Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 1, 2021, Pages. 5926 - 5943
Received 15 December 2020; Accepted 05 January 2021.

A Comprehensive and Systematic Literature Review of Computational Intelligence Algorithms to Diagnose and Predict Female Infertility

T. Satya Kiranmai, Department of Information Technology, Chaitanya Bharathi Institute of Technology (A), Hyderabad, India**.**

P.V.Lakshmi, Professor, Department of CSE GITAM (Deemed to be University) Visakhapatnam, India.

Abstract*:*

Recent researches inferred that one out of ten couple are facing fertility problem and it is observed that most of the cases address female infertility as a common health disorder due to the effect of several genetical, environmental and psychological factors. Efficient medical practices like IVF (In-vitro fertilization) are widely available in treating infertility problems in couples. As IVF procedure is influenced by several clinical attributes, predicting the rate of success is considered as prime concern. Apart from that, diagnosing ovulation disorders caused due to hormonal imbalances is considered as a challenging aspect. Technology assisted medication plays a vital role in addressing such problems. This article presents the comprehensive review of existing research studies that include a vast range of computational intelligence and machine learning algorithms addressing the problem of infertility in women. In addition, this article comprehensively illustrates a systematic analysis of existing research studies to identify the research gaps and challenging issues in devising computational intelligence algorithms for diagnosing and treating the infertility in women.

Keywords: Female infertility, PCOS, IVF, Machine learning, Computational intelligence.

1. Introduction

Now a day's, health care and services are being automated with the advent of artificial intelligence and machine learning while diagnosing the hidden features of diseases. Machine learning may be used in medical applications wherever necessary which includes the robotic surgeries, disease diagnosis and automated patient treatment. Research on diagnosing infertility has drawn the interest of several multidisciplinary researchers' recent years. Current generations of mankind evidenced infertility as a significant problem among couples due to their stressful daily routines. Infertility is universally defined as the failure of conceiving pregnancy even after 12 months of unprotected sexual activity [1]. Several factors affect the fertility in women, the major among them includes uncertain and infrequent ovulation, hormonal imbalances that effect the process of ovulation which in general called as polycystic ovary syndrome (PCOS) and pelvic inflammatory disease (PID) a sexually transmitted disease that effects the endometrium layer. Furthermore, the most common factor is additional growth of a tissue on the endometrium layer.

Most of the cases the diagnosis for the problems of infertility will be identified using ultrasound imaging in this context if we devise an algorithm or decision-making framework that could be utilized while making a decision in the treatment options that could be more appropriate

for the infertile couple that could be an added advantage for the medical professionals in enhancing heir fertility rates. In such cases the clinical data that is crewed as input to such computational intelligence algorithm may include the demographic data, images of ultrasound scans in addition to the decisions related to the treatment options made by different medial professional to different patients. Machines that use these kinds of algorithms may include the parametric values that are computed based on the prior information obtained from the various medical data sources that include the historical data of infertile couples succeeded in conceiving. It is inferred form various medical sources that there are several techniques for ovulation induction that are imparted to the computational intelligent algorithms during the decision-making process.

Markov chain-based probability transition matrices are considered as one of the prime methods to obtain and analyse the parametric values for such computational intelligence algorithms [2]. Application of Markov chain model in the context of infertility related data could be found in [3] besides that the application of stochastic process in the bio-medical applications are sourced from [4]. Diagnostic machines that handle this kind of algorithms along with diverse parametric values results in the enhancement of treating options. Machine learning enthused computational intelligence algorithms will enhance the chances of conceiving on computing the probabilities that include the characteristics at patient level when combined with the treatment options.

The ability to conceive depends on the duration of the sexual exposure, coitus frequency and the age of the pair. The typical young, old couples have 25% probability of getting unprotected relationships after 1 month; 70% of the couples have six months' conceivability and 90% of the couples have a possibility of 1 year's conceivability. After one and a half year or two years, about 5% of couples can conceive [5], [6]. The triggers are similarly responsible for both men and women. Most infertile couples have one of these three primary causes – male, ovulatory or tubal peritoneal. The literature suggests that vaginism and dyspareunia are more prominent in women aged 20-24 years[7]. The sexual interaction cycle plays an important role in promoting fertility as it includes sequential physical and emotional changes that occur when a person is sexually excited[8]. In normal physiology, the two gonadotropin hormones, follicle stimulating hormone (FSH) and luteinizing hormone (LH), are formed in a hypo physiological environment and regulated by their release by a gonadotropin hypothalamus-releasing (GnRH). The hypothalamus starts at the beginning of a new cycle to release GnRH, which works on the pituitary gland to release FSH and LH. Both hormones activate the ovary and produce the follicles. Around 30-40 follicles begin to develop monthly in response to FSH, but only one mature egg is released per month. This includes the communication of the message in the form of ovarian, pituitary and hypothalamic hormones. The mature follicle produces a growing amount of oestrogen released by the granulosa cells in the follicle when the Egg is ripe. The estrogen emitted by the dominant follicle rises rapidly with the development of the egg until the oestrogen is released into the blood. The elevated level of oestrogen induces the hypophysis to release a significant volume of LH, resulting in the LH wave. This LH works on the mature follicle, allowing it to break free from the mature egg (ovulation).

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 1, 2021, Pages. 5926 - 5943
Received 15 December 2020; Accepted 05 January 2021.

1.1 Causes for Infertility in Women

The causes of female infertility can be classified into three large categories according to the Centre of Disease Control (CDC 2013)[9], including faulty ovulation, transport and implantations. These types are described in detail below.

Deficient ovulation: It is due to endocrine disorders. Hypothalamus and pituitary gland disfunction can contribute to the secretion of excess prolactin, preventing ovulation. Additional endocrine glands, including thyroid and adrenals, can also delay ovulation. If the corpus luteum does not contain enough progesterone to thicken the uterine lining, the fertilized egg may not implant, which results in infertility.

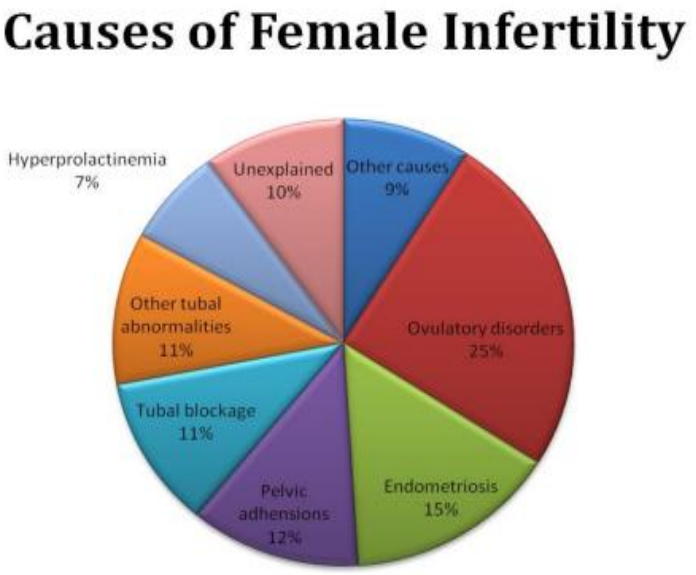

Figure 1.1 Causes for infertility in female

Physical disorders: Certain physical disorders include obesity, anorexia nervosa and excessive exercise, and later the menstrual cycle can contribute to overweight or malnutrition, such that the couples are infertile.

Ovarian disease: Polycystic ovarian disease (PCO) may result in fertility due to increased testosterone and LH levels and decreased glucose intake by the pancreas by muscle, fat and liver cells. Low FSH levels also inhibit the development of ovarian follicle eggs, resulting in ovarian ovarian cysts filled with fluid that ultimately cover the entire ovaries and prevent pregnancy. The ovarian disease as seen in Figure 1.2.

Endometriosis: It is a condition in which the tissue that forms half of the uterus is located in other regions such as ovaries, fallopian tubes, uterine tissue and pelvic cavity[10]. Cysts form and ovulation happens as it affects ovaries. It may also form scars that block Falopian tubes or interfere with fertilized egg implantation. It also includes painful and extreme cycles of ovulation pain.

Blocking of the tubes [11]: Fallopian tubes are the location where the egg fertilizes; thus they should be safe and accessible. The sperm does not enter the egg to fertilize it when the

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 1, 2021, Pages. 5926 - 5943
Received 15 December 2020; Accepted 05 January 2021.

tubes are blocked or damaged[12]. Anatomical defects, pervasive surgery in your womb or in the tubes and infections related to sexually transmitted disorders, especially chlamydia or gonorrhoea, may be the cause of tubal blockage.

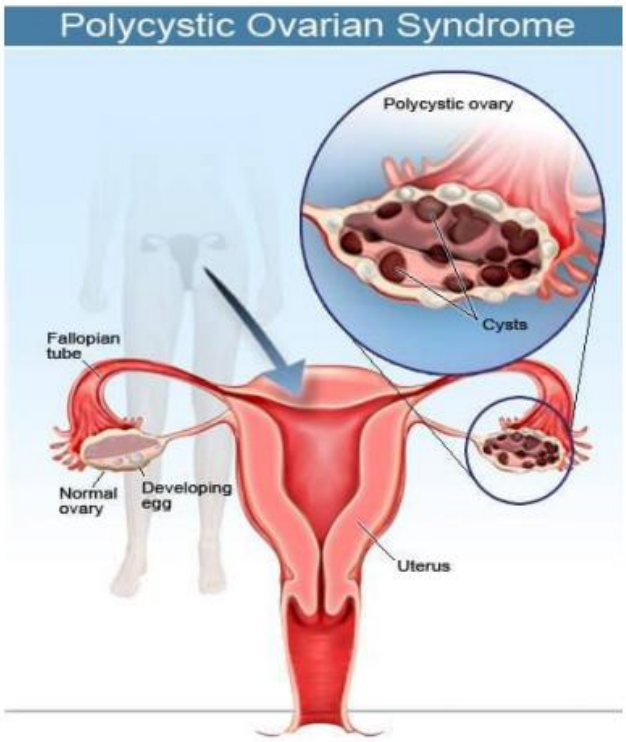

Figure 1.2: PCOS – Ovarian disorder

Ovum: Pelvic inflammatory disease (PID), gonorrhoea, peritonitis, earlier surgery on the tubes, and tubular adhesions can lead to a pause in the concepts; as a result, eggs are not released or stuck and hence the pregnancy is retarded.

Scar tissue after abdominal operation: The presence of the abdominal tissue can change the movement of the ovaries, fallopian tubes, and uterus after abdominal intervention, leading to infertility. Sperm: Presence of vaginism or dyspareunia may avoid fertilization and make the couple infertile.

Cervix: Trauma, surgery, cancer, cervical mucus anti-sperm antibodies may even delay birth.

Defective implantation: Congenital anomalies and fibroids, congenital uterine defects, such as bicornuate uterus and cervix fibroids, may change the implantation of zygota and trigger infertility. We address the causes of female infertility. Examinations of the causes of male infertility are further discussed.

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 1, 2021, Pages. 5926 - 5943
Received 15 December 2020; Accepted 05 January 2021.

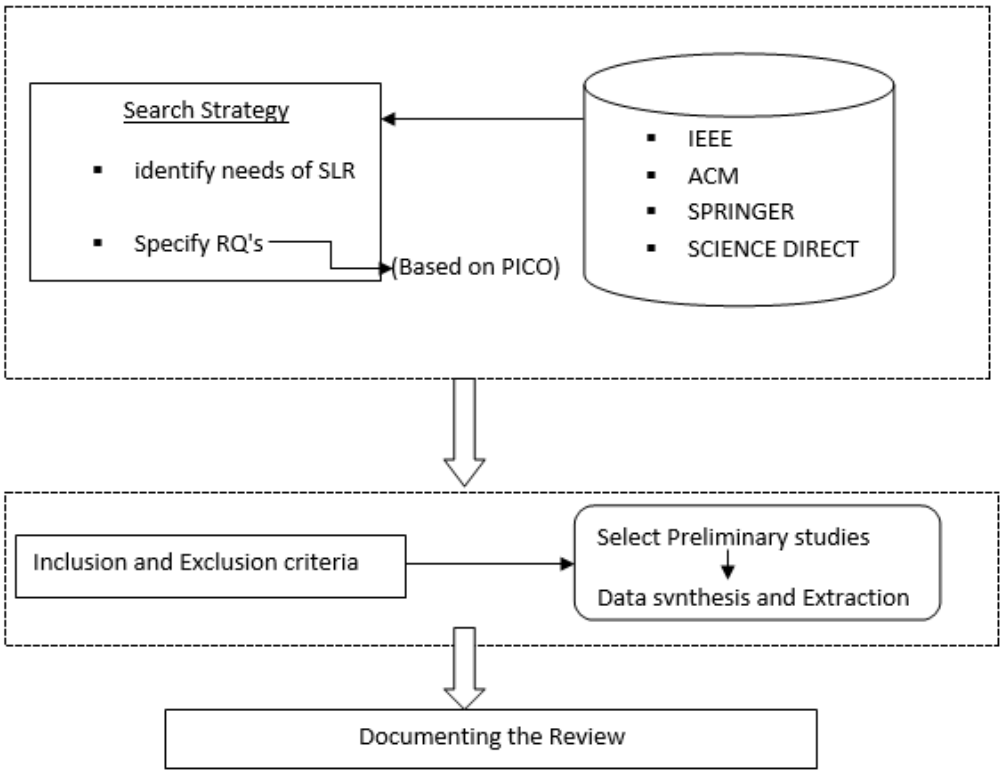

Figure 2. SLR (Systematic literature review Protocol)

The main objective of this SLR is to enable the next generation researchers with the research gaps and the scope of research in medical data analysis in specific to female infertility. Despite of several published SLRs addressing various challenges of medical diagnosis in technological perspective [14], It is observed that there is need of well interpreted SLR addressing various computational intelligence algorithms deployed in analysing the causes of infertility. The review protocol is defined based on three phase architecture in which initially Research Questions(RQs) are developed based on the systematic search carried by incorporated specific keywords related to the female infertility over indexed journal data bases to identify the related studies in the context of the efficient algorithms and models deployed for predicting and diagnosing the reasons for female infertility. Selection of preliminary studies related to the female infertility is made in the second phase of review protocol based on defined inclusion and exclusion criteria. In the third phase the detailed review is documented and interpreted to identify the future scope of research.

2. Research Methodology

2.1 Search strategy

The search strategy is formulated based on the application of defined set of keywords related to computational intelligence algorithms to assess infertility of female over the indexed databases like IEEE, ACM, SPINGER and SCIENCE DIRECT etc., the combination of search strings utilized for preliminary article search is depicted in table1:

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 1, 2021, Pages. 5926 - 5943
Received 15 December 2020; Accepted 05 January 2021.

Table:1 Combination of search strings to identify relevant articles from scientific databases

Female infertility**OR**Endometrial **issues OR** PCOS**OR** Fallopian tubes**OR**IVF
AND
Ultrasound Scans**OR** Genetic information**OR**Demographic data**OR** Clinical data
AND
Computational intelligence algorithms **OR** heuristic algorithms **OR** machine learning techniques**OR**Data Mining techniques
AND
. Algorithm Specific Methods **OR** Data specific method **OR** Hybrid Mechanisms
AND
Systematic Study **OR** SLR **OR** Mapping Study **OR** Review

In the initial cases based on above search strings 130 relevant research papers addressing the domain of female infertility were identified within a range of a decade (2009-2020) directly from scientific databases the dissemination of the articles over various databases is depicted in table:2

Table: 2 Dissemination of identified articles over various scientific databases

S.No	Database	No. Of Papers
1	IEEE	55
2	ACM	22
3	SPRINGER	29
4	SCIENCE DIRECT	24
Total		130

2.2 Defining Research Questions

Formulating research questions is considered to be a Vitol aspect in the context of interpreting an SLR as it includes the motivation and need for conducting a systematic review in the context of female infertility. With the motivation gained by [15] PICO strategy is adopted while designing a well-built research question in the context of yielding the high-level evidence.

Table 3: Formulation of Research Questions

Acronym	Definition	Motivation	Research question
P	**Problem**	Gain insights and attain knowledge in the context and causes for female uncertainty	**RQ1:**Identify different causes and treatment options in the context of infertile women.
I	**Intervention**	Analyse state of art mechanisms that addressing computational	**RQ2:** What are various existing mechanisms and algorithms that

		intelligence algorithms for female infertility	are devised while addressing female infertility.
C	**Comparison**	Compare various Algorithms and techniques available	**RQ3:** Illustrate the performance analyses of various existing algorithms that enhances of accuracy of prediction during infertility of women
O	**Outcome**	Understand the necessity of developing automated algorithms addressing the solutions to predict female infertility	**RQ4:** What is the evidence to support the scope of research in the design and development of efficient algorithms

2.3 Preliminary selection

In the preliminary phase of the selection 130 articles are identified from the scientific databases using the search strings indicated in table: 1 based on which the initial scrutiny of the articles is performed by analysing the relevancy of the article title towards the problem statement. In the second stage of the scrutiny the articles filtered based on the inclusion and exclusion criteria framed as shown in table: 4

Table 4: Inclusion/Exclusion criteria

Inclusion criteria	Exclusion criteria
Articles that include Algorithms centric methods and mathematical assertions designed to devise solutions for enhancing the prediction accuracy of infertility in women	Articles with an ambiguity in the context of the implementation tools and data sets
Articles developed based on the evidential research that are formulated that well defined implementation details and simulation results with the inclusion of tools and databases required for research in diagnosing female infertility	White papers and lecture notes regarding published in the context of architectural perspective of diagnosing female infertility
Articles that are primarily implemented in the bio-computer science domain	Articles that are written in other than the English language
Articles that are written in the English language	

In the second phase of selection 26 research articles that are abided by the inclusion and exclusion criteria are selected for documenting the review where it to be noted that these articles are furnished with required information that enhance the scope of research in that

highlights the solutions for diagnosing, predicting and recommend treatment options for infertility in women.

3. Review of Computational intelligence algorithms addressing infertility in women

3.1 Research on intelligent algorithms to enhance prediction accuracy of IVF

The early prediction of the result of an IVF procedure is relevant for medical practitioners as well as for their patients for determining potential outcomes. Many experiments in literature use machine learning methods for the estimation of IVF pregnancy. Kaufmann et al. (1997) [16] used ANN as one of the earlier researches. First, they used T-test to classify four attributes (patient age, number of recovered eggs, number of transferred embryos and freezing of embryos), which they felt were most important and used these as the inputs to train a neural network and achieve a 59 per cent accuracy in the test data collection

Jurisica et al. (1998) [17] introduced a case-based rationalization system which took advantage of prior practice to recommend potential improvements to an IVF treatment plan to increase future overall success rates. In order to predict the result of embryo implantation in IVF, Uyar et al. (2015) [18] tested various classifier forms utilizing 18 attributes in total: 9 related to the patient's characteristics and 9 related to the embryo. The classificatory tested included the following: SVM, DT, naive bays (NB), nearest neighbours (KNN), multilayer perceptron (MLP) and radial network feature (RBF). In order to test the efficiency of the classifiers, they used the Receiving Operational Characteristics (ROC) and obtained stronger outcomes, utilizing NB and RBF classifiers that achieved the 0.739 (±0.036) and 0.712 (±0.036) areas under the ROC curve (AUC). In a later analysis the authors decreased the characteristic range to 11 in line with their relative value, which contributed to marginally improved results with NB achieving 80,4 percent precision and AUC 0,756 (±0,036).

The authors analyzed ten characteristics and experimented with three separate classifications, namely SVM, ANN and DT, in (Nanni et al., 2010) [19] and their versions. It used the sequential forward floating selection method (SFFS) for the selection of features and concluded that the ensemble approach was predicted more exactly based on a random subspace of the DTs with just three features (patient age, sub endometrial volume and endometrial vascularization / flow index). Their best AUC model was 0.85; nevertheless, their analysis was focused on a dataset with comparatively few treatment cycles. Durairaj and Thamilselvan (2010) used ANN to forecast IVF performance. A dataset of 14 attributes was compiled, eight of which were chosen to train the algorithm, and a cumulative accuracy of 73 percent was achieved. The study was further developed later on by creating an applied Rough Set Theory (RST) and ANN data mining method utilizing a broader dataset of additional attributes. RST was used to purify and minimize results, resulting in 9 influential attributes among the 27 initial data set attributes. The RST-based attributes were then used to train an ANN classification that could accurately predict the outcomes of pregnancy in 90% of cases.

Milewski et al. (2013) [20] contrasted the prediction capabilities of both Multivariable Logistic Regression (MLR) and Multilayer Perceptron (MLP) models. Their research found that, although the measurement of MLR was more fitting for theoretical purposes, the

description of MLP was more effective in clinical prediction. The MLR calculation was done by an MLP classification of 26 inputs, three layers of a 40-6-2 secret structure and one output to estimate IVF outcomes by a margin of 0.061.

In [21] proposed to predict the performance of a procedure using a rating-based algorithm a modern approach known as the Success Estimates using Ranking Algorithm (SERA). RIMARC is the specific algorithm used in their analysis (Ranking Instances by Maximizing the Area under ROC Curve). They evaluated three models, namely SERA, NB and RF, where 84.4% and 0.833 percent of SERA performed other models with precision. Hafiz et al. (2017) [22] checked the effects of IVF using a dataset comprising 29 attributes using five techniques, including SVM, RF, recursive partitioning (RPART), adaptive boosting and nearest neighbour, and stated that the RF and RPART were outperforming other methods achieving AUC of 0.8423 and 0.8205 and respectively.

Ramasamy et al. (2017) [23] suggested, in a recent review, a hybrid algorithm incorporating Ant Colony and Relative Reduct, known as an ACRR algorithm, to pre-process the IVF data, reducing the number of attributes from 41 to 18. Applying ACRR in combination with classifiers of NB, MLP, RBF and J48 showed that NB worked better with an accuracy of 90.35%. The authors suggested in (Corani et al., 2013) [24] a Bayesian embryo model network model, assuming that a uterus and a workable embryo are required for a healthy pregnancy. Their models were represented by two E (embryo) and U (uterine) variables respectively in the probabilistic calculation, and the Bayesian network consisted of a guided diagram that links the nodes representing these variables. The network parameters were calculated using the EM procedure, which maximizes the posterior likelihood of the parameters (i.e., probability of the parameters given some observed data). They concluded that their parameter estimation approach produced a stronger AUC (0.834) than the conventional MAP calculation (most definitely a posteriori). Their inference was focused on a limited dataset, though.

In (Choi et al., 2013) [25], the authors suggested a tree-based statistical model that is tailored to the clinical evidence for a patient's first progress in IVF therapy. The model was created and validated using data of several clinics without having a particular mandatory format or collection of clinical protocols for each clinic.

Morales et al. (2008) [26] recommended a Bayesian classifier with 20 features to help identify the most successful embryos to be transferred to the female womb. Furthermore, Guh et al. (2011) [27] suggested a hybrid approach which incorporates both GA and DT techniques. The DT obtained a sensitivity and specificities of 21.5 and 95.0 percent using all 67 IVF attributes included in the analysis using 5 times cross-validation. The GA algorithm was used to concurrently determine the best set of characteristics and learning parameters for the DT model. This GA and DT integration decreased the amount of essential features to 28 (e.g. age of patients, the number of embryos implanted, the number of frozen embryos and the days of culture of the embryo). Multiple DT models were produced from which the top ten models were chosen. The best DT model had a 73.2% predictive precision with 71.6% and 73.9% respectively of the associated sensitivity and specificity.

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 1, 2021, Pages. 5926 - 5943
Received 15 December 2020; Accepted 05 January 2021.

In (Manna et al. 2013) [28] the authors also employed a methodology for the consistency evaluation and rating of embryos or oocytes for IVF care based on a neural network picture study. The protocol was evaluated on 269 oocytes and 269 identical embryos from 104 women and a contrast with the previously proposed TRACE algorithm was made (Total Recognition by Adaptive Classification Experiments). Although previous studies have shown an AUC of 0.83 obtained from neural networks, further research is required with larger datasets and comparisons with other methods of machine learning. While literary works indicate that machine learning strategies are promising to drive IVF therapy plans, their efficiency needs to be strengthened to render medical practitioners welcomed. Table 5 summarizes significant past and recent studies which attempted to apply various techniques of machine learning to predict the outcome of IVF. As the table demonstrates, current works remain unreliable such that the techniques established can be accurately utilized by the practitioners. Further analysis is also required to increase the precision of prediction to an appropriate degree.

Table 5. A Glimpse of research on enhancing the prediction accuracy of IVF using machine learning techniques

Author & Year	Technique utilized	Feature selection/ extraction technique used	Features selected	Performance analysis	Validation
Durairaj& Ramasamy, 2017[23]	Naïve Bayes, Multilayer Perceptron (MLP), RBF, J48	A total of 18 features was selected from 41 features after applying Ant Colonized Relative Reduct Algorithm (ACRRA)	18 attributes	The best accuracy was achieved by MLP, which was 90.35%	Not indicated
Hafiz et al., 2017 [22]	Support vector machines (SVM), recursive partitioning (RPART), random forest (RF), Adaptive boosting (Adaboost), and 1-nearest neighbor are used	No feature selection	-	RF and RPART produced the best results in this study. AUC attained is 0.8423 and 0.8205, respectively	5-fold cross validation
Uyar et al., 2015 [18]	Six classification techniques were used: NB, k-nearest neighbor, DT, support vector machines, ANN, and RBF. Best accuracy was	Input characteristics were assessed using information gain feature weighting and forward feature selection methods	18 attributes	Accuracy 80.4% by NB	Two-third of the data was used for training and one-third for testing. Same procedure was repeated 10 times using random

	achieved by NB.				split on data, each time maintaining the same training/test ratio.
Güvenir et al., 2015 [29]	Success estimation using a ranking algorithm (SERA) which is RIMARC (ranking instances by maximizing the area under ROC curve)	No feature selection	-	AUC – 0.833; Accuracy – 84.4%	10- fold crossvalidation
Guh et al., 2011[27]	Integrating genetic algorithm (GA) and DT techniques	Expert domain is incorporated to select the 38 attributes that are likely to influence the IVF outcome - chosen from the original 69 attributes.	38 attributes	Accuracy 73.2%	Five-fold cross validation
Uyar et al., 2010 [30]	Naive Bayes (NB) and Radial Basis Function (RBF) provided the better results among six classifiers	No feature selection techniques was used	Nine features related to patient and 9 features related to embryo were used	AUC (Area Under ROC) attained by NB and RBF are : 0.739 (± 0.036) and 0.712 (±0.036), respectively	Randomly train-test dataset were selected and repeated the process 10 times

3.2 Research on clinical data Pre-processing, classification and prediction using data mining techniques

The processing of infertility clinical data was suggested by a feature selection mechanism focused on ant colony optimization algorithm [31]. The optimum data set was identified after a number of iterations of PSO applications and filter-based process. Vector Support Machine (SVM) has been used to verify this method's output. The findings indicate that the properties obtained with ACO are greater than other projections of cancer. The specification of the proper combination of the pre-processing system has a major effect on the data set classification accuracy. Ant Miner was used for pre-processing consisting of the following steps, including elimination of instances of missed values, discretion and collection of attributes. Various sub-set approaches were used such as the evaluation of relief attributes, the selection of a sub-set function dependent on correlation, the assessments of accuracy sub-set, the measurement of chi-Squared attributes; the evaluation of benefit ration attributes; the evaluation of OneR attributes; and the evaluation of Symmetrical ambiguity attributes. For assessment, the 10-fold cross validation was used. This approach was used to test a total of 25 medical data sets and the findings were positive.

The HD diagnostic method of Fuzzy Logic[32] is focused on criteria such as cholesterol, blood pressure, diabetes, sex and age. Experimental findings demonstrated 92 percent precision. Diabetes disorder was identified using the Probabilistic Artificial Neural Network [64]. The accuracy in teaching and research attained for this model was 89.56% and 81.49%, respectively

A cluster model that incorporates wavelet transformation and Fuzzy Logic interval type-2 [33] was suggested. These modules are integrated to cope better with high dimensionality and insecurity. Fuzzy C-Means Clustering consists of unmonitored learning and genetic algorithm dependent on tuning parameters interval type-2. These logical machine modules have large computing costs and convert wavelets to reduce these computational costs. The classification model proposed achieves a precision of 97.88 per cent for the detection of breast cancer.

A Fuzzy-Rough classification model for collection of features and Multitree Genetic Programme, for the identification of intension trends with brain signal was developed[34]. In order to overcome the complexities and technical difficulties, it was suggested to combine the fuzzy Regular Additive Model (SAM) with a genetic algorithm (GA) called GSAM. It consisted of three continuous steps, for example rule initialization, enhancement of rules and tuning of parameters.

The findings for embryo implantation were tested by six separate approaches such as the Naïve Bayes, the Closest Neighbour, the Decision Tree and Help Vector Machines, the Multi-Layer Perceptron and the Radial Base Function Network [35]. Naïve Bayes and the Radial Base Feature Network have been found to operate best for all six. A hybrid [36] framework has been applied to predict infertility care based on the Rough Collection and Artificial Neural Network. The principle of rough set was extended to locate the reduction set. It was used to minimize the number of variables used as an input to the neural network as a pre-processing tool. This method was successful for medical data of large and medium scale.

In addition with Support Vector Machine, an improved ACO algorithm was proposed to implement the latest data classification (ERURACO-SVM) [37] process. The efficiency of the approximate ACO algorithm has been enhanced by the implementation of pheromone evaporation rate strategy and pheromone upgrade laws. Parallelism, global optimization, positive feedback system and high robustness of improved ACO algorithm have strengthened the learning efficiency and generalizing capabilities of the SVM model. Experimental studies have shown that the ERURACO-SVM is more reliable and more universal.

The Back Propagation Neural Network (RS-BPNN) Rough Range indiscernibility [38] approach is suggested. It is separated into two steps. The first move is to use the indiscernibility relation approach to treat the missing values. In the second level, the selected data sets are categorized by BNN. The system has been tested using UCI hepatitis, breast cancer of Wisersia and Statlog HD info. The precision was 94.3%, 98.6% and 90.4% respectively for hepatitis, breast cancer and HD. For the diagnosis of heart failure genetic

algorithm and Fuzzy Logic have been used. The writers proposed that clinicians should be able to utilize this system to detect heart failure. Genetic algorithm is used to determine the best solution. For Fuzzification the Gaussian membership function and Centroid approach are used to defuzzify the output value. Output metrics such as precision, sensitivity, specificity and confusion matrix were used for the evaluation of device performance. The precision reached was 86%. Compared to current work, the precision was improved by 1.54%.

A hybrid intelligent classification model [39] was proposed for medical details. The model includes the neural network Fuzzy Min-Max, the classification and regression tree and the Random Forest algorithm. The Fuzzy Min–Max Neural Network is responsible for gradual testing, the classification and regression tree increases the comprehension and the predictive output is enhanced using the Random Forest algorithm. The WBCD data collection is one of the medical datasets utilized for diagnostic assessments. They registered an accuracy of 98,84 percent for the diagnosis of breast cancer.

Megha R, et al [40] suggested a hybrid disease prevention strategy that merged REP bagging with help vector machine bagging. The dataset was initially cleaned and the MRMR Feature Selection Algorithm was used to retrieve features. The SVM description was then introduced and the findings analysed. If the findings are positive, the data set will be checked or the data set will be identified again by means of the bagging classifier. Rajwant Kaur et al [41] suggested a hybrid modelling strategy based on risk factors for heart failure. Genetic Algorithm (GA) and SVM Classifier were used. The hybrid methodology employed GA's global optimization benefit to initialize SVM weights. The framework was based on MATLAB and precision was 95%.

3.3 Research studies addressing the problem of PCOS

Of the various issues around us, the topic relating to women's reproductive health was chosen as a field of concern because of its relevance. A comprehensive survey of PCOS studies and diagnostic systems was performed. Literature claims that roughly 10% of Indian womenThe multifaceted endocrine condition called Polycystic Ovary Syndrome (PCOS) is impaired by reproductive age[16]. It is an important source of anovulatory infertility and increases insulin tolerance, obesity, cardiovascular and psychosocial conditions risk[42]. The symptoms of PCOS can differ between patients. Some of themmenstrual disorders, acne, overweight, elevated infertility, extreme hair loss, face front balding, increased hair development [43].

PCOS will historically be believed if more follicles are found in an ovary than twelve per unit area and radiologically visible[15]. Some scholars recommended changing the cut from 12follicles to 20 or abandoning ultrasound in favour of others.Biomarkers, such as anti-muller hormone serum (AMH). The detection of PCOS is simple and involves only a few well-standardized diagnostic procedures. Diagnosis of PCOS is sometimes postponed, which impacts the well-being of patients adversely. Escobar et al. [45] propose that care is symptomatic, long-term and complex, and can be tailored to evolving conditions and the particular patient's preferences and aspirations. Joham et al. in[46] found the PCOS-infertility-rate relationship among people in this population, and the usage of fertility

hormone therapy in women who record PCOS has been considerably higher. Taking into consideration the prevalence of PCOS and theHealth and economic burden of infertility, PCOS diagnostic techniques and fertility factors are significant. It is because the miscarriage is 15- fold higher in female PCOS reporters, irrespective of BMI.

There is a two-way correlation between obesity and PCOS. Both aggravate each other in a cyclical fashion that never stops. Essah, P.A. Essah, P.A. And J.E's Nestler reports that obesity prevalence is 30 to 75 percent in PCOS women [47]. Clinical PCOS validation is typically conducted by Rotterdam guidelines [48] or by PCOS business requirements. A cross-sectional research by Brower et al.[49] revealed that menstrual mismanagement may be used to predict the existence and likely magnitude of insulin resistance in patients with PCOS. Many scientific studies conducted in the PCOS diagnosis utilizing ultrasound scan characteristics and picture analysis techniques for PCOS diagnosis. Any reports have utilized psychiatric and metabolic disorder characteristics.

4. Observations made from the Literature review

PCOS is the most prevalent form of endocrine condition in women of reproductive age. Which may contribute to anovulation and infertility. The diagnosis criteria require biomarkers for the condition, including clinical and metabolic parameters. Among the different algorithms used, the efficiency of RF algorithms is superior. This automated algorithm will allow the doctor to save substantial time in testing patients and thus reduce the time required to diagnose PCOS.Implications for this work from the scientific expert survey show that advances in medical data are still challenging in the medical and healthcare sectors. Further it is encouraged to analyse and devise valuable novel methodologies such as vitamin D effects on PCOS, experiments that show the impact of PCOS on premature working/abortions, attempts to detect the amount of lean PCOS patients, etc., needs to be carried out in the future.

Provided the attributes found, the multi-layer perceptron was a successful algorithm for prediction of infertility in women however it is suspected that greater precision may be obtained through raising the number of records used and finding other related characteristics that may help to predict female infertility. Rule-induced algorithms may also be used to map the association between the chosen attributes to assess the possibility of infertility for women using decision-trees algorithm.Selecting methods with classifiers such as SVM, MLP or RF classifiers in the sense of IVF care has produced stronger predictive efficiency than those published on literature. In addition, medical professionals may be easily guided to determine pregnancy results and counsel patients for better IVF care. Studies considering multicollinearity will boost the collection strategy of features and therefore the outcome of pregnancies.

5. Research Gaps and Open issues

It is evidenced from the review that; utmost of the techniques was developed addressing two major criteria in the research for efficient diagnosis of infertility in women. The primary one among them is the enhancing the prediction accuracy of IVF treatment on selecting and

extracting the features from ultrasound treatments. The second one is addressing the problem and diagnosis of PCOS in the women of varied ages. Most of the research have neglected many other vital aspects likeendometriosis, early detection of the blockage of fallopian tubes etc.

On the offset it is understood from the literature that development of a treatment recommendation system based on the symptoms of the infertility in women could be considered as a worthy area of research because, in several cases it is observed that treating infertility could be a multi-criteria decision-making problem as there is a need of addressing several causes of infertility. Optimizing the feature selection could be a promising area that enhances the prediction accuracy in case of worthy treatments like IVF.

6. Conclusion

In this article, a comprehensive and systematic literature review is illustrated to analyse the open issues and research challenges involved in devising computational intelligence algorithms for analysing the clinical data obtained during the IVF treatment and diagnosis of PCOS etc. The main objective of the study is to explore research gaps that could drive the attention of the budding researchers while in this specific research area. In this context, a systematic study is initialized based on the standard procedure so as to clearly identify the open challenges. On reviewing worthy and standard articles from the indexed databases it is inferred that infertility in women is influenced by several factors but most of the research studies have focused on only PCOS and IVF treatment. This scenario indicates that, there is still lot of scope for research in developing intelligent algorithms to handle clinical data related endometriosis and other issues that effect fertility in women. Further more it is observed that there is still lot more scope in developing a treatment recommendation system that analysis historical data of infertile couples succeeded after the treatment, to enhance the prediction rate of the successful treatments for infertility.

7. References:

[1] McDonnell J, Goverde AJ, Rutten FF, Vermeiden JP. Multivariate Markov Chain analysis of the probability of pregnancy in infertile couples undergoing assisted reproduction. Hum Reprod. 2002;17(1):103–6

[2] Fiddelers AA, Dirksen CD, Dumoulin JC, van Montfoort A, Land JA, Janssen JM, et al. Cost-effectiveness of seven IVF strategies: results of a Markov decision-analytic model. Hum Reprod. 2009;24(7):1648–55.

[3] Rao ASRS, Diamond MP. Role of Markov modeling approaches to understand the impact of infertility treatments. Reprod Sci. 2017;11:1538–43.

[4] Hsieh MH, Meng MV, Turek PJ. Markov modeling of vasectomy reversal and ART for infertility: how do obstructive interval and female partner age influence cost effectiveness? FertilSteril. 2007;88(4):840–6

[5] Kakarla N, Bradshaw K, "Evaluation and Management of the Infertile Couple", *The Global Library of Women's Medicine,* 2008

[6] Remah M Kamel, "Management of the infertile couple: an evidencebased protocol", *Reproductive Biology Endocrionol*, Vol 8(21), 2010.

[7] NaeimehTayebi, SeyedMojtabaYassiniArdakani, "Incidence and Prevalence of the Sexual Dysfunctions in Infertile Women", *EuropeanJournal of General Medicine*, Vol 6 (2), pp 74-77, 2009.
[8] Diane Fraser, Margaret Cooper, "Myles' Textbook for Midwives", *Churchill Livingstone, Elsevier*, 2009
[9] Shahnaz Anwar and Ayesha Anwar, "Infertility: A Review on Causes, Treatment and Management", *Women's Health and Gynaecology,* Vol 2 (6), pp 1-5, 2016.
[10] Online URL: http://www.nordfertility.com/en/articles/causes-of-femaleinfertility#Other causes, Last Accessed: June 2017.
[11] Rachel Gurevich, "Blocked Fallopian Tubes: Symptoms and Treatment", Available Online: https://www.verywell.com/all-aboutblocked-fallopian-tubes-1959927, Last accessed: June 2017.
[12] Blocked Fallopian Tubes: Causes, Symptoms and Treatment, Available Online: http://indiraivf.com/blocked-fallopian-tubes-causes-symptomstreatment/, Last Accessed: June 2017.
[13] Hussmann H. (1994) Formal Foundations for Pragmatic Software Engineering Methods. In: Wolfinger B. (eds) InnovationenbeiRechen- und Kommunikationssystemen. Informatikaktuell. Springer, Berlin, Heidelberg
[14] Barbara Kitchenham, O. Pearl Brereton, David Budgen, Mark Turner, John Bailey, Stephen Linkman, Systematic literature reviews in software engineering – A systematic literature review, Information and Software Technology, Volume 51, Issue 1, 2009,
[15] Evaluation of PICO as a Knowledge Representation for Clinical QuestionsXiaoli Huang, Jimmy Lin, Dina Demner-Fushman AMIA AnnuSymp Proc. 2006; 2006: 359–363.
[16] Kaufmann SJ, Eastaugh JL, Snowden S, Smye SW, Sharma V (1997) The application of neural networks in predicting the outcome of in-vitro fertilization. Human Reproduction 12(7):1454-1457
[17] Jurisica I, Mylopoulos J, Glasgow J, Shapiro H, Casper RF (1998) Case-based reasoning in IVF: prediction and knowledge mining. Artificial Intelligence in Medicine 12(1):1-24
[18] Uyar A, Ayse B, Ciray HN (2015) Predictive modeling of implantation outcome in an in vitro fertilization setting: an application of machine learning methods. Medical Decision Making 35(6): 714-725.
[19] Nanni L, Lumini A, Manna C (2010) A data mining approach for predicting the pregnancy rate in human assisted reproduction. Advanced Computational Intelligence Paradigms in Healthcare 5: 97- 111
[20] 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. Studies in Logic, Grammar and Rhetoric 35(1):39-48
[21] Hassan R, Begg R, Taylor S, Kumar DK (2006) HMM-Fuzzy model for recognition of gait changes due to trip-related falls. In: 28th Annual International Conference on Engineering in Medicine and Biology Society (EMBC), pp 1216-1219
[22] Hafiz P, Nematollahi M, Boostani R, Bahia NJ (2017) Predicting implantation outcome of In Vitro fertilization and intracytoplasmic sperm injection using data mining techniques. Fertility and Sterility 11(3):184-190
[23] Ramasamy N, Durairaj M (2017) Feature Reduction by Improvised Hybrid Algorithm for Predicting the IVF Success Rate. Journal of Advanced Research in Computer Science 8(1):37-40

[24] Corani G, Magli C, Giusti A, Gianaroli L, Gambardella LM (2013) A Bayesian network model for predicting pregnancy after in vitro fertilization. Computers in biology and medicine 43(11):1783- 1792

[25] Choi B, Bosch E, Lannon BM, Leveille MC, Wong WH, Leader A, Pellicer A, Penzias AS, Yao MW (2013) Personalized prediction of first-cycle in vitro fertilization success. Fertility and Sterility 99(7):1905-1911

[26] Morales DA, Bengoetxea E, Larrañaga P, García M, Franco Y, Fresnada M, Merino M (2008) Bayesian classification for the selection of in vitro human embryos using morphological and clinical data. Computer Methods and Programs in Biomedicine 90(2):104-116

[27] Guh R, Wu TCJ, Weng SP (2011) Integrating genetic algorithm and decision tree learning for assistance in predicting in vitro fertilization outcomes. Expert Systems with Applications 38(4):4437-4449

[28] Manna C, Nanni L, Lumini A, Pappalardo S (2013) Artificial intelligence techniques for embryo and oocyte classification. Reproductive Biomedicine online 26(1):42-49

[29] Guvenir HA, Misirli G, Dilbaz S, Ozdegirmenci O, Demir B, Dilbaz B (2015) Estimating the chance of success in IVF treatment using a ranking algorithm. Medical & Biological Engineering and Computing 53(9): 911-920

[30] Uyar A, Bener A, Ciray HN, Bahceci M (2010) ROC based evaluation and comparison of classifiers for IVF implantation prediction. Electronic Healthcare 108-111

[31] Sabeena S and Sarojinin B, "Optimal Feature Subset Selection using Ant Colony Optimization", *Indian Journal of Science and Technology,* Vol 8(35), pp 1-5, 2015.

[32] NeeruPathania and Ritika, "Implementation of Fuzzy Controller for Diagnose of Patient Heart Disease", *International Journal of Innovative Science, Engineering and Technology,* Vol 2(4), pp 694-698, 2015.

[33] Nguyen, Khosravi T, Creighton A D and Nahavandi S, "Medical data classification using interval type 2 fuzzy logic system and wavelets", *Applied Soft Computing, Elsevier,* Vol 30, pp 812-822, 2015.

[34] Jonh Hyun Lee, JavadRahimipourAnaraki, Chang WookAhn and Jinung An, "Efficient classification system based on Fuzzy-Rough Feature Selection and Multitree Genetic Programming for intension pattern recognition using brain signal", *International Journal of ExpertSystems with Applications,* Vol 42(3), pp 1644-1651, 2015.

[35] AsliUyar, AyseBener, Nadir Ciray H, Mustafa Bahceci, "ROC Based Evaluation and Comparison of Classifiers for IVF Implantation Prediction", *Proceedings of International Conference on ElectronicHealthcare, Istanbul, Turkey, Lecture Notes of the Institute for ComputerScience, Social Science and Engineering, LNICST,* Vol 27, pp 108-111, 2009.

[36] Durairaj M, Meena K, "A hybrid prediction system using Rough Sets and Artificial Neural Networks", *International Journal of InnovativeTechnology and Creative Engineering,* Vol 1 (7), pp 18-23, 2011

[37] Hongpeng Zhu and Xiaohong Li, "Research on a New Method based on Improved ACO Algorithm and SVM Model for Data Classification", *International Journal of Database Theory and Application,* Vol 9(1), pp 217-226, 2016.

[38] KindieBiredagnNahato, Khanna Nehemiah Harichandran and Kannan Arputharaj, "Knowledge mining from clinical data sets using Rough Sets and Back Propagation Neural Network", *Computational andMathematical methods in Medicine,* Vol 20(15), pp 1-14, 2015

[39] ManjeevanSeera and Chee Peng Lim, "A hybrid intelligent system for medical data classification", *Expert Systems with Applications, Elsevier,* Vol 41, pp 2239-2249, 2014

[40] MeghaRathi and Vikas Pareek, "Disease prediction tool: an integrated hybrid data mining approach for health care", *International Journal ofComputer Science and Information Technology*, Vol 6(6), pp 32-40, 2016.

[41] Rajwant Kaur and Sukhpreet Kaur, "Prediction of Heart Disease Based on Risk Factors using Genetic SVM Classifier*", International Journal ofAdvanced Research in Computer Science and Software Engineering*, Vol 5 (12), pp 205-208, 2015.

[42] Dumesic, D.A., Oberfield, S.E., Stener-Victorin, E., Marshall, J.C., Laven, J.S. and Legro, R.S., 2015. Scientific statement on the diagnostic criteria, epidemiology, pathophysiology, and molecular genetics of polycystic ovary syndrome. *Endocrine reviews*, *36*(5), pp.487-525.

[43] Cheng, J.J. and Mahalingaiah, S., 2018. Data mining and classification of polycystic ovaries in pelvic ultrasound reports. *bioRxiv*, p.254870.

[44] Lawrence, M.J., Eramian, M.G., Pierson, R.A. and Neufeld, E., 2007, May. Computer assisted detection of polycystic ovary morphology in ultrasound images. In Fourth Canadian Conference on Computer and Robot Vision (CRV'07) (pp. 105-112). IEEE.

[45] Escobar-Morreale, H.F., 2018. Polycystic ovary syndrome: definition, aetiology, diagnosis and treatment. *Nature ReviewsEndocrinology*, *14*(5), p.270.

[46] Joham, A.E., Teede, H.J., Ranasinha, S., Zoungas, S. and Boyle, J., 2015. Prevalence of infertility and use of fertility treatment in women with polycystic ovary syndrome: data from a large community-based cohort study. *Journal of women's health*, *24*(4), pp.299-307

[47] Essah, P.A. and Nestler, J.E., 2006. The metabolic syndrome in polycystic ovary syndrome. *Journal of endocrinologicalinvestigation*, *29*(3), pp.270-280

[48] Dhayat, N.A., Marti, N., Kollmann, Z., Troendle, A., Bally, L., Escher, G., Grössl, M., Ackermann, D., Ponte, B., Pruijm, M. and Müller, M., 2018. Urinary steroid profiling in women hints at a diagnostic signature of the polycystic ovary syndrome: A pilot study considering neglected steroid metabolites. *PloS one*, *13*(10), p.e0203903.

[49] Brower, M., Brennan, K., Pall, M. and Azziz, R., 2013. The severity of menstrual dysfunction as a predictor of insulin resistance in PCOS. The Journal of Clinical Endocrinology & Metabolism, 98(12), pp.E1967-E1971