Article - Engineering, Technology and Techniques

Dynamic Model for Assisted Reproductive Technology Outcome Prediction

Ranjini Kothandaraman[1]*
https://orcid.org/0000-0003-4101-927X

Suruliandi Andavar[2]
http://orcid.org/0000-0002-2863-1555

Raja Soosaimarian Peter Raj[3]
https://orcid.org/0000-0002-7216-2207

[1]ManonmaniamSundaranar University, Department of Computer Science and Engineering,Tirunelveli, Tamil Nadu, India; [2]ManonmaniamSundaranar University,Department of Computer Science and Engineering, Tirunelveli, Tamil Nadu, India; [3]Vellore Institute of Technology, School of Computer Science and Engineering, Vellore, Tamil Nadu, India.

*Correspondence: ranjinirajalingam@gmail.com (R. K.).

HIGHLIGHTS

- Proposed a dynamic model for ART outcome prediction

- Partial implementation of the model with the help of machine learning incremental classifier

- Performance evaluation of the model with the state-of-art methods.

Abstract: Infertility is becoming a growing issue in almost all countries. Assisted Reproductive Technologies (ART) are recent development in treating infertility that give hope to the infertile couples. However, the pregnancy rates achieved with the aid of ART is considerably low, as success in ART is not only based on the treatment but also on many other controllable and uncontrollable biological, social, and environmental features. High expenditures and painful process of ART cycles are the two major barriers for opting for ART. Moreover, ART treatments are not covered by any health insurance schemes. Computational prediction models could be used to improve the success rate by predicting the treatment outcome, before the start of an ART cycle. This may suggest the couples and the doctors to decide on the next course of action i.e. either to opt for ART or opt for correcting determinants or quit the ART. With the intension to improve the success rate of ART by providing decision support system to the physicians as well to the patients before entering into the treatment this research work proposes a dynamic model for ART outcome prediction using Machine Learning (ML) techniques. The proposed dynamic model is partially implemented with the help of an ensemble of heterogeneous incremental classifier and its performance is compared with state-of-art classifiers such as Naïve Bayes (NB), Random Forest (RF), K-star etc.,using ART dataset. Performance of the model is evaluated with various metrics such as accuracy, Precision Recall Curve (PRC), Receiver Operating Characteristic (ROC), F-Measure etc., However, ROC cure area is taken as the chief metric. Evaluation results shows that the model achieves the performance with the ROC area value of 94.1 %.

INTRODUCTION

Infertility, now-a-days, is becoming a public health issue in almost all countries due to the changes in the lifestyle of the people [1]. This is to be considered as one of the most important social problems as it is going to cause a structural impact and imbalance in the next generation [2]. World Health Organization states that one in every four couples, either men or women or both, are affected by infertility [3]. In India 27.5 million couples are suffering from infertility based on the survey held on 2019 [4]. The treatment to infertility is becoming a booming business as more and more couples are suffering difficulties to conceive [5].

Assisted Reproductive Technology (ART) is a medical procedure which is considered as a last recourse for infertile couples [6].ART includes all the treatments that handle human sperms, occytes or embryo in-vitro to establish pregnancy. In-Vitro Fertilization (IVF) and Intra Cytoplasmic Sperm Injection (ICSI) are the most common methods under ART where the female germ cells (oocytes) are inseminated by sperm/s under laboratory condition. Besides the expense and pain, the probability of success of ART is very low as it involves various determinants like age of the women, condition of the uterus, number and the quality of the oocytes retrieved, morphology of sperms, quality of embryo developed and many more [7-8]. Also, there may be chances for tribulations in every stage of the process from conception to delivery of the baby.

Moreover, the cost and the emotion beyond every cycle affect the success rate of the treatment. Repeated attempts of the treatment affect the physical and mental health of the couples too. Presently, decision making in the ART treatment is usually based on the combination of patient particular characteristics and the physician's knowledge and clinical experiences. For most clinicians, the syntheses of previous experience with the current situation become almost intuitive with time.

The probability of success of ART can be increased by perfecting the determinants that affect fertility by identifying them by their level of significance and treating them [9-10]. But it is a complicated task for the doctors and embryologist to correlate all the determinants since the number of determinants is significantly large and have complex inter relationships. This necessitates studies trying to forecast the probability of success of the treatment by analyzing the complicated interlink between the determinants using automated tools to supplement the efforts taken by the doctors and the patients in achieving higher success rate.

Machine Learning (ML) is one of the tools which are used for prediction.The literature studied makes it clear that ML classifiers are used to build models to predict possible outcomes. The classifiers used in most studies worked with static data, though ART data is likely to be dynamic, with the influencing variables changing with respect to environmental characteristics. Consequently, a dynamic predictive model that describes all the influencing attributes is lacking.

Hence the objective of this research work is to propose a dynamic model for ART outcome prediction. The part of the proposed dynamic model is implemented with a dynamic machine learning classifier named ensemble of heterogeneous incremental classifiers which was proposed in [11]. Finally the performance of the proposed dynamic model is check with other state-of-art classifiers used for ART outcome prediction.

Related work

Chen and coauthors [12], proposed two phase Particle Swam Optimization (PSO) approach to predict IVF outcome using 654 IVF cycle data with 10 attributes. They also obtain an optimal rule set from the model by encoding decision rules into a particle position. Ruey-ShiangGuhand and coauthors [13], introduced a hybrid algorithm by integrating genetic algorithm and Decision Tree (DT) C4.5 to tailor the IVF process by analyzing 5275 records with 69 attributes. Thamilselvan and Durairaj[14], aimed to predict the success rate of IVF using Artificial Neural Network (ANN). They construct a Multilayer Perceptron (MLP) having 8 input layers by giving 8 attributes. The number of nodes in the hidden layer varied as per the validation of data. The output layer is for producing success rate of IVF treatment. The system is trained using Back Propagation Algorithm. Milewskiandand and coauthors [15], compared ANN with Logistic Regression for IVF dataset from USA with 26 attributes and argued that ANN may be better suited for predictive model for clinical treatment. Also stated that Logistic Regression was is good in selection of determinants as well as to find the degree of influence of the determinants on the final result. Giivenirand and coauthors [16], with an aim to determine the attributes and their particular values that affect the outcome of an IVF treatment, proposed a Success Estimation Ranking Algorithm (SERA) that is implemented using the Ranking Instance by Maximizing the Area Under Receiver Operating Characteristics Curve (RIMARC). Ramasamy [17], introduced an improvised

hybrid algorithm which combines the existing Ant Colony and Relative Reduct Algorithm for feature reduction of IVF database containing 42 attributes. They argued that their proposed algorithm achieved its target of reducing the features to minimum number without compromising the core knowledge of the system to estimate the system success rate. Raefandand and coauthors [18] and Hafiz and coauthors [19], tried to choose the best predictive model for calculating the probability of IVF/ ICSI success for couples using a comparative study among various classifiers namely Support Vector Machine (SVM), Recursive Partitioning (RPART), Random Forest (RF), Adaptive Boosting (AdaBoost) and One – Nearest Neighbour (1NN). According to their dataset of 486 patients with 29 attributes they found that RF and RPART outperformed the other comparable methods. Hassan and coauthors [20], proposed hill climbing feature (attribute) selection algorithm coupled with automated classification using machine learning techniques to increase the accuracy in IVF pregnancy prediction. They use 25 attributes and 5 machine learning classifiers namely MLP, SVM, C4.5, Classification and Regression Tree (CART), and RF.

A brief Critical Overview of the Existing Methods

A critical overview of the classifiers used for ART related work are discussed here. The classifier C4.5 Decision Tree well handled both numerical and categorical data and it was used to induce rule set for prediction. Adaptive boosting and ensembles enhanced Decision Tree Learning. Artificial Neural Network performed well in finding clinical results whereas it was poor in finding influential attribute. Logistic Regression found correlations between the attributes and it suited best to find the influencing attribute and to what degree it influenced the final result. But its performance degraded when the attributes were dependent. Naïve Bayes (NB) outperformed Logistic Regression in exactness and it produced similar performance for both oversampling and under sampling of classes. Radial Basis Function also produced similar performance as that of NB but most of the researchers recommended Naïve Bayes for ART outcome prediction. RPART was mainly used to assign score to the variable [21]. Most of these classifiers are taken for comparison with the proposed method.

Motivation and Justification of the work

Infertile couples face lot of mental and physical suffering, and ART is considered as a last resort. However, the limitations of ART like its low success rate, physical pain, and the considerable expense, putting it beyond the reach of most people. Repeated attempts of the treatment affect the health of the couples too. Having a system in place that is capable of predicting ART outcomes, prior to commencing treatment may help doctors and patients in deciding whether to opt for the treatment or attempt to improve the determinants for success.

From the literature it is understood that there exists some models to predict the outcome of ART treatment by analysing limited amount of data collected from their regional fertility centers. Also that, all the models developed so for are static in nature and there is no way for refinement [21]. In reality, the ART data is dynamic in nature and change as and when new determinants are identified by physicians which may lead to new patterns. Building new models as and when the data is updated is a tedious task.

Hence it is observed that there is a need for a dynamic model that supplements doctors' and patients' effort in achieving higher success rate in ART by analysing all the determinants which affect reproductive fitness. The expected model should consider all determinants that could be measured during fertility treatment to give prediction. The model should be self revisable on getting the data set with a new pattern. The proposed model may help to identify the success rate and suggest the patients for undergoing different types of treatment required for the individual. Motivated by this, this paper proposes a dynamic fertility model which has the capability to predict the success probability of a couple before going for the treatment as well as give suggestions to go for further treatments. The proposed model has the capability to refine itself when its prediction went wrong or it gets a new pattern. In machine learning, these kinds of dynamic model can be developed with the help of incremental classifiers which has the capability to learn from the historical data and give prediction about the future and also able to update itself on receiving new data. Hence it is justified that this research work proposes a dynamic model for ART outcome prediction and part of the proposed model is implemented with the help of incremental classifiers in ML.

Outline of the work done

The outline of the work done is shown in Figure. 1, which starts with the architecture to generate a dynamic model for ART outcome prediction, which is followed by the implementation of the proposed model

with the help of an ensemble of heterogeneous incremental classifier and ends with the performance evaluation of that classifier with the existing state-of-art classifiers.

The rest of the paper is organized as follows. Section 2 describes the Materials and methods of proposed Model Generation; Section 3 discusses the results of the partial implementation of the model and section 4: Concludes the findings.

MATERIAL AND METHODS

This section explains the architecture of the proposed model, with the aim to improve the success rate of ART especially in IVF/ICSI by giving suggestions to doctors as well as to the patients by an automated system which has the capability to learn itself.

Pre-processing – SMOTE

Low success rate of ART makes the dataset imbalance with less number of data for positive outcome. The performance of the classifiers gets trembled with imbalanced data and the usual performance metric accuracy may mislead. Hence experiments are carried out to check the right performance evaluation metric for the imbalanced data and found that ROC shows unbiased result even for imbalanced dataset [24]. Imbalance in the dataset can be handled by sampling techniques. Hence it analysed three sampling techniques namely Undersampling, Oversampling and Synthetic Minority Oversampling Technique (SMOTE) for various proposition of data and found that SMOTE is the best sampling method for balancing the imbalanced ART dataset. So the ART dataset is balanced using SMOTE [24]**.** In SMOTE, synthetic samples of minority class will be generated to balance the dataset.

Architecture of the proposed model

The proposed machine learning model is expected to work in three phases.
   a. Predictive Model Generation Phase
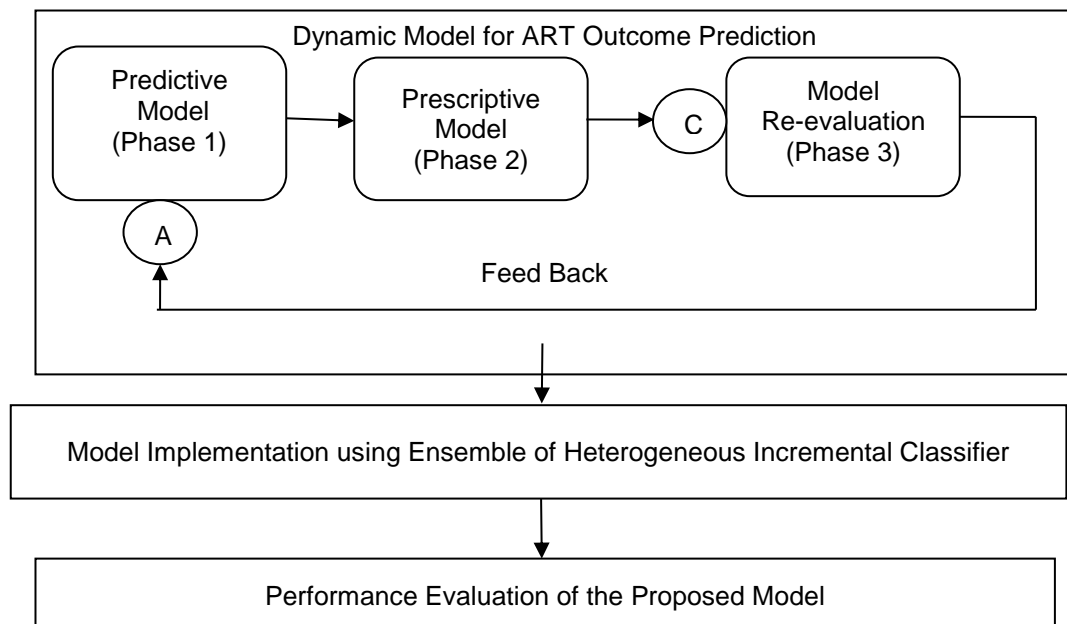   b. Prescriptive Model Generation Phase
   c. Model Re-evaluation Phase

Figure 1. Outline of the work done

Predictive Model Generation

Predictive modelingis a process that uses machine learning and probability to forecast outcomes. Eachmodelis made up of a number of predictors, which are variables that are likely to influence future results. Once data has been collected for relevant predictors, a statisticalmodelis formulated.

From the literature the various determinants (features/variables/factors/predictors/ attributes) that affect fertility are identified and some of the important factors are listed in Table 1.

Table 1. Infertility determinants identified from the literature

General factors	Hormone factor	Male factor
The general factors that are considered during treatment	Hormone levels that affects the chances of a couple achieving a pregnancy	Determinants relating to man that adversely affect the chances of a couple achieving a pregnancy.
1.Woman age 2.Man age 3.Duration of inf0ertility 4. Primary / secondary infertility 5.Body Mass Index (BMI) 6. Women Occupation 7. Men Occupation 8. Woman Education 9. Men Education	1.Follicle-stimulating hormone(FSH) 2. Luteinizing hormone (LH) 3. Prolactin 4. Dehydroepiandrosterone sulfate (DHEA-S) 5. Testosterone 6. Progesterone 7. Human chorionic gonadotropin (**h**CG) level – day 11 8. Anti- mullerian 9. Inhibin A 10. Thyroid Stimulating Hormone (TSH) 11. prolactin	1.Sperm quality 2. Total sperm count or sperm concentration 3. Morphology or normal forms 4. Motility or progressive motility 5. Quality of motility 6. Semen Volume 7. History of male urethritis 8.Varicocele 9. General Health 10. Allergies 11. Previous Treatment 12. Surgical Treatment
Female factor	**IVF/ ICSI factors**	**Other factors**
Determinants relating to woman that adversely affects the chances of a couple achieving a pregnancy.	Factors related to ART treatment that determines the chances of a couple achieving a pregnancy.	The other day-to-day activities that may affects the chances of a couple achieving a pregnancy
1. Ovarian factor 2. Tubal factor 3. Cervical factor (mucus) 4. History of Previous pregnancy / Complication 5. Previous childbirth 6. Ovarian size 7. Fibroid 8. Duration of ovarian stimulation 9. Endometriosis 10. Sub and intra endometrial vascular signals 11. Endometrium thickness 12. Endometrium morphology 13. Menstrual Cycle Detail 14. General Health 15. Fertility Medications 16. Surgical Treatment	1. Number of Previous Cycle 2. Number of embryos 3.Morphology score of the best and second best embryo 4.Fertilization rate 5.Method of fertilization 6.Ovulation Stimulation 7.Number of good quality embryos 8. Day of embryo transfer 9.Number of good quality embryos transferred 10.Number of retrieved oocytes 11.Number of pre-Ovulatory follicles 12.Proportion of fertilized oocytes 13. Physician performing Embryo Transfer 14. Number of frozen Embryo 15. Sperm Penetration assay 16. Outcome	1. Daily coffee 2. Smoking habits (current/former) 3. Alcohol 4. Tobacco 5. Stress measures 6. Relaxation method 7. Psychological Treatment 8. Exercise 9. Unknown factors 10. Residential Area 11. Family History 12. Weight Change 13. Others

In the ART outcome Predictive Model, historical data will be collected from the fertility centers. The data must include all the factors that affect fertility for both men and women. These factors will be collectively called as Reproductive Fitness Factors. The history of treatment and the medication the couples have underwent and their results should also be noted. Other than this, the social and environmental behavior of the patient must also be taken into consideration.

Once this data is collected, predictive model for fertility can be generated by following the below steps.
Step 1: Preprocess the raw data and convert it into the needed form. Since it is medical data the preprocessing is only to check whether any validation or transformation is required in the data.

Step 2: Store the processed data in the data storage
Step 3: Identify the significant feature set based on ART related Rules
Step 4: Apply Data Analytics using Machine Learning Algorithm. For Predictive Model generation Supervised Learning will be more useful.
Step 5: Identify the patterns and probability of success for that patterns
Step 6: Store the patterns as success and Failure Pattern based on the probability.
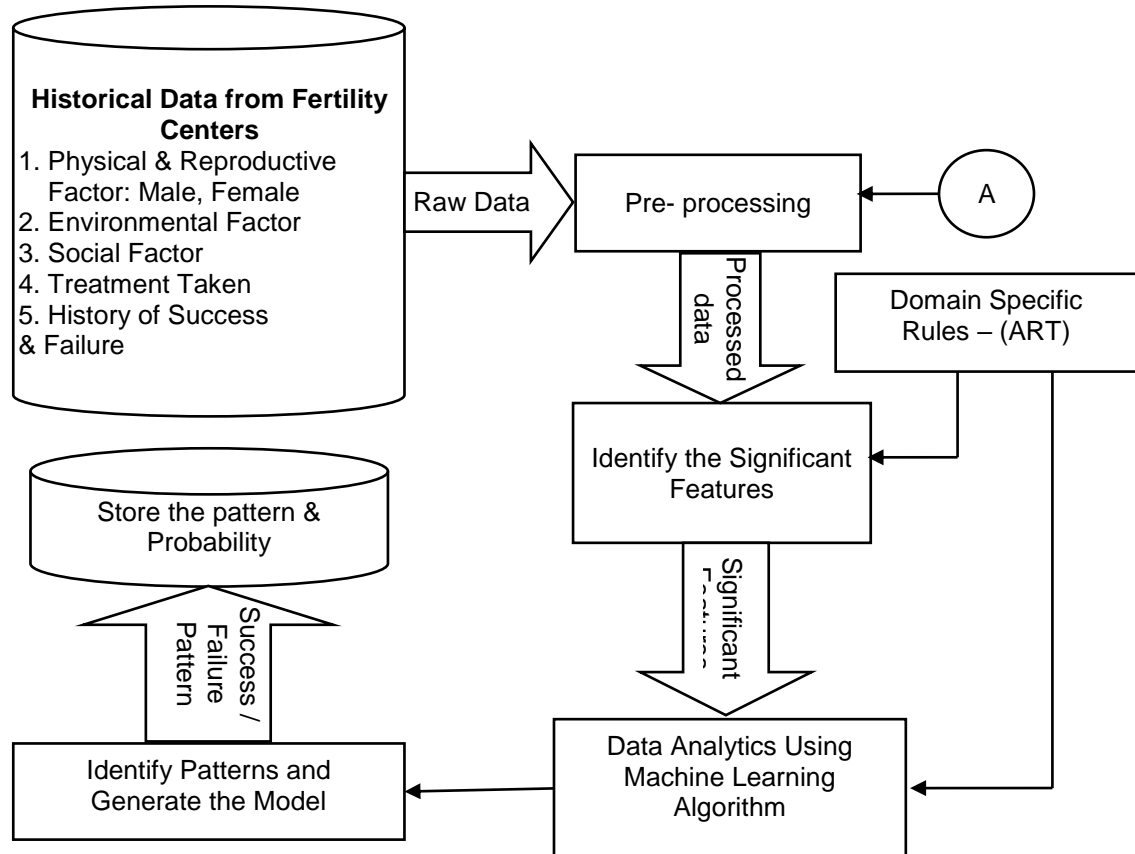Figure. 2 depicts the steps involved in predictive model generation

Figure 2. Fertility Predictive Model Generation

Prescriptive Model Generation

Once the predictive model is generated, the model will have the capability to tell the success probability of a couple. The prescriptive model will function as a testing model which will recommend a specific couple to go for IVF/ICSI treatment in the current cycle or not. The model also has the capability to prescribe course of actions (treatment/ medication) to be taken for that specific couples. The steps involved in the prescriptive model are given below.

Step1: New couples details will be taken as input
Step 2: Preprocess that input and convert it into needed form
Step 3: Check with the stored patterns in phase 1 (predictive model generation) to find whether the pattern is found or not

Step 3.1: If the pattern is found, check whether it is a positive (Success) pattern or Negative (Failure) pattern.
Step 3.1.1: If it is a positive pattern recommend the couples to go for IVF/ ICSI treatment and prescribe course of action to be taken based on the pattern.
Step 3.1.2: If it is a negative pattern the couples will not be recommended to go for IVF/ICSI treatment in the current cycle. Instead course of treatment and medication will be prescribed to make the pattern positive.

Step 3.2: If the pattern is not found means, the couples will be recommended for treatment and the outcome of the treatment will be followed up and the pattern will be updated in the database

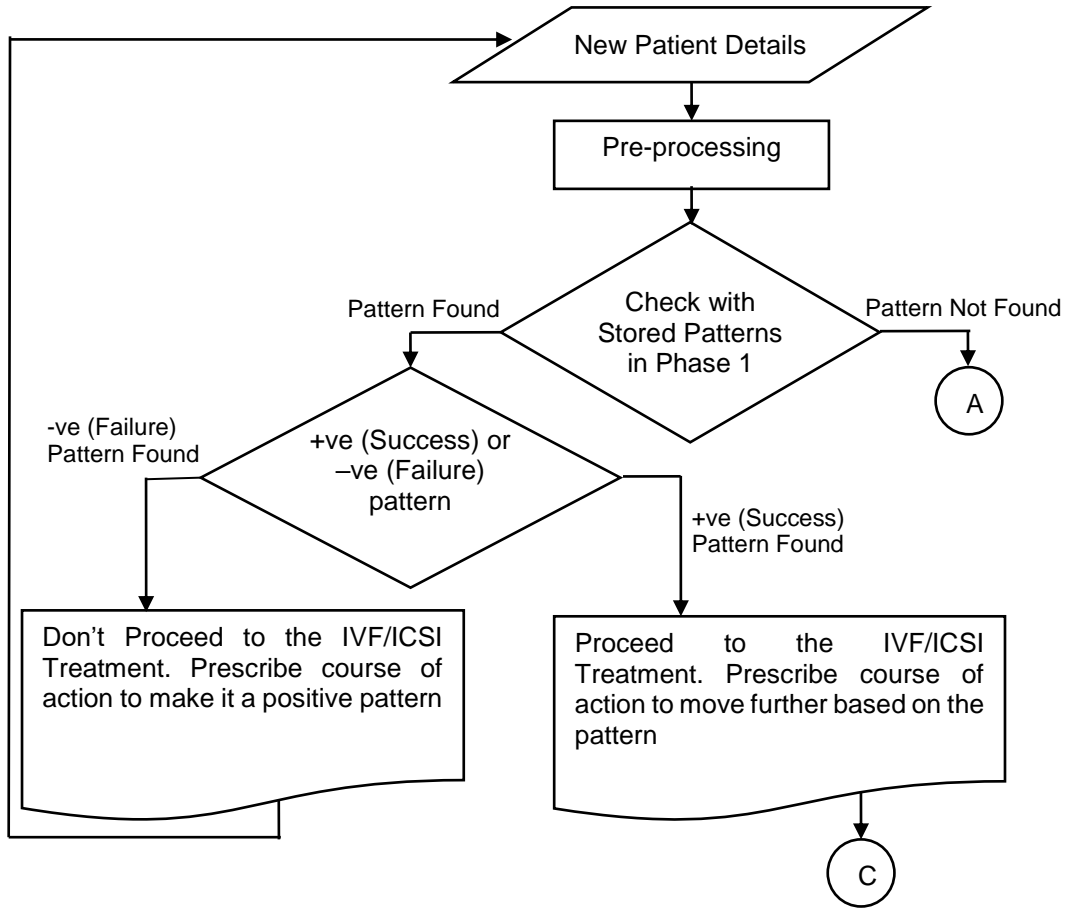The diagram representing the above steps is given in the Figure 3.

Figure 3. Prescriptive Model Generation

Model Reevaluation

Model Reevaluation is self checking phase. In this phase the system will check whether the system recommended couples get positive result. If the IVF/ICSI treatment result is positive for the recommended couples, it is identified as the success of the system so the success probability of that specific pattern of the couples will be updated. Or otherwise, if the system recommended couple get the negative result, it indicates the absence of some more determinants which needs to explore yet.

After identifying the new determinant, if the couples get positive result the pattern will be updated with the newly added determinant or added as a new pattern based on the need. This process will be repeated by sending feedback to the predictive phase. The re-evaluation phase is depicted in Figure 4.
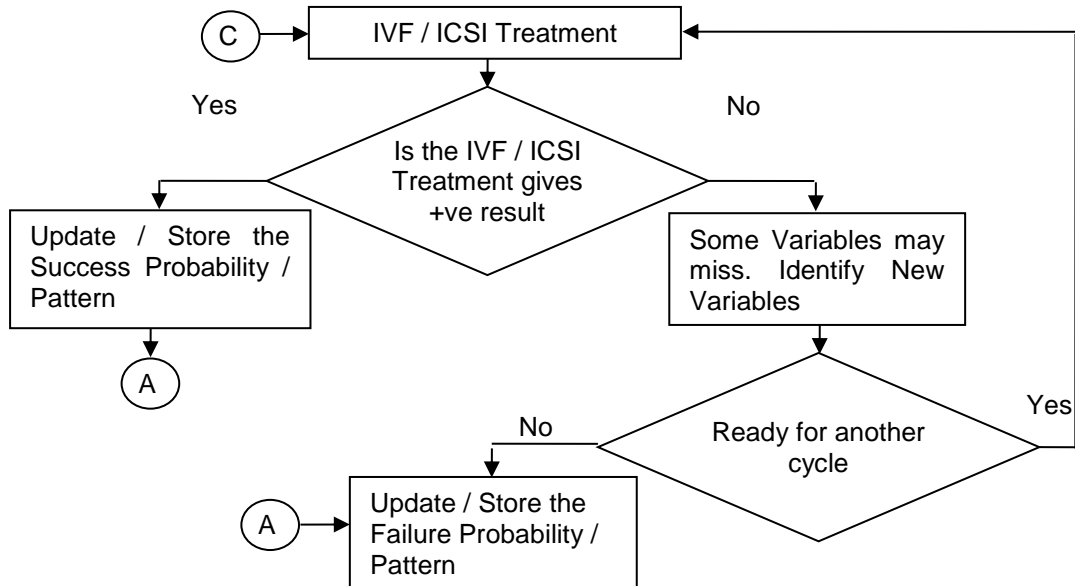
Figure 4. Re-evaluation Model

Ensemble of Heterogeneous Incremental Classifiers (EHIC)

The proposed idea of ART outcome prediction model is implemented with the help of an ensemble of heterogeneous incremental classifiers which was proposed in [11] that was built by combining Instance Based (IB1) Learner and Averaged One Dependence Estimators (A1DE) Updatable. The algorithm for the dynamic incrementally updatable ensemble learner is discussed hereunder.

$\textbf{\textit{Algorithm}}$: Incrementally Updatable Ensemble Learner
$\textbf{\textit{Parameters}}$:

      Base Classifiers   : IB1, A1DE Updatable
      Combination Rule $\prod$: Product of Probability
      H_t — Pediction of Base Classifier t
      H^* — Pediction of Ensemble Classifier
      T — Number of Base Classifiers : 2

1. $\textbf{\textit{Input}}$: $Training\ Set\ D = \{\ x_i\ ,y_i\}_{i=1}^m$
2. $\textbf{\textit{Output}}$: $Prediction\ of\ Ensemble\ Classifier\ H^*$
3. $Step\ 1$: $Learn\ Base\ Classifiers$
4. $for\ t = 1\ to\ T\ do$
5. $Learn\ H_t\ based\ on\ D$
6. $end\ for$
7. $Step\ 2$: $Combine\ using\ Combination\ Rule\ for\ Ensemble$
8. $for\ t = 1\ to\ T\ do$
9. $H^* = \prod H_t$
10. $end\ for$

Hyper parameter Settings for the EHIC with IB1 and A1DE Updatable

The IB1 classifier searches neighbouring instances to find the similarity between instances using distance measures. The A1DE updatable learner considers the relationship between attributes, super parents, and class. There is a possibility of setting differing values for the different parameters used in both the classifiers. The optimal performance of the classifiers is obtained by setting the correct parameters for them. In order to find the optimal parameters, experiments are carried out by setting different values for the options available for the classifiers [11]. The final chosen values of the parameters are shown in Table 2.

Table 2. Parameters for the EHIC.

Parameter Name	Available Values	Chosen Value	Description
Search Method	1. Linear Search 2. KD Tree 3. Ball Tree 4. Cover Tree	Linear Search	Used to searches the neighbouring instance
Distance Measure	1. Chebyshev distance 2. Euclidean distance 3. Manhattan distance 4. Minkowski distance	Euclidean distance	Used to find the similarity between the instances
Distance Weighting	1. 1/ distance 2. 1- distance	1/ distance	To give more influence on the predicted value for neighbors closer to the data point
Weighted A1DE	1. True 2. False	False	Weights are calculated based on mutual information between attribute and the class
Combination Rule	1. Average of Probability 2. Product of Probability 3. Majority Voting 4. Minimum Probability 5. Maximum Probability	Product of Probability	Shows how the classifier are combined together to produce the results

The proposed ensemble model is thus built by setting the values for the parameters as shown in Table 2, and the two base learners, namely, IB1 and A1DE updatable learner, are combined by product of the probability voting method.

RESULTS

A critical analysis of the exiting ML methods that are used for ART outcome prediction is done in [21]. This section will check the ability of the proposed dynamic model building for ART outcome prediction using EHIC by comparing it with the already existing classifiers that are identified in the literature for ART related research.

Dataset

A dynamic ART dataset which was maintained by Human Fertilization and Embryo Authority (HFEA) are used for experiment purpose [22]. Initially, 16383 records with 52 attributes are taken for the study. This dataset contains most of the determinants that are listed in Table 1. Only few determinants related to the personal details of the couples such as occupation, education and routine behaviors that are categorized as other factors are missing. In the fertility dataset, for best results, some data are converted into nominal data for experimental purposes.

Performance Metrics

A Confusion Matrix (CM) is used to illustrate the performance of the classifier for a dataset whose class values are known. It clearly shows the confusion between classes, where one class is commonly mislabeled as another. Most performance measures are calculated from the CM. The CM for ART predictive model is given in Table 3, and the performance metrics in Table 4.

Table 3. The Confusion Matrix

	Live Birth – P (1)(Predicted)	*Not Live Birth* – N(0)(Predicted)
Live Birth (Actual) – P (1)	True Positive (TP)	False Negative (FN)
Not Live Birth (Actual) – N(0)	False Positive (FP)	True Negative (TN)

Table 4. Performance Metrics

Measure	Formula	Explanation
Accuracy	$\dfrac{(TP + TN)}{(TP + TN + FP + FN)}$	Ratio of correct predictions to total predictions.
True Positive Rate (TPR) / Recall / Sensitivity	$\dfrac{TP}{TP + FN}$	Measure of completeness or quantity
False Positive Rate (FPR) /(1 - Specificity)	$\dfrac{FP}{FP + TN}$	Measures the false alarm rate
True Negative Rate (Specificity)	$\dfrac{TN}{TN + FP}$	Proportion of actual negatives correctly identified as negatives
Precision / Positive Predictive Value	$\dfrac{TP}{TP + FP}$	Measure of exactness or quality. It means the percentage of the results which are relevant
F-Measure	$\dfrac{2 * \text{Recall} * \text{Precision}}{\text{Recall} + \text{Precision}}$	It is the harmonic average of the precision and recall.
Mean Absolute Error (MAE)	$\dfrac{\sum_{i=1}^{n} abs(y_i - x_i)}{n}$	Measures the average of absolute errors in a set of predictions, where y_i is the prediction and x_i is the true value.
ROC	The Receiver Operating Characteristics (ROC) curve shows the relationship between Sensitivity and (1 - Specificity). The area under the ROC curve is ROC Area or Area Under ROC Curve (AUC).	
PRC	A Precision-Recall Curve (PRC) is a plot of the precision (y-axis) and the recall (x-axis) for different thresholds, much like the ROC curve.	

Even though various metric are measured in the experiment, this research work focus on ROC metric based on the study [23].

A Performance Comparison of the Proposed Model with state-of-art Classifiers used for ART

The performance of the implemented model is checked with the other classifiers which are already used for ART related studies. The performance of the classifiers may be compared accurately only when they are checked with a common experimental setup. Hence all the selected classifiers are applied to the same ART dataset taken for the study. Moreover, all the classifiers are evaluated by 10 fold cross validation invariant of the validation techniques used for the classifier in the corresponding study. The results of the evaluation are shown in Table 5.

Table 5. A Performance Comparison of the Proposed Model with other state-of-art classifiers used for ART.

Classifier	Accuracy	ROC Area	Time (s)	MAE	PRC	F-Measure
PART	88.68	89.4	37.73	0.148	89.6	88.4
Naïve Bayes (NB)	76.24	78.6	0.03	0.25	82.9	76.3
Bayes Net	76.26	78.4	0.83	0.25	83	76.3
1-NN	87.75	90.9	0.01	0.13	91.7	87.0
Ada Boost (Decision Stump)	79.36	68.4	0.72	0.31	74.6	74.2
Random Forest (RF)	88	93.4	0.24	0.14	93.2	87.6
J48	86.69	83.8	1.23	0.19	85.6	86.0
Support Vector Machine	84.02	73.5	8255	0.15	76.1	82.6
Reduced Error Pruning (REP) Tree	85.7	83.7	0.91	0.20	85.5	84.9
LibSVM	80.7	64.4	161.56	0.19	70.7	76.6
Logistic Regression	84.5	86.9	746.43	0.22	89.3	83.4
Multilayer Perceptron (MLP)	76.2	73.5	1605.28	0.23	78.3	72.3
Proposed Model (EHIC)	**85**	**94.1**	**0.2**	**0.15**	**94.1**	**84.7**

The results show that the proposed model outperforms other classifiers used in the literature for the ROC Area metric. Moreover the proposed model has the capability to learn in an incremental manner where most of the existing classifiers lack the property of incremental learning. It is understood that next to the proposed model (EHIC) Random Forest shows optimum performance. The literature surveyed also insists that RF classifier is the best performing classifier for ART [21]. It is promising to see that the proposed model outperform the RF.

A Performance Comparison of Proposed Model with other Incremental Classifiers for ART dataset

The objective of this work is to build a dynamic model for ART outcome prediction. Hence the performance of the proposed model using EHIC is checked with other available incremental classifiers which have the ability to refine it-self on getting new data without forgetting the existing knowledge. The results of the evaluation are shown in Table 6.

Table 6. A Performance Comparison of Proposed Model with other Incremental Classifiers for ART dataset

Classifiers	Performance Metric (%)					
	Accuracy	ROC Area	F-Measure	PRC Area	Time (S) Taken	MAE
Stochastic Gradient Descent (SGD)	77.1	76.1	76.9	70.6	17.17	0.23
Radial Basis Function Network	77.89	77.6	0.87	76.5	0.31	76.7
Stochastic Primal Estimated sub-GrAdientSOlver for SVM (SPegasos)	77.6	75.5	76.9	70.8	24.02	0.2
Naïve Bayes (NB) Updatable	71.4	77.7	71.1	78.2	0.01	0.3
Local Weighted Learning (LWL)	67.06	74.7	60.6	76.0	0.02	0.41
K-Star	82.7	94	82.1	93.9	0.02	0.19
IB1	85.5	92.3	85.3	91.4	0.01	0.15
A1DE Updatable	77.2	84.4	76.2	84.8	0.12	0.26
Proposed Model (EHIC)	**85**	**94.1**	**84.7**	**94.1**	**0.2**	**0.15**

The performance of the proposed model is checked with other incremental classifiers available. The results from Table 6 make it plain that the proposed model outperforms the other incremental classifiers and occupies the top position with the ROC Area of 94.1. It also reinforces the fact that combining more than one classifier improves the prediction performance. Next to the proposed ensemble K-Star classifier shows optimum performance.

A Performance comparison of the proposed model with and without feature selection methods

With the aim to select the influencing feature that impacts the outcome of ART result, Feature Selection (FS) method is applied to the ART dataset. The FS methods chosen include the methods already used in the literature for ART related study as well as some of the promising FS methods performed well for incremental classifiers. The performance of the proposed model is evaluated with and without FS method and the results are in shown in Table 7.

Table 7. A Performance comparison of the proposed ensemble model with and without feature selection methods

FS Method	Evaluator	Search Method	No. of Selected Attributes	Performance Metrics (%)		
				Accuracy	F-Measure	ROC
Without FS	-	-	52	85	84.7	94.1
Filter Method	Correlated Feature Selection	Best First	7	73	71.2	75.3
		Greedy	7	73	71.2	75.3
		PSO	9	74.9	73.7	79.4
	Information Gain	Ranker	13	80.43	79.8	86.6
	Chi-Squared	Ranker	30	83.8	83.4	91.6
Wrapper Method	Proposed Model (EHIC)	Greedy	15	78.2	77.7	85.6
		Best First	15	78.2	77.7	85.6
		Incremental Wrapper Subset Selection (IWSS)	**36**	**85.8**	**85.7**	**94.6**
		IWSS Embedded with NB	18	81.1	80.4	87.8

The results of Table 6 inferred that IWSS method performed well among the other methods. The significant features (predictors) that are selected by most of the FS methods are *patient's age, cervical factor,*

embryos transferred, sperm immunology factor and sperm motility,, Endometriosis, Eggs Micro-injected. It is noted that the performance of the proposed model increased with 0.5 for ROC and the number of features needed to achieve this performance is 36.

Evaluating the performance of proposed model (EHIC) for other dataset

The experiments make it clear that the proposed model performs well for ART dataset.However, in order to check the applicability of the proposed machine learning model to other kind of datasets, it is compared with some benchmark datasets. The description of the dataset used for comparisons are given in Table 8.

Table 8. Description of Datasets.

Dataset Name	Number of Instance	Number of Attributes	Number of Classes
Prisoner	463	31	3
Vote	435	17	2
Breast Cancer	498	10	2
Hypothyroid	3772	30	4
Credit-g	1750	21	2

The performance of the proposed model is checked for various benchmark dataset and also for the dataset downloaded from git-hub website [24] is shown in the Table 9.

Table 9. Performance of proposed model (EHIC) for other dataset

Dataset	Performance Metric (%)					
	Accuracy	ROC Area	F-Measure	PRC Area	Time (s)	Error
Prisoner	97	99.9	97	99.8	0.001	0.02
Vote	93.6	98.7	93.6	98.8	0.001	0.06
Breast Cancer	87.6	94.2	87.6	94.0	0.04	0.13
Hypothyroid	93.8	94.0	92.8	95.9	0.03	0.03
Credit Card	85.9	96.3	85.6	95.7	0.01	0.14

From the results of Table 9, it is understood that the proposed ML model gives optimum performance for other datasets also. In fact, the performance of the proposed ensemble for other dataset is better than that of fertility dataset. The reason for this may be the other dataset taken for comparisons may be less complex than that of fertility dataset and this can be evident from Table 8. From the results it is inferred that, when the complexity of the dataset is low, the performance of the chosen model increases. This experiment also proves that the proposed model will be applied to other domains also.

Computational Complexity of the Proposed Model

After finding the fact that, the proposed model performs well for all the datasets taken for the study, it is imperative to check the computational complexity of the proposedmodel. Hence it is planned to calculate time and space complexity of the ensemble, both during training and classification time. The computational complexity of the proposed model is calculated based on the base learners IB1 and A1DE Updatable.

Table 10. Computational complexity of Algorithms.

Algorithm	Training		Classification		Ref
	Time	Space	Time	Space	
AIDE Updatable	$O(tn^2)$	$O(k(nv^2))$	$O(kn^2)$	$O(k(nv^2))$	[25]
IB1	$O(ts)$	$O(s)$	$O(s)$	$O(s)$	[26,27]
EHIC	$O(tn^2) + O(ts)$	$Max(O(k(nv^2)), O(s))$	$O(kn^2) + O(s)$	$Max(O(k(nv^2)), O(s))$	

$k$is the number of classes
$n$is the number of attributes
$v$is the average number of values for an attribute
$t$is the number of training examples
$s$is the number of instances retained in the subset (Concept Descriptor)

The training time taken for the proposed model depends on the number of training samples taken and the number of attributes. The time taken for classification depends on the number of classes and the number

of attributes. The space complexity of the model depends on the number of training samples, the number of distinct values for each attribute and the number of available classes.

DISCUSSION AND CONCLUSION

This paper discussed a novel approach for generating a dynamic model for predicting the outcome of ART and implemented a major part of the model with the help of an ensemble of heterogeneous incremental classifier and analyzed the performance of the model with the help of an ART dataset. Experiments are carried out with respect to the state-of-art models and achieve the result of 94.1 for ROC area and prove the prediction efficiency of the model.

The strength of this research work is that the proposed dynamic model for ART outcome prediction using machine learning can be used as a clinical decision support tool to physicians as well as to the infertile couples to consider the chances of success before the treatment procedure. Since the proposed dynamic model can integrate the experiences of all experts and the history of treatments into a single computational tool, learns from the past cases, and analyzes several patient records, it can make predictions in minimum time with less subjectivity, human bias, and with higher precision. The results of prediction by the proposed model could help to make more accurate decisions by the physicians and minimize the current challenges of manual observation and repeated attempts in ART treatments. The limitation of this work is the implemented ML model using EHIC is working well for nominal data. Its performance may vary for continuous data. The proposed model achieved the specified performance for SMOTE balanced data. It may vary for any other balancing techniques. The EHIC will have the incremental property only when the incremental classifiers are combined together.

In future for handling imbalance dataset other techniques like Cost Sensitive Learning that highlights the imbalanced learning problem by using cost matrix method which describes the cost for misclassification in a particular scenario may be employed. Moreover, the causes of infertility factor are considered as a binary variable in this research. If continuous or fuzzy values of the factor are available, fuzzy sets can be incorporated in this factor to find the accurate level of intensity of the cause which in turn may improve the efficiency of the model and also helps to give prescription.

REFERENCES

1. Tyagi, Infertility is rising in India's populace [Internet]; 2019 [cited 2020 Aug 8]Available from: https://www.deccanerald.com/opinion/panorama/infertility-is-rising-in-india-s-populace-788978.html.
2. Patil AS. A review of soft computing used in assisted reproductive techniques (ART). Int J Eng Trends Appl (IJETA). 2015;2(3):88-93.
3. WHO , Infertility [Internet]; 2020 [cited 2020 Nov 12] Available from: https://www.who.int/news-room/fact-sheets/detail/infertility
4. Sharma, Kalpana, 27.5 million couples in India suffering from infertility [Internet]; 2018 [cited 2019 April 4] Available from:https://timesofindia. indiatimes.com/life-style/parenting/getting-pregnant/27-5-million-couples-in-india-suffering-from-infertility/articleshow/63938393.cms
5. Donna Rosato, How High-Tech Baby Making Fuels the Infertility Market Boom [Internet]; 2014 [cited 2018 Oct. 8] Available from: http://time.com/money /2955345/high-tech-baby-making-is-fueling-a-market-boom/,
6. What is Assisted Reproductive Technology (ART) [Internet]; 2020 [cited 2020 Nov 12]Available from: https://www.cdc.gov/art/whatis.html.
7. Greil AL, Slauson K. Blevins, and J. McQuillan. 2010. The experience of infertility: A review of recent literature. Sociology of Health & Illness.;32:140-62.
8. Zarinara A, Zeraati H, Kamali K, Mohammad K, Shahnazari P, Akhondi MM. Models predicting success of infertility treatment: a systematic review. J Reprod Infertil. 2016 Apr;17(2):68.
9. Fertility Guideline (Second Draft For Consultation, August 2003), National Collaborating Centre for Women's and Children's Health, 2003
10. Wendy Kuohung, Mark D Hornstein, Evaluation of female infertility – UpToDate [Internet]; 2018 [cited 2019 Apr. 22]. Available from : www.uptodate.com/contents/evaluation-of-female-infertility.
11. Ranjini K, Suruliandi A, Raja SP. An Ensemble of Heterogeneous Incremental Classifiers for Assisted Reproductive Technology Outcome Prediction. IEEE Trans. Comput. Soc. Syst. 2020 Nov 3.
12. Chen CC, Hsu CC, Cheng YC, Li ST. Knowledge discovery on in vitro fertilization clinical data using particle swarm optimization. In 2009 Ninth IEEE International Conference on Bioinformatics and BioEngineering 2009 Jun 22, pp. 278-83.

13. Guh RS, Wu TC, Weng SP. Integrating genetic algorithm and decision tree learning for assistance in predicting in vitro fertilization outcomes. Expert. Syst. Appl. 2011 Apr 1;38(4):4437-49.

14. Durairaj M, Thamilselvan P. Applications of artificial neural network for IVF data analysis and prediction. J. Eng. Appl. Sci. 2013 Sep;2(9):11-5.

15. Milewski R, Milewska AJ, Więsak T, Morgan A. Comparison of artificial neural networks and logistic regression analysis in pregnancy prediction using the in vitro fertilization treatment.Stud. Log. Gramm. Rhetor. 2013 Dec 1;35(1):39-48.

16. Güvenir HA, Misirli G, Dilbaz S, Ozdegirmenci O, Demir B, Dilbaz B. Estimating the chance of success in IVF treatment using a ranking algorithm. Med Biol Eng Comput,. 2015 Sep;53(9):911-20.

17. Ramasamy N. Feature Reduction by Improvised Hybrid Algorithm for Predicting the IVF Success Rate. Int**.** J. Adv. Res. in Comput. Sci. 2017 Jan 1;8(1).

18. Raef B, Maleki M and Ferdousi R. Computational prediction of implantation outcome after embryo transfer. Health Informatics J., Sage Journals, 2020; 26(3):1810-26

19. Hafiz P, Nematollahi M, Boostani R, Jahromi BN. Predicting implantation outcome of in vitro fertilization and intracytoplasmic sperm injection using data mining techniques. Int J Fertil Steril. 2017 Oct;11(3):184.

20. Hassan MR, Al-Insaif S, Hossain MI, Kamruzzaman J. A machine learning approach for prediction of pregnancy outcome following IVF treatment. Neural comput. and appl.. 2020 Apr;32(7):2283-97.

21. RanjiniK, Suruliandi A and Raja SP. Machine Learning Techniques for Assisted Reproductive Technology: A Review J. Circuits, Syst. Comput, 2020; 29(8):2030010.

22. Human Fertilization and Embryo Authority - [Cited: Aug 2, 2018] Available from :https://www.hfea.gov.uk,

23. SuruliandiA, Ranjini K and Raja SP, Balancing Assisted Reproductive Technology Dataset for Improving the Efficiency of Incremental Classifiers and Feature Selection Techniques. *J. Circuits, Syst. Comput, 2020; DOI:* https://doi.org/10.1142/S021812 662130 0075

24. Crime Propensity Prediction Data Set - [Cited: Apr. 11, 2020] Available from: https://github.com/Benjamindavid03/CrimePropensityPrediction

25. WebbG Boughton and Wang. Not so naive Bayes: aggregating one-dependence estimators.Mach. Learn.2005;58(1): 5-24.

26. WilsonDR and MartinezTR.Reduction techniques for instance-based learning algorithms. Mach. Learn.2000; 38(3): 257-86.

27. AhaDW, KiblerD and Albert MK. Instance-based learning algorithms.Mach. Learn.,1991; 6(1):37-66.