A Survey on the Machine Learning Techniques used in IVF Treatment to Improve the Success Rate

Ms. Babitha
Assistant Professor, Dept.of CSE
St.Joseph Engineering College
Mangalore, Karnataka, India

***Abstract*— In vitro fertilization (IVF) is one of the widely used assisted reproductive technologies which help the couple suffering from infertility related issues to have a child. During IVF, an egg is removed from the woman's ovaries and fertilised with sperm in a laboratory. The fertilised egg, called an embryo, is then returned to the woman's womb to grow and develop. Since this treatment involves an application of several hormones and medicines to both female and male patients, it is very complicated and stressful process. Even after undergoing this costly treatment, there is no guarantee that the couple will get the positive result. There are many cases in which the IVF cycle fails and thereafter people will lose their hope of having a child. So, it is essential to have some method in which it is possible to predict the possibilities of having success in the treatment. The best possible management of the in-vitro fertilization treatment and patient advice is crucial for both patients and medical practitioners. The ultimate aim of infertile couple is the success of IVF treatment which depends on a number of influencing attributes. Without the automated tools, it is difficult for the doctors to assess any influencing trend of the attributes and factors which can lead successful pregnancy. There were many studies conducted in this area in which different machine learning classification techniques such as artificial neural networks (ANN), support vector machines (SVM), Decision trees, naive bayes, K-nearest neighbour (KNN), multi layer perceptrons(MLP) , Random Forest were used. Some of the work focuses on helping the doctors to understand the trends and patters to help them in suggesting the patient to go for IVF treatment or not. Some of them concentrate on helping the embryologist to choose the right embryo which will result in successful pregnancy. Some authors also worked towards optimizing the algorithm by selecting the correct number of features.**

Keywords—Artificial neural networks, Support vector machines, decision trees, Feature selection, Classification, Naïve Bayes classifier

I. INTRODUCTION

Infertility is the inability of a person, animal or plant to reproduce by natural means. According to world health organization (WHO), one in every four couples in developing countries had been found to be affected by infertility. Times of India reported that there are 27.5 million couples in India suffering from infertility. There are many medical factors and some unknown reasons stopping many of these couples to enjoy the bliss of having their family. In the past, there was no solution for this problem. But, thanks to the advancement in medical science and technology, today there are many options that can help them fulfill their dream. One such popular method is IVF. IVF stands for In Vitro Fertilization and it is one of the types of assisted Reproductive technology (ART) being practiced in every corner of the globe that has helped many childless people. IVF has undergone various modifications from its birth in 1978.

Infertility can occur in either of the couple and in some cases, this problem can arise from both. IVF is not the first option for treating the infertility. Fertility drugs, artificial insemination treatments like IUI and surgery are some of the options available. When these methods don't work, IVF will be recommended.

IVF is a complex process in which initially a female partner is given a course of drugs that can stimulate 12 to 15 mature eggs or oocytes. An ultra sound and blood tests are conducted to check if the ovarian stimulation is taking place properly or not. A hormone injection is given two days prior to egg collection which helps in maturation of eggs. These eggs are collected from women uterus and the sperm of male partner is also collected. The sperm and eggs are put together in the lab by the embryologist and they are fertilized in the lab. Two to five days post, the doctor places one or two embryos in the uterus. If there are additional embryos, these can be used later for the next IVF cycle when the first cycle fails. Around two weeks later, blood test is conducted to know if the test result is positive or negative.

Though this treatment is being used from 1978, there are many risks involved in the treatment. Some of them are multiple births, premature delivery, higher stress level etc. Main thing is there is no guarantee of success in this treatment even after undergoing this complex and expensive process which will really cause emotional and economical burden on infertile couple. Since there are almost 50 to 60 female and male factors which decides the success of this treatment, doctors find it difficult to suggest the couple to go for the treatment or not. There are some machine learning algorithms in which a model is built which helps the doctors to suggest the couple. Another important factor in which IVF success rate depends is the embryo selection. Most of the times, doctors fail to select the most viable embryos to transfer into the uterus of the women it will lead to failure. There are some machine learning models which help the doctors to select the most viable embryos. In some cases, doctors select more than one embryo to plant into uterus in

order to increase the success rate, but that will lead to multiple pregnancies.

In this survey, various machine learning techniques used at various phases of IVF treatment to boost the success rate of the treatment are studied along with its performance. These techniques are described in the next section

II. ROLE OF MACHINE LEARNING TECHNIQUES

Machine learning is the scientific study of algorithms and statistical models that computer systems use to effectively perform a specific task without using explicit instructions, relying on patterns and inference instead. It is a subset of artificial intelligence. Artificial intelligence (AI), machine learning and deep learning are taking the health care industry by storm. They are the practical tools that can help the companies optimize their service provision, improve the standard of care, generate more revenue and decrease risk. Almost all the companies in health care space have already begun to use the technology in practice.

The machine techniques used in the various works discussed in this paper can be classified into three types based on the purpose

1. Techniques which builds model to help doctors to suggest the patients whether to go for the IVF treatment or not.
2. Models which predicts the success of IVF result helps the infertile patients to know what is their chance of having a child when they undergo IVF treatment
3. Models which helps the embryologist to choose the top quality embryo to plant into women uterus which will increase the chance of having pregnancy and also tries to decrease the risk of multiple pregnancies.

.

III. MACHINE LEARNIG TECHNIQUES

Machine learning algorithms can be broadly classified into 3 types as shown in figure 1.

- **Supervised learning:** Algorithm builds a mathematical model from a set of data that contains both the inputs and desired outputs.
- **Un-supervised learning**: The algorithm builds a mathematical model from a set of data which contains only inputs and no desired output labels.
- **Semi supervised learning**: learning algorithms develop models from incomplete training data where portion of data doesn't have labels

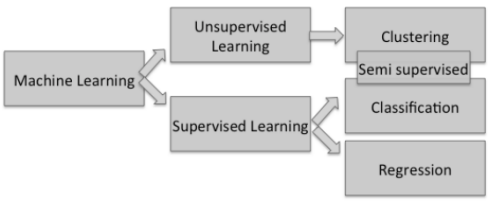

Fig 1. Machine learning methods

This paper focuses on increasing the success rate of IVF treatment by creating a model which predicts the outcome of treatment is success or failure by studying various parameters. This is considered to be a classification problem. Classification algorithms are supervised algorithms in which outputs are restricted to a limited set of value. In our problem it is 1 for success and 0 for failure. An algorithm that implements classification is known as classifier. It refers to the mathematical function implemented by classification algorithm that maps input data to a category. Some of the classifiers used in predicting the outcome of IVF treatment are described below:

A.

Naive bayes classifier

Naive bayes classifiers are a family of simple probabilistic classifiers based on applying bayes theorem with strong independence assumption between the features.

Bayes' theorem finds the probability of an event occurring given the probability of another event that has already occurred. Bayes' theorem is stated mathematically as the following equation

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \quad (1)$$

Where A and B are events and P(B)≠0 , P(A) is the prior probability , P(A|B) is a posterior probability.

In [1],G .Korani et al in their work proposed Embryo-Uterine model in which they have assumed it is necessary to have both a receptive uterus and viable embryo to result in successful pregnancy. They considered θ_e and θ_u respectively the probabilities of the embryo to be viable and of the uterus to be receptive The EU model estimates the probability of pregnancy after the transfer of a single embryo using the formula (2) where θ_e is the probability of the embryo to be viable and θ_u is the probability of uterus to be receptive assuming the independence of viability and receptivity.

$$P(\text{Success}) = \theta_e * \theta_u \quad (2)$$

But this formula failed to estimate the probability of multiple pregnancies when more than one embryo was transferred. This technique is also considered as "partial observable" problem as we will not be able to identify out of two features, which feature led to failure.

In the next version of this work, two parent model was constructed by adding some more features like age of the female, score of viable embryos, number of IVF ,cycle if the patient had undergone IVF cycles before and ICSI (intracytoplasmic sperm injection). The Bayesian network constructed is shown in figure 2

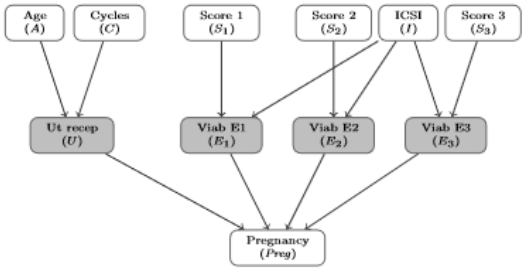

Fig 2: Bayesian structure of two parent model

Scores were given to embryos by observing the morphological features of embryos at specific intervals which helps doctors to select the most viable embryos in order to improve the success rate by decreasing the rate of multiple births. This model yields better results as it considers the factors of both the partners. In this analysis it is also found that embryo viability is critical factor in the success of IVF treatment. But just by considering the morphological features it is not easy to identify the quality of embryos. In the recent work, time lapse images are considered to score the embryos to identify the most viable one. In [14], author has proposed embryo selection using time lapse monitoring. With this technology, the embryos can be monitored without removing them from the incubator. A camera is built into the incubator and takes pictures of the embryos at preset intervals. With the help of the proper software, a video can be made that depicts their development. This type of monitoring allows for the collection of much more information on the timing of the cleavages and the dynamics of the morphologic changes. But having a machine learning model which selects the viable embryos through these images would be better.

B. *Decision tree classifier*

It is a simple and widely used classification technique. This classifier poses a series of carefully crafted questions about the attributes of the test record. Each time it receives the answer, a follow up question is asked until a conclusion is drawn about the class label.

In [4], hybrid intelligence method which integrates genetic algorithm and decision learning techniques for knowledge mining of an IVF medical database was developed. The 28 most significant attributes for determining the pregnancy rate (e.g., patient's age, number of embryo transferred, number of frozen embryos, and culture days of embryo) and their combinative relationships (represented by if–then rules) were identified through this method. This model had a predicted accuracy of 73% and the corresponding sensitivity and specificity were 71% and 73% respectively. This expert system helps the doctors to answer the question of "what are the chances of having success in the treatment?" posed by the infertile couple undergoing treatment. However, this treatment doesn't help the doctors to choose most viable embryos.

C. *Random forest*

Random forests or random decision forests are an ensemble learning method for classification that operates by constructing a multitude of decision trees at training time and outputting the class that is the mode of classes or mean prediction of the individual trees. It aggregates the vote of different decision trees to decide the final class.

In [12], random forest model was constructed using 25 attributes and the prediction ability of this model was checked using the performance metrics such as accuracy rate, F-measure and AUC under ROC curve. Same model was again tested by minimizing the number of features to 16. Feature selection algorithm in [12] also tested the other models like multi layer perceptron (MLP), support vector machine. As a conclusion, this work identified the most influential attributes which decide the success of treatment. Some of these are 'age', 'indication of fertility factor', 'method of sperm collection', 'follicies on day 14', 'Embryo transfer day'.

D. *Support Vector machines (SVM)*

In machine learning, support vector machines are supervised learning models with associated learning algorithms that analyze the data used for classification. In this model, given labeled data, the algorithm outputs hyper plane which categorizes new examples. RIMARC (Ranking Instances by maximizing the ROC curve) algorithm is based on support vector machines. This is a binary classification methodology that ranks the instances based upon how likely they are to have a positive label. This is based on ROC curve and attempts to maximize the area under ROC curve. This algorithm learns a ranking function which is a linear combination of non-linear score functions constructed for each feature separately.

In [5] author proposed a SERA (success estimation using ranking instances) which uses RIMARC algorithm for ranking the features. In this work, author has considered 64 independent features in which 52 of them are related to female and 12 of them are related to male and one dependent feature called Result.

TABLE I. FEMALE ATTRIBUTES

Age	Blood_type	Height	Weight	BMI	Tubal factor
Ovulatory_Dysfunction	Unexplained Infertility	Severe_Pelvic_Adhesion	Laparoscopy	Hysteroscopy	Hystroscopic_surgery
Abdominal_surgery_category	Abdominal_Surgery	Abdominal_Surgery_Category	Abdominal_Surgery_Category	Ovarian_Surgery	Cyst_Aspiration
Tubal_Surgery	Uterine_Surgery	Endometriosis	Cycle_No.	Baseline_FSH	Embryocryo
Baseline_LH	Baseline_E2	Gravida	Aabortus	Living childrenn	Age_related infertility
Diabetes mellitus	Hypertension	Duration_Infertility	Hydrosalpinx	Office_Hysteroscopic_Incision	Abdominal suregry
PCOS	HSG*_Cavity	HSG*_Tubes	Office_Hysteroscopy	Office_Hysteroscopic_Procedure	Hydrosalpinx
Total_Antral_Follicle_Count	Right_Ovarian_Antral_Follicle_Count	Thyroid_Disease	Anemia	Laparotomy	Left_Ovarian follicle_count
Hyperprolactinemia	Hepatitis	Endometrioma_surgery	Localization_Myoma_uteri		

Table I shows the variables related to Female considered in the study

TABLE II. MALE ATTRIBUTES

Male_Karyotype	Male_Genital_Surgery	Sperm_Count
Male_Age	Semen_Analysis_Category	Sperm_Motility

Male_Blood_Type	Male_FSH	Total_Progressive_ Sperm_Count
Sperm_Morphology	Testicular_Biopsy	Testicular sperm extraction

Table II shows the variables related to Male considered in the study. The RIMARC algorithm calculates the feature weights and creates rules that are in a human readable form and easy for the doctors to interpret. SERA algorithm outperformed other classifiers used in the study in terms of Area under curve and accuracy.

Artificial neural networks

Artificial neural network is an interconnected group of nodes, inspired by a simplification of neurons in a brain. It is based on the collection of interconnected units or nodes called artificial neurons. Each neuron and edges of ANN typically have a weight that adjusts as the learning proceeds. In recent years, application of ANNs found in a number of fields including medical, biological, and engineering, robotics, social and business applications. ANNs can effectively replace traditional statistical prediction methods with better accuracy. In [2], the different factors such as female age, duration of infertility ,BMI(Body mass index), previous surgery, previous pregnancy, endometriosis, tubal causes, ovulatory factor, sperm concentration , sperm vitality, number of oocytes retrieved, number of embryos transferred are considered for constructing a artificial neural networks.

Figure 3 show the artificial neural network constructed using these features.

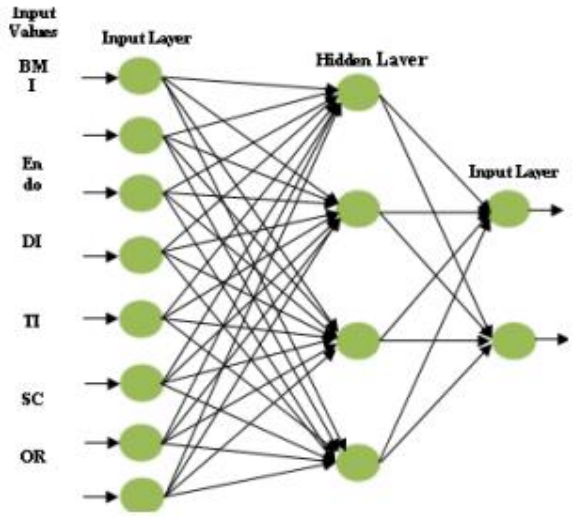

Fig 3. Artificial neural network

Here DI=duration of infertility, SC = sperm concentration, OR=Number of oocytes retrieved, ET=Embryos transferred, TI=Tubal infertility

The dataset from the infertile couples who successfully conceived after IVF treatment has been used for training the network. Model constructed was the multi layer perceptron with a single hidden layer. This ANN reported 79% of accuracy.

In [6], neural network was constructed using the variables of age, number of eggs recovered, number of embryos transferred and whether there was embryo freezing. Overall the network managed to achieve an accuracy of 59%

IV. Evaluation of the Techniques

For selecting the good classifier from the set of classifiers cross validation is the most widely method. The selection of features for the classifier is also one important factor to get better results. In [12], author proposes hill climbing approach for choosing the best subset of features. This algorithm initially selects the most influential feature set that positively improves the classification results and then repeatedly adds one feature at a time to the selected feature set that positively improves the classification results or provides the least reduction in classification accuracies.

To evaluate the performance of the various classifiers, the following metrics can be used

1) Classification accuracy: ratio of all samples that are classified to the total number of test samples
2) Sensitivity: ability of the classifier to accurately predict the successful pregnancy
3) Specificity: ability of the classifier to accurately predict the unsuccessful pregnancy
4) Area under AUC curve: ROC curve is plotted as true positive rates at y axis and false positive rates at x-axis.

In neural networks performance varies with number of hidden layers, initial weights and learning parameters.

V. Comparison

The table III shows the comparison between the different machine learning techniques studied in this survey.

TABLE III . COMPARISON OF ML TECHNIQUES

Study reference	ML technique used	No. Of attributes selected	Performance
Kaufmann et al[6]	Artificial neural network	4	59% accuracy
Corani G et al[1]	Naive Bayes	2(EU model) 5(Two parent model)	AUC 67.0 for EU model AUC 83.0 for double parent model
Durairaj and thamilselvan[2]	Artificial neural network	8	73% accuracy
Guvenir et al.[5]	Support vector machine	64	84% accuracy
Uyar et al[13][11]	Naive Bayes	18	80.4% accuracy
Guh et al[4]	Integration of genetic algorithm and decision trees	38	73.2% accuracy
Ramaswamy N et al[10]	Multilayer perceptron, Naive Bayes	18	MLP-90.35% accuracy

VI. CONCLUSION

From this survey, it is found that the machine learning techniques can be used at different phases of IVF treatment. Some of the work discussed in the paper focuses on developing a model which helps the couples to take the decision whether to go for an IVF treatment or not .Patients data such as age, BMI, sperm count etc are used to train the model. One more group of work takes into account some additional features like the number of oocytes retrieved, number of embryos transferred and builds a model which will tell what is the probability of having pregnancy in the treatment. This will help the patients to mentally prepare themselves to face the result. Another set of work focuses on helping the doctors to choose the most viable embryos by observing their morphological features or by time lapse images. Using this technique, embryologist can take a right decision on selecting the embryo which will result in pregnancy. There are various features to be considered for building the model. Some models consider up to 64 features whereas some considers as low as 5-6 features. And also there are many classifiers available. Selecting the features and classifier is a challenging task. Different parameters like accuracy rate, specificity, sensitivity, area under ROC curve are some of the performance metrics used to compare the classifiers.

REFERENCES

[1] Corani G, Magli C, Giusti A, Gianaroli L, Gambardella LM (2013) A Bayesian network model for predicting pregnancy after in vitro fertilization. Comput Biol Med 43(11):1783–1792
[2] Durairaj M, Thamilselvan P (2013) Applications of artificial neural network for IVF data analysis and prediction. J Eng Comput Appl Sci 2(9):11–15
[3] Durairaj M, Nandhakumar R (2014) An integrated methodology of artificial neural network and rough set theory for analyzing IVF data. In: 2014 International conference on intelligent computing applications (ICICA), pp 126–129
[4] Guh R, Wu TCJ, Weng SP (2011) Integrating genetic algorithm and decision tree learning for assistance in predicting in vitro fertilization outcomes. Expert Syst Appl 38(4):4437–4449
[5] Gu ¨venir HA, Misirli G, Dilbaz S, Ozdegirmenci O, Demir B, Dilbaz B (2015) Estimating the chance of success in IVF treatment using a ranking algorithm. Med Biol Eng Comput 53(9):911–920
[6] Kaufmann SJ, Eastaugh JL, Snowden S, Smye SW, Sharma V (1997) The application of neural networks in predicting the outcome of in vitro fertilization. Hum Reprod 12(7):1454–1457
[7] Manna C, Nanni L, Lumini A, Pappalardo S (2013) Artificial intelligence techniques for embryo and oocyte classification. Reprod Biomed Online 26(1):42–49
[8] Milewski R, Milewska AJ, Wiesak T, Morgan A (2013) Comparison of artificial neural networks and logistic regression analysis in pregnancy prediction using the in vitro fertilization treatment. Stud Log Gramm Rhetor 35(1):39–48
[9] Morales DA, Bengoetxea E, Larran ˜aga P, Garcı ´a M, Franco Y, Fresnada M, Merino M (2008) Bayesian classification for the selection of in vitro human embryos using morphological and clinical data. Comput Methods Programs Biomed 90(2):104–116
[10] Ramasamy N, Durairaj M (2017) Feature reduction by improvised hybrid algorithm for predicting the IVF success rate. J Adv Res Comput Sci 8(1):37–40
[11] Uyar A, Bener A, Ciray HN, Bahceci M (2010) ROC based evaluation and comparison of classifiers for IVF implantation prediction. In: Electronic healthcare, pp 108–111
[12] Md Rafiul Hassan, Sadiq Al-Insaif, M. Imtiaz, Joaeder Kamruzzaman, “A machine learning approach for prediction of pregnancy outcome following IVF treatment”,Neural computing and applications, August 2018
[13] Uyar A, Ayse B, Ciray HN (2015) Predictive modeling of implantation outcome in an in vitro fertilization setting: an application of machine learning methods. Med Decis Mak 35(6):714–725
[14] Peter Kovacs ,Embryo selection:the role of time-lapse monitoring ,BioMed Central ,2014 Dec 15