. After obtaining the Reduct data set and stored in R, the best ant table is updated. The Final termination criteria are checked. If the termination criteria are reached, the optimal data is stored, and the process is stopped. Else, the pheromone level is updated and the process starts from initialization.

Algorithm: **Ant Colonized Relative Reduct Algorithm** ACRRA(C-> Conditional Features, D-> Decision Features)

Input: Data set

Algorithm:

Step 1: *Initialize Phermone, R*
Step 2: *Create Ants*
Step 3: *Construct a Solution for each Ant*
Step 4: *Select Feature Subset*
Step 5: $R \leftarrow C$
Step 6: $\forall_a \ \varepsilon C$
Step 7: *If* $K_{R\{a\}}(D) == 1$
$R \leftarrow R\{a\}$
Go to Step 8
Else
Go to Step 5
Step 8: *Update the best Ant Value*
Step 9: *Check for Termination Criteria*
If yes go to Step 11
Else go to step 10
Step 10: *Update the Pheromone level*
Go to step 3
Step 11: *End*
Output: Optimal Reduct Data set == R

Figure 1: Depicts the framework of the proposed ACRR Algorithm

IV. RESULTS AND DISCUSSION

The Existing Algorithms like Genetic Algorithm (GA), Ant Colony Optimization (ACO) Algorithm, Particle Swarm Optimization (PSO) Algorithm, Relative Reduct (RR) Algorithm and Quick Reduct (QR) Algorithm are taken for study with the proposed Ant Colony Relative Reduct (ACRR) Algorithm.

TABLE III: Comparison of accuracy of different algorithm with different classifiers

Original Attributes				Attributes obtained by using GA				Attributes obtained by using the PSO Algorithm			
NB	MLPN	RBF	J48	NB	MLPN	RBF	J48	NB	MLPN	RBF	J48
72.81	95.61	77.19	73.68	67.54	91.22	71.05	79.83	66.67	83.33	68.42	69.30

TABLE IV. COMPARISON OF ACCURACY OF DIFFERENT ALGORITHM WITH DIFFEENT CLAS

Attributes obtained by using the ACO Algorithm				Attributes obtained by using RR Algorithm				Attributes obtained by using QR Algorithm			
NB	MLPN	RBF	J48	NB	MLPN	RBF	J48	NB	MLPN	RBF	J48
66.67	92.98	74.56	73.68	69.30	87.72	75.44	73.68	64.04	85.09	72.81	69.30

A total of 18 Features is selected from 41 Features after applying ACRRA. The selected features are tested with the respective algorithms. The reduced Features are listed in Table III.

TABLE V: list of Attributes reduced by existing and proposed Algorithm

Original Attributes	GA	ACO	PSO	RR	QR	ACRR
Age	✓	✓	✓		✓	✓
Endometriosis	✓	✓	✓	✓	✓	✓
Ovulatory Factor	✓	✓	✓	✓	✓	✓
Hormonal Factor	✓	✓	✓	✓	✓	
Crevical Factor	✓	✓	✓	✓	✓	✓
Unexplained Factor	✓	✓	✓	✓	✓	✓
Semen Ejaculate Volume	✓	✓	✓	✓	✓	
Liquefaction Time	✓				✓	
Sperm Concentration	✓	✓	✓	✓	✓	✓
Sperm motility	✓	✓	✓	✓	✓	✓
Sperm vitality		✓	✓	✓	✓	
Sperm morphology	✓	✓		✓	✓	
No. of oocytes retrieved						✓
No. of embryos transferred						✓
Male factor only	✓	✓		✓		✓
Severe male factor						
Female factor only	✓	✓		✓		
Combined factor						
IVF Treatment	✓	✓	✓	✓	✓	✓

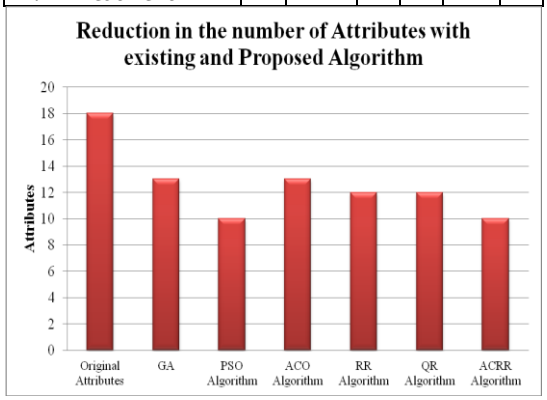

Figure 2: Reduction in the number of Attributes

Figure 2 depicts the number of features obtained by using the existing and proposed algorithms.

TABLE VI. Comparison of accuracy of different algorithm with diffeent classifiers

	Attributes obtained by using Proposed ACRR Algorithm			
	NB	MLPN	RBF	J48
Accuracy	75.44	90.35	78.07	73.68

The accuracy obtained while classifying the attributes obtained using every classifier is listed in Table 4, 5 and Table 6 respectively

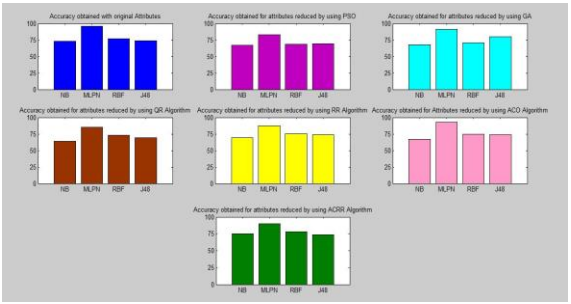

Figure 3: Comparison of Accuracy for different Classifiers by using attributes obtained by Existing and Proposed Algorithm

Figure 3 depicts the comparison of Accuracy for different classifiers by using the attributes obtained by using the Existing and Proposed Algorithm for Feature Reduction.

V. CONCLUSION

Feature Reduction is an important task in Data Mining. Reducing the unwanted features will improve the classification accuracy and minimizes the time taken for classification too. The existing algorithms GA, PSO, ACO, QR and RR are taken for study and a new Algorithm ACRRA is proposed. ACO Algorithm perfomred better than the PSO and GA. The performance of RR Algorithm gives better performance than the QR Algorithm. The proposed ACRRA algorithm is developed by combining the core features of ACO Algorithm and RR Algorithm, which performs better than the existing Algorithms. The proposed ACRR Algorithm improves the accuracy and also performs better with other metrics, which is evident from the results obtained.

REFERENCES

Al-Shawi S.G., Z.K. Al-Younis and N.F.A. Al-Kareem. Study of cumin antibacterial and antioxidant activity

of alcoholic and aqueous extracts. Pak. J. Biotechnol. 14(2): 227 -231 (2017)

Asli U., A. Bener, H.N. Ciray and M. Bahceci, ROC Based Evaluation and Comparison of Classifiers for IVF Implantation Prediction, Institute of Computer Sci-ences, Social Informatics and Telecommunication Eng-ineering, LNICST 27, Pp. 108-111 (2010).

Claudio M., L. Nanni, A. Lumini, S. Pappalardo, Artificial intelligence Techniques for Embryo and Oocyte Classification, Elsevier, Reproductive BioMedicine (2013).

Durairaj M. and P. Thamilselvan, Applications of Artificial NeuralNetwork for IVF Data Analysis and Prediction. Journal of Engineering, Computers and Applied Sciences 2(9): (2013).

Durairaj, M. and K. Meena, A Hybrid Prediction System Using Rough Setsand Artificial Neural Networks. International Journal of Innovative Technology and Creative Engineering 1(7): (2011).

Durairaj, M. and N. Ramasamy, A Comparison of the Perceptive Approaches for Preprocessing the Data Set for Predicting Fertility Success Rate. International Journal of Control theory and Applications 9 (27): (2016)

Durairaj, M. and N. Ramasamy, Intelligent Prediction Methods and Techniques Using Disease Diagnosis in Medical Database: A Review. International Journal of Control theory and Applications 8(5): (2015).

Durairaj, M. and R. Nandhakumar, An Integrated Methodology of Artificial Neural Network and Rough Set Theory for Analyzing IVF Data, IEEE International Conference on Intelligent Computing Application (2014)

Durairaj, M. and R. Nandhakumar, Data Mining Application on IVF Data For The Selection of Influential Parameters on Fertility. IJEAT 2(6): (2013).

Kaufmann, S.J., J.L. Eastaugh, S. Snowden, S.W. Smye and V. Sharma, The application of neural net-work in predicting the outcome of in-vitrofertilization. Human Reproduction 2(7): 1454-1457 (1997).

Milewski, R., Milewska, A. J., Czerniecki, J., Leśniewska, M., and S. Wołczyński, Analysis of the demographic profile of patients treated for infertility using assisted reproductive techniques in 2005–2010. Ginekologia Polska 84(7): 609–614 (2013).

Uzair Ahmad, Evaluating seed rate, cutting and nitrogen level study of yield and yield components of triticale. Pak. J. Biotechnol. 14(2): 193 - 204 (2017).