# A review of machine learning methods to build predictive models for male reproductive health

**Ariawan Adimoelja[1], Wayan Firdaus Mahmudy[2], Diva Kurnianingtyas[2]**
[1]Doctoral Program in Medical Science, Faculty of Medicine, Universitas Brawijaya, Malang, Indonesia
[2]Department of Informatics Engineering, Faculty of Computers Science, Universitas Brawijaya, Malang, Indonesia

**Article Info**

***Article history:***

Received Dec 19, 2023
Revised Mar 10, 2024
Accepted Mar 21, 2024

***Keywords:***

Artificial intelligence
Artificial neural network
Fuzzy inference system
Male health
Medical

**ABSTRACT**

Developing of artificial intelligence (AI) technology in the medical sector, especially in the part of male reproduction and infertility, is growing rapidly. In both supervised learning and unsupervised learning, AI has been tested and applied to medical personnel to treat their patients. Calculations from simple to complex probability and a combination of some different methods have conducted results of accurate and precise. The results can help determine the condition of male infertility. Artificial neural network (ANN) and fuzzy inference system (FIS) are AI techniques applied to male health issues. ANN is adequate for processing large amounts of combined data in a short time. ANN also has a high level of accuracy and excellent adaptive capabilities. Afterwards, FIS can reflect problems using models with easy to understand, flexible, and also competent to model complex linear functions for decision-making. Based on the advantages of ANN and FIS, it is hoped acquiring prediction results of better and more accurate in male health issues.



***Corresponding Author:***

Wayan Firdaus Mahmudy
Department of Informatics Engineering, Faculty of Computer Science, Brawijaya University
Jl. Veteran No.10-11, Ketawanggede, Lowokwaru, Malang, Indonesia 65145
Email: wayanfm@ub.ac.id

## 1. INTRODUCTION

Difficulty in conceiving offspring in couples of childbearing age is still a problem. Although the development of therapy in the medical field to overcome the infertility issue has been very developed. In the last half-century, the resolution of infertility problems has been more emphasized in the process of handling therapy in the reproductive system of both men and women [1]. The therapy stage not only takes a lot of time but also costs and sometimes does not produce results. The problem becomes an impractical and difficult obstacle for young couples. For young couples experiencing fertility issues, the treating procedure refers to the many stages of examination. As a result, many are reluctant to have their fertility checked because it is very impractical. The success of fertility treatment is not only determined by the number of examination procedures but also by the doctor's expertise in exploring the patient's medical record. Time constraints are a contributing factor to the failure of treatment because there are often many records of medical and family that have not been clarified. Based on this explanation, an artificial intelligence (AI) approach can make early predictions carefully on fertility issues [2], [3]. AI can be applied as a means of supporting clinical guidance to improve the accuracy of diagnosis and reduce the possibility of misdiagnosis. AI can be utilized as a repository of necessary medical data. By utilizing AI, it can find relationships between the necessary data to produce a significant prediction solution [4].

The use of AI has grown rapidly in various fields including the medical field. Initially, AI was still experimental, now it has successfully become an implementation in various fields. Technological advances in AI and machine learning (ML) have been widely utilized by the medical sector to help the computer decision-making process system [2], [5]–[8]. The science of artificial intelligence extends widely in the medical field, for example to the detection of coronary artery diseases [9], breast cancer detection [10], sexually transmitted disease identification [11], and classification of male and female infertility problems [12]–[15].

ML as a part of AI has the ability to recognize and make predictions based on certain patterns from a large and complex data set. ML is very useful for use in the medical field because it is closely related to various kinds of data. Artificial neural network (ANN), a type of ML, has become a favorite because of its ability and flexibility in classification, clustering, pattern recognition, and prediction processes in various disciplines. ANN is also very compatible when combined with conventional regression and other statistical modeling [16], [17]. ANN can be used to evaluate factors in data analysis such as accuracy, calculation speed, performance, scalability fault tolerance, and convergence [18], [19]. The specialty of ANNs is their very rapid ability to process various data derived from various combinations of similar implementations in very large quantities in a short time. In addition, ANNs also have a high degree of accuracy and excellent adaptive capabilities [18]. Therefore, ANN is one option that can be considered in the medical sector which has a lot of data from various sources that need to be combined and synchronized. A combination of statistical prediction modeling and neural network (NN) capabilities can be used to detect early fertility issues. In addition, it is very useful because it results in better cost-effectiveness. However, making an accurate prediction model is not easy, many things must be considered so that a statistical prediction model can proceed significantly.

This research will present a systematic review of several things that are taken into consideration in constructing ANN and fuzzy inference systems (FIS) for infertility prediction. In addition, this review will discuss various prediction models that have been used in the medical field to help streamline the diagnosis procedures. Finally, recommendations will be given for constructing an accurate prediction model by combining various AI methods.

## 2. CLASSIFICATION AND PREDICTION

### 2.1. Methods for classification and prediction

In developing ML modeling, several types of classifiers must be considered to provide the predictions given to achieve accurate results. According to Tsai *et al.* [20], ML algorithms used to predict something can be grouped into 4 major parts, namely pattern classification, single classifiers, hybrid classifiers, and ensemble classifiers as shown in Figure 1. Algorithms in single classifiers are K-nearest neighbour (K-NN), support vector machine (SVM), ANN, self-organizing map (SOM), decision tree, naïve Bayes network, genetic algorithm (GA), and fuzzy logic.

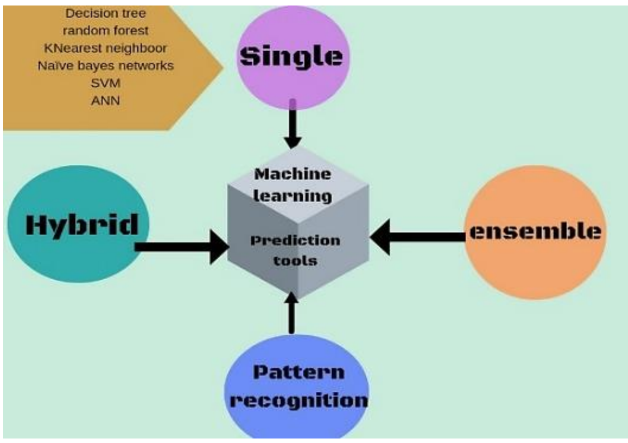


Figure 1. ML as a predictive tool

Wang *et al.* [4] compared the advantages and limitations of four different single classifiers namely decision tree, random forest, SVM, and naïve Bayes classifier. The following comparison of the four single classifiers is shown in Table 1. K-nearest neighbor (K-NN) is known as the simplest and most traditional method for classifying samples. However, K-NN is rarely used. K-NN does not have a training stage, causing biased results when dealing with complex data [20]. In contrast, ANN is now the most widely used algorithmic method due to its ability to handle large amounts of complex data. ANNs in general have good capabilities and performance for clustering classification and prediction in various disciplines [21]–[23]. Fuzzy logic is better known as a component of soft computing than as part of predictive classification. However, fuzzy logic is now starting to be widely used as a predictive modeling tool [18], [24], [25]. Fuzzy logic works by mapping the problem from input to output. The advantages are easy to understand, flexible, able to model complex linear functions, and map of the experience of experts to be used directly in a decision-making process [18].

Single classifiers were the most widely used predictive classification algorithm from 2000-2007. However, those single classifiers were considered less accurate for modeling comparisons and evaluations. Using a mixture of two or more classification methods such as hybrid and ensemble classifiers is considered to have a better future for modeling. The reason is that both methods complement each other's shortcomings. Finally, a better level of prediction accuracy is obtained [20], [26].

Table 1. Comparison of single classifiers [4]

| | Decision tree | Random forest | Support vector machine | Naïve Bayes classifier |
|---|---|---|---|---|
| Advantages | Easy to understand and interpret | Easy to understand and interpret, Able to correct the overfitting problem in the decision tree | Process large and complex amounts of data from various sources Able to solve linear and non-linear problems | Easy to understand and interpret Requires little training data |
| Shortcoming | Overfitting | Maintenance process takes a long time | Hard to train | Shortcoming |
| Performance | | | Highest | |
| Accuracy | | Highest | High | |
| Technique used | | | Hyperplane | Conditional probability |
| Reference | [27]–[31] | [32] | [33]–[39] | [40], [41] |

## 2.2. Predictive models for male reproductive health and infertility

Various methods that can predict male infertility are urgently needed. This aims to make men increasingly aware that infertility problems do not only occur in women. Certain diseases or an unhealthy lifestyle can contribute to their fertility. Based on the explanation, many applications emerged as preventive action by predicting male fertility [18], [25]. Based on a systematic review conducted by Zarinara *et al.* [1], there are many prediction methodologies to measure fertility rates based on statistical methods and NN algorithms. However, much research is still needed to produce a valid prediction. A prediction model can be applied clinically if it has gone through a statistical evaluation and a good validation value. It is intended that therapy recommendations follow the prediction model and theoretical approach. Incorrect predictions can lead to adverse consequences, both for doctors and patients. Therefore, a predictive model can be used clinically when it conforms to the truth of the theory, has been through many evaluations, and has high validity. In addition, the results must be recorded and evaluated periodically because if there is an anomaly it can be solved.

In the last decade, there has been a lot of literature discussing the creation of a framework capable of predicting fertility rates [1], [4], [12], [18], [23]. The following Table 2 shows some of the literature that discusses predictive models related to male fertility rates. Furthermore, Table 3 shows the accuracy of several predicting models for male infertility that have been done in the past decade.

Table 2. Research using the FIS and fuzzy logic

| Researcher | Method | Issue | Outcome |
|---|---|---|---|
| [42] | – MLP<br>– SVM<br>– Decision tree | Relationship between lifestyle and semen quality | MLP and SVM showed high accuracy |
| [12] | – Logistic regression<br>– SVM<br>– Random forest | Relationship of varicocelectomy sperm parameters with semen analysis data, hormonal, and clinical data | A clinical calculator application can be used for patient counseling |
| [43] | – Decision tree<br>– Random forest<br>– Naïve Bayes<br>– K-NN<br>– SVM<br>– Superlearner | Effectiveness of 6 different algorithms with different split ratios | Random forest produces the highest level of accuracy compared to the others and the use of different split ratios produces different levels of accuracy |
| [18] | Fuzzy logic toolbox from MATLAB software | Measurement of infertility risk of Nigerian men based on 8 non-invasive variables affecting their fertility | Fuzzy logic can be used as a consideration for one of the predictive tools in determining the level of male fertility |

Table 3. Comparison accuracy of several predicting models

| Reference | Predicting model algorithm | Research topic | Number of samples | Accuracy (%) |
|---|---|---|---|---|
| Goodson | SVM (motility) | Classify motility sperm rate of individual sperm | - | 89.9 |
| Tseng (2013) | SVM (morphology) | Classify sperm morphology using sperm contour | 80 | 87.5 |
| Dash | ensemble learning: | Seminal analysis using lifestyle factors | 100 | 89 |
| | Voting classifier | | | 88 |
| | Bagging classifier | | | |
| Koch | Supervised learning | Diagnosing the risk factors of male infertile disease | 329 | 97 |
| Oyegoke | Fuzzy logic | Risk factors of male infertility | | Not stated |
| Ory, J | Random forest | Predicting sperm parameter upgrading after Varicocele repair | - | 72 |
| Akinsal | ANN | Predicting chromosomal abnormalities in Azoospermia male | - | 95 |

According to Oyegoke *et al.* [18], further research is needed on male fertility rates by a combination of ML classifications integrated with fuzzy logic abilities. The combination of ML and fuzzy logic is believed to be able to make decision-making processes in the medical sector more effective and accurate. Although fuzzy logic has been widely applied, cases for predicting male fertility are still very rare. Designing a model for predicting male fertility rates to be used independently without having to go through a series of therapies is a breakthrough with provides opportunities for further research.

## 3. ARTIFICIAL NEURAL NETWORK AND FUZZY INFERENCE SYSTEM CONCEPT

### 3.1. Supervised versus unsupervised learning

Supervised learning (SL) and unsupervised learning (UL) are part of ML. SL has been widely used in research on reproductive health. SL is a computer learning process that still requires human intervention in its operation. SL processes data based on datasets in the form of training data and testing data. Computers are trained first based on data that has been labeled and monitored in data processing. The SL algorithm is made using labeled data features to predict an output that has been prepared beforehand. The weakness of SL is that the labeling of its data features must be done by humans, it cannot work automatically on its own, so it requires a lot of time for data processing and is prone to human errors. Algorithms in SL are usually for data classification or numerical prediction/regression. The algorithms for classification can be used logistic regression, decision trees, random forest, K-NN, SVM, NN, and naïve Bayes. Algorithms for numerical prediction or regression such as linear regression, decision trees, NN, SVM, and trees. Among the various existing algorithms, ANN and SVM are known as the most widely implemented algorithms because they can process large and complex amounts of data with a higher level of accuracy compared to other SL algorithms.

Meanwhile, UL is a computer learning process without using labeled data and minimal human intervention in its operation. There is no training data set and no feature targets as output in UL. The computer learns it independently without any initial supervision. UL is an algorithm whose modeling focuses on hidden structures and the relationships between data sets. UL does not require labeled data as input data but only needs to include data features to be used as training data. UL can be used to predict outcomes that may be beyond predictions. UL is very well used for modeling tasks related to association and clustering [4]. Furthermore, Clustering data is grouped into small groups that have the same patterns of properties and characteristics as components analysis and clustering algorithms, for example, K-means clustering, hierarchical clustering, and association rule-learning [44], [45]. UL is usually used for deep learning and has been widely applied in the manufacture of motorized vehicle robot automation, voice recognition, and pattern recognition [4], [44]. Research in the reproductive health sector using UL is still rare [46].

### 3.2. Famous neural network and fuzzy inference system and its applications

ANN was originally designed to imitate how the neurons in the human brain work. ANN consists of nodes that represent neurons in the human brain. Each node has an activation function and a certain definition as the output [47]. One of the most widely used ANNs in the last decade is the back propagation neural network (BPNN). BPNN is an algorithm using guided training data. Therefore, BPNN is part of SL. BPNN consists of three layers, namely input layers, hidden layers, and output layers. BPNN works through three phases, namely the feed-forward phase, the backpropagation phase, and the weight modification phase. First, the input signal will be propagated forward to the output through a pre-weighted hidden layer. Afterward, BPNN utilizes the error output to change the weight value through backpropagation. Output that does not have an activation value of zero is considered an error. The output error will undergo weight modification to acquire the smallest error. Activation values apply binary or bipolar sigmoid formulas. BPNN consists of three types of data, namely training data, test data, and predictive data. Meanwhile, to determine the number of hidden neurons, trial and error experiments are needed.

One form of self learning neural network (SLNN) is the Kohonen NN. Kohenen NN is used to find invisible connections or similarities between data in a data set. Warpechowski *et al.* [48] applies the Kohenen NN method to compare the level of knowledge among medical students at Bialystok University, Poland. In this research, Kohenen NN was used to analyze material about reproductive health and its relationship with lifestyle. The research material applied is a questionnaire method with an answer scale between 1-10. The questionnaire is grouped into two major sections. The first part is general questions about the respondent. The second part is questions related to the factors that influence lifestyle that affect reproductive health. Furthermore, the second part is separated into several more specific groups, such as knowledge about the lifestyle effects on the psychology of fertility and the impact of food on fertility. The first part is general data of respondents grouped by age group, major, and courses taken. After the clustering process is complete, all data is entered into the database and analyzed by Kohonen NN. According to Warpechowski *et al.* [48], Kohonen NN has the ability to self-learn. This means are the data entered does not need to be made into a special pattern beforehand. From the data analysis, certain new patterns will result. The reason is that the Kohonen NN topology has many neurons that can sort and make new classifications. Therefore, Kohonen NN is suitable for predicting multidimensional input data such as medical data. Kohonen NN is used to find correlations between data that are hidden and cannot be mapped through general statistical methods.

In its implementation, Kohonen NN uses 15% of the data set for group testing and validation. Meanwhile, the remaining 70% of the data is used for training data. From this research, it was found that Kohonen NN can be used to find unseen correlations. Based on the standard statistical modeling results, future research is recommended to use more data variations as input.

ANN is widely applied for modeling in the medical field. However, one of the challenges is the instability of the data set which causes bias and traps gradient descent at a local minimum. Some researchers tried to use a metaheuristic algorithm by continuously updating the weight values to minimize the error. However, this method was not successful in overcoming the local minimum problem [23]. To overcome this problem, Yibre and Koçer [23] tried to conduct experimental research using the feed forward neural network (FFNN) approach combined with the artificial algae algorithm (AAA) to determine the accuracy of the prediction. Afterward, synthetic minority over-sampling technique (SMOTE) is applied to stabilize the data set. The experimental results will be compared with single classification algorithms such as multi-layer perceptron (MLP), naïve-Bayes, SVM, K-NN, and random forest. This approach shows significant results in overcoming the problem of biased data. Data imbalance usually occurs when too many types of sample data are used. The solution is to balance the data set. The solution will not only optimize the capability of SL algorithms but also improve the generalization process.

AAA is a basic algorithm commonly used in population studies. AAA is formed by imitating the way microalgae move to survive. AAA shows better performance, compared to the metaheuristic algorithm, when used to solve continuous optimization problems. There are three stages in AAA, namely evolutionary, adaptation, and helical movement. The algae's ability to reach the light is considered the global optimal point. Before starting the main process, the algorithm starts with an initial solution and tests its working level. Next, the colony size of the algae is calculated based on the formula. AAA imitates the helical movement of algae cells. Helical motion is the movement of an algae cell as it moves from one position to another on the water surface to absorb light. This movement depends on the surface friction and energy level. The higher the friction, then the higher the frequency of helical motion to increase local search. When the surface friction is low the coverage area of the algae cell will widen. The gravity of AAA is considered zero. The position of each alga depends on the dragging force of the algae movement in the liquid. Based on the principle of algae cell movement, linear and angular movement equations are formed. The evolutionary process in algae cells occurs when algae cells get enough food and can form colonies. Each algae cell will divide into two new cells. Thus, the algae colony will grow bigger. Algae cells that cannot divide will die. If there is a growth of algae colony, it will give a better prediction result. Each artificial algae gets an initiation value of zero. Furthermore, colonies that can grow will give optimal results, while unsuccessful colonies will get hungry and unable to grow, which means the prediction results are poor. The purpose of using AAA is to find the optimal weight value so that mean squared error (MSE) and FFNN can run accurately. This optimal weight is used in the training, validation, and testing process.

FFNN is one of the ANN methods where the information flow process is only one way from input, hidden to output. FFNN is a form of SL method for classification. FFNN is the most widely used algorithm for a variety of issues such as disease diagnosis, pattern classification, and determination of misidentification. The number of nodes in the input layer of FFNN is equal to the number of features contained in the classification data set. The weight calculation is conducted in the hidden layer with a range between 0-1 with a sigmoid activation function. FFNN output results sometimes do not match what is expected. The error that occurs is calculated as the difference between the target output and the actual output. In this case, propagated back is conducted to update the weights. An illustration can be shown in Figure 2.

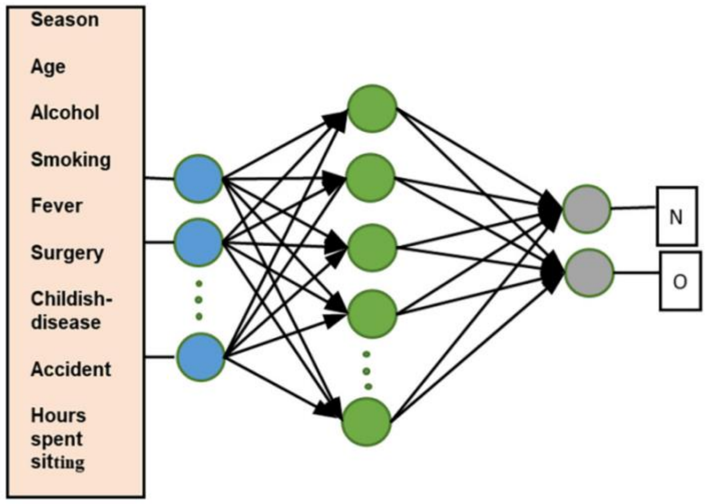


Figure 2. FFNN architecture [23]

Fuzzy logic is not a new approach, first introduced in 1965. Fuzzy methods are suitable for conditions where the boundaries are unclear and difficult to understand. Therefore, fuzzy logic was first implemented in the medical sector [24]. The reason why this method is suitable to be applied in the medical field is because decision-making deals with a lot of uncertainty, difficult to formulate and measure [49]. Fuzzy logic is associated with logic to make the right decision under conditions of uncertainty and incompleteness of the situation. Fuzzy logic is a form of mapping from input to expected output. According to Arji's paper, it is explained that almost 50% of the research uses the FIS method [24].

Based on Figure 3, FIS has a high evaluation performance indicator or a high level of accuracy. FIS modeling is based on IF_THEN rules based on fuzzy sets. FIS is a prediction model that can be used well to predict behavior in a variety of uncertain conditions, such as in modeling infectious disease conditions that have many uncertain conditions [24]. There are three requirements for FIS modeling, namely having i) a membership function, ii) a fuzzy set operation procedure, and iii) an inference procedure [50].

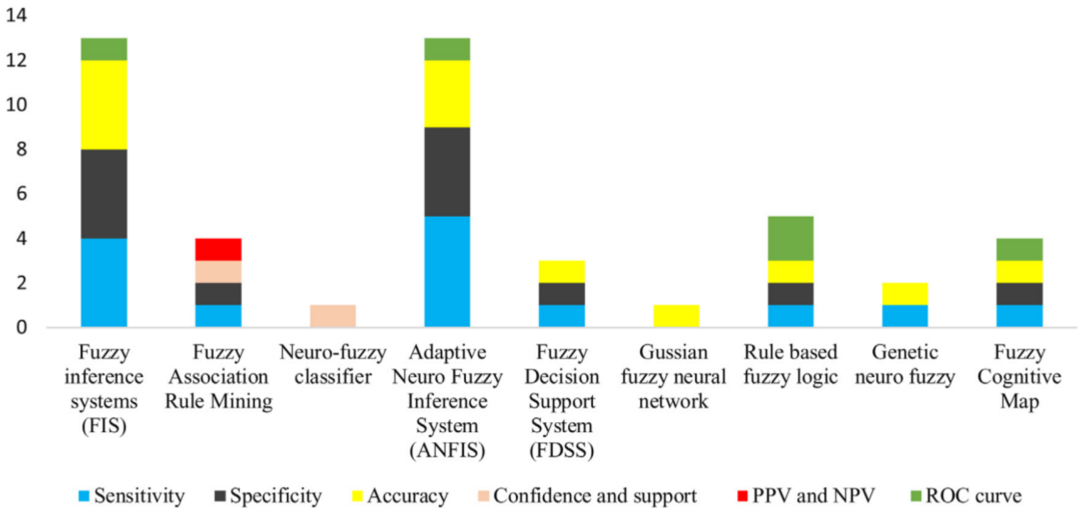


Figure 3. Algorithms performance comparison [24]

FIS has four stages of the modeling process, namely fuzzification, weighting, assessment of inferences procedures, and de-fuzzification [24]. Fuzzification is the process of breaking data into input variables that can be accepted by fuzzy logic. This process is used to map the interval [a, c] into a fuzzy value that matches the mathematical equation. Weighting is the process of labeling membership functions (3 triangular membership or 2 triangular membership) usually worth between 0, 1, or 2. Assessment of inferences procedures is the process of testing each input variable and its membership function. De-fuzzification is the process of returning variables to their original data [18].

Three FIS methods can be used, namely the method of Mamdani, Tsukamoto, and Sugeno. The difference between the method of Mamdani and Sugeno is the input and output variables in the Mamdani have the same fuzzy statement, while in the Sugeno there is a nonlinear affiliation. Therefore, choosing the right method affects the expected results [24]. FIS and fuzzy logic have been implemented in various studies, which are shown in Table 4.

Table 4. Research using FIS and fuzzy logic [51]

| Researcher | Year | Disease | Fuzzy technique | Research title | Conclusion |
|---|---|---|---|---|---|
| [52] | 2015 | Peritonitis | FIS | Developing a fuzzy inference system for diagnosing peritonitis | The proposed FIS technique enables physicians to easily diagnose peritonitis |
| [53] | 2017 | Urinary tract infections (UTIs) | FIS | To achieve an improvement in constructing the outcomes of the microscopic urine examination by means of fuzzy logic methods. | The fuzzy logic foretells the existence of urinary tract infection in the patient by microscopically examined parameters |
| [54] | 2012 | Airborne transmission infection | Fuzzy logic | Use of the fuzzy logic model to predict airborne transmission infection | The fuzzy logic model demonstrates a similar threat of infection for secondary cases compared to the Gaussian dispersion technique. |
| [18] | 2019 | Infertility in man | Fuzzy logic | A predictive model for the risk of infertility in men using fuzzy logic | Fuzzy logic with a triangular membership function could predict male infertility risk |

### 3.3. Hybrid methods

The hybrid classifier is a modeling that is a combination of two or more ML techniques to produce a system with better performance and results [20], [51] groups hybrid classifications into several types based on the combination modeling technique used. A comparison of hybrid ML methods is shown in Table 5. The first type is hybrid modeling using a combination of two kinds of functional components. The first component is usually to process raw data as input and produce output in semi-finished results. The output will be followed by processing using the second component to produce the final result [55]. The second type uses a combination of two different classification techniques to deliver a new technique. An example in the research of [26] combines several fuzzy techniques with a decision tree to construct a new technique called the multiple fuzzy frequent pattern (MFFP) tree.

Table 5. Result of method comparison in hybrid ML [51]

| Method | Usage | Accuracy | Reliability | Sustainability |
|---|---|---|---|---|
| Hybrid WNN-ARIMA | Estimate | +++ | +++ | +++ |
| WNN | Estimate | ++ | ++ | ++ |
| ARIMA | Estimate | ++ | + | + |
| Hybrid EE-ANT-WNN | Estimate | +++ | +++ | +++ |
| Hybrid the optimized multi-stage method | Estimate | +++ | +++ | +++ |
| BAGNBT | Estimate | +++ | +++ | +++ |
| SVM | Estimate | ++ | + | + |
| NBT | Estimate | + | + | + |
| RFNBT | Estimate | ++ | ++ | ++ |
| EEMD-ELM-GOA | Estimate | +++ | +++ | +++ |
| HybPAS | Estimate | +++ | +++ | +++ |

The third type applies several clustering methods to process the initial data sample. The data sample generates a quality training data sample. Training data is used to cluster larger data, usually by combining supervised and unsupervised. One example of the use of a hybrid by combining supervised and unsupervised is in the study of Niño *et al*. [56]. The research was conducted by combining the classification of single-unit decision trees followed by clustering [26]. The fourth type is the integration of two classification techniques to obtain optimal results and create better model predictions. The impact can increase the results that were previously optimal to be more optimal.

### 3.4. Ensemble methods

According to Tsai *et al.* [20], ensemble classifiers combine several algorithm weaknesses to provide different treatments by trying to provide various training samples to obtain a more significant performance increase. Meanwhile, Ardabili *et al.* [51] stated that ensemble was created by combining various grouping techniques such as bagging and boosting with several single classifications from ML to increase the accuracy of the prediction model. The single classification commonly used is the decision tree. The ensemble method is included in the SL algorithm. This method makes it possible to train different algorithms by creating flexible training data to obtain the highest accuracy results. Table 6 provides a complete illustration of the ensemble method comparison in terms of accuracy, reliability, and sustainability. According to Dash and Ray [42], ensemble classifiers provide a higher level of accuracy compared to the use of a single classifier because it combines the advantages of several single classifier algorithms. With ensemble classifiers, it is possible to combine the results from each classifier and make the final decision [57].

Table 6. Result of method comparison in ensemble ML [51]

| Method | Usage | Accuracy | Reliability | Sustainability |
|---|---|---|---|---|
| Ensemble TSM | Estimate | +++ | +++ | +++ |
| Ensemble GDBT | Estimate | +++ | +++ | +++ |
| Ensemble EBFTM | Estimate | +++ | +++ | +++ |
| Random forest | Estimate | +++ | ++ | ++ |
| BUT | Estimate | ++ | ++ | ++ |
| ICEEMDAN-ELM | Estimate | +++ | +++ | +++ |
| ICEEMDAN-OSELM | Estimate | ++ | + | + |
| ICEEMDAN-RF | Estimate | ++ | ++ | ++ |
| Ensemble KF-DA | Estimate | ++ | ++ | ++ |
| Ensemble bagging-boosting | Estimate | +++ | +++ | +++ |
| DTFNN | Estimate | ++ | + | + |

Dash and Ray [42] divide ensemble classifiers into four types, namely bagging, random forest classifiers, extra tree (ET) classifier, and voting classifiers. The bagging technique or bootstrap aggregating is an effective technique, working by deriving missing functions from various components such as the probability error of a classification base. Random forest technique is based on combining several twig algorithms to obtain classification and regression [58], [59]. ET classifier is a collection of randomized trees that have been created. Then combined into an ensemble form which has greater randomization power. The main difference from the ET classifier is that it is easy to choose the intersection points at random. In addition, this technique has the ability to reduce classification variances in general to improve classifier voting accuracy. This technique in principle tries to use a different approach than the existing classifier models to create a more accurate ensemble model. Voting selection is based on major voting and soft voting or average probability. Dash and Ray [42] created a voting classifier by combining soft voting techniques with traditional bagging classifiers. At the beginning of the research, they used several types of single classifiers and then voted to decide which one was the best to combine with the bagging ensemble technique. For the soft voting technique, 'p' is used to determine the probability of each single classifier and adds the weight 'w' to calculate the average value of the various combinations of single classifier. The voting classifier has an accuracy rate of 89% compared to other prediction models [42].

## 4. CONCLUSION

AI in the medical sector is undergoing rapid progress. Likewise, the algorithm application is used to help medical personnel know the condition of their patients as well as an aid in making medical decisions. ANN is one of the commonly implemented algorithms for more accurate prediction calculations with the weighting of each process. In its implementation, BPNN is applied to correct prediction errors with a backward step. The results from BPNN tend to represent actual conditions compared to other methods. Male reproductive and infertility issues are a problem combination of conditions and circumstances that are interrelated and influence each other. Various risk factors impact the results of sperm examination. Therefore, it is necessary to choose a suitable algorithm for precise and accurate prediction process. ANN and FIS are choices of solutions in building a model to predict the outcomes of complex problems involving many interrelated factors. Selection and combination of machine learning methods can provide more accurate predictions with minimum errors.

## REFERENCES


[1] A. Zarinara, H. Zeraati, K. Kamali, K. Mohammad, P. Shahnazari, and M. M. Akhondi, "Models predicting success of infertility treatment: A systematic review," *Journal of Reproduction and Infertility*, vol. 17, no. 2, pp. 68–81, 2016.

[2] Y. Ma *et al.*, "IAS-FET: An intelligent assistant system and an online platform for enhancing successful rate of in-vitro fertilization embryo transfer technology based on clinical features," *Computer Methods and Programs in Biomedicine*, vol. 245, 2024, doi: 10.1016/j.cmpb.2024.108050.

[3] P. Bagane, S. Kulshretha, M. Mishra, Y. Asrani, and V. Maheswari, "IVF success prediction using machine learning techniques: a comparative study," *International Journal of Intelligent Systems and Applications in Engineering*, vol. 12, no. 4, pp. 819–829, 2024.

[4] R. Wang *et al.*, "Artificial intelligence in reproductive medicine," *Reproduction*, vol. 158, no. 4, pp. R139–R154, 2019, doi: 10.1530/REP-18-0523.

[5] L. Atallah, M. Nabian, L. Brochini, and P. J. Amelung, "Machine learning for benchmarking critical care outcomes," *Healthcare Informatics Research*, vol. 29, no. 4, pp. 301–314, 2023, doi: 10.4258/hir.2023.29.4.301.

[6] J. Choi *et al.*, "Current status and key issues of data management in tertiary hospitals: a case study of seoul national university hospital," *Healthcare Informatics Research*, vol. 29, no. 3, pp. 209–217, 2023, doi: 10.4258/hir.2023.29.3.209.

[7] S. Lenka, S. K. Pradhan, S. P. Nayak, and S. C. Nayak, "Forecasting quality of service (QoS) parameters of iot based web services with an artificial electric field algorithm trained artificial neural network," *Karbala International Journal of Modern Science*, vol. 9, no. 1, pp. 137–148, 2023, doi: 10.33640/2405-609X.3282.

[8] A. Y. S. Lau and P. Staccini, "Artificial intelligence in health: new opportunities, challenges, and practical implications," *Yearbook of medical informatics*, vol. 28, no. 1, pp. 174–178, 2019, doi: 10.1055/s-0039-1677935.

[9] K. S. Nugroho, A. Y. Sukmadewa, A. Vidianto, and W. F. Mahmudy, "Effective predictive modelling for coronary artery diseases using support vector machine," *IAES International Journal of Artificial Intelligence*, vol. 11, no. 1, pp. 345–355, 2022, doi: 10.11591/ijai.v11.i1.pp345-355.

[10] A. Ridok, N. Widodo, W. F. Mahmudy, and M. Rifa'i, "A hybrid feature selection on AIRS method for identifying breast cancer diseases," *International Journal of Electrical and Computer Engineering*, vol. 11, no. 1, pp. 728–735, 2021, doi: 10.11591/ijece.v11i1.pp728-735.

[11] G. E. Yuliastuti, A. N. Alfiyatin, A. M. Rizki, A. Hamdianah, H. Taufiq, and W. F. Mahmudy, "Performance analysis of data mining methods for sexually transmitted disease classification," *International Journal of Electrical and Computer Engineering*, vol. 8, no. 5, pp. 3933–3939, 2018, doi: 10.11591/ijece.v8i5.pp3933-3939.

[12] J. Ory *et al.*, "Artificial intelligence-based machine learning models predict sperm parameter upgrading after varicocele repair: a multi-institutional analysis," *World Journal of Men's Health*, vol. 40, 2022, doi: 10.5534/wjmh.210159.

[13] R. Subha, B. R. Nayana, R. Radhakrishnan, and P. Sumalatha, "Computational intelligence for early detection of infertility in women," *Engineering Applications of Artificial Intelligence*, vol. 127, 2024, doi: 10.1016/j.engappai.2023.107400.

[14] J. C. A. Suárez and C. P. F. Ramos, "Identification of sources of male sterility in the Colombian Coffee Collection for the genetic improvement of Coffea arabica L.," *PLoS ONE*, vol. 18, no. 9, 2023, doi: 10.1371/journal.pone.0291264.

[15] H. Krenz *et al.*, "Machine learning based prediction models in male reproductive health: Development of a proof-of-concept model for Klinefelter Syndrome in azoospermic patients," *Andrology*, vol. 10, no. 3, pp. 534–544, 2022, doi: 10.1111/andr.13141.

[16] Y. Li, Q. Zhang, and S. W. Yoon, "Gaussian process regression-based learning rate optimization in convolutional neural networks for medical images classification," *Expert Systems with Applications*, vol. 184, 2021, doi: 10.1016/j.eswa.2021.115357.

[17] S. A. E. -Mienye, E. Esenogho, and T. G. Swart, "Integrating enhanced sparse autoencoder-based artificial neural network technique and softmax regression for medical diagnosis," *Electronics*, vol. 9, no. 11, pp. 1–13, 2020, doi: 10.3390/electronics9111963.

[18] T. Oyegoke, A. Amoo, J. Balogun, T. Alo, and P. Idowu, "A predictive model for the risk of infertility in men using fuzzy logic," *International Journal of Computers*, vol. 4, pp. 27–37, 2019.

[19] A. Mozaffari, M. Emami, and A. Fathi, "A comprehensive investigation into the performance, robustness, scalability and convergence of chaos-enhanced evolutionary algorithms with boundary constraints," *Artificial Intelligence Review*, vol. 52, no. 4, pp. 2319–2380, 2019, doi: 10.1007/s10462-018-9616-4.

[20] C. F. Tsai, Y. F. Hsu, C. Y. Lin, and W. Y. Lin, "Intrusion detection by machine learning: a review," *Expert Systems with Applications*, vol. 36, no. 10, pp. 11994–12000, 2009, doi: 10.1016/j.eswa.2009.05.029.

[21] M. Marji, A. Widodo, M. Marjono, W. F. Mahmudy, and M. M. Arifin, "Comparison of multi-layer perceptron and support vector machine methods on rainfall data with optimal parameter tuning," *International Journal of Advanced Computer Science and Applications*, vol. 14, no. 7, pp. 414-419, 2024, doi: 10.14569/IJACSA.2023.0140745.

[22] J. Sak and M. Suchodolska, "Artificial intelligence in nutrients science research: a review," *Nutrients*, vol. 13, no. 2, pp. 1–17, 2021, doi: 10.3390/nu13020322.

[23] A. M. Yibre and B. Koçer, "Semen quality predictive model using feed forwarded neural network trained by learning-based artificial algae algorithm," *Engineering Science and Technology, an International Journal*, vol. 24, no. 2, pp. 310–318, 2021, doi: 10.1016/j.jestch.2020.09.001.

[24] G. Arji *et al.*, "Fuzzy logic approach for infectious disease diagnosis: A methodical evaluation, literature and classification," *Biocybernetics and Biomedical Engineering*, vol. 39, no. 4, pp. 937–955, 2019, doi: 10.1016/j.bbe.2019.09.004.

[25] K. Mittal, A. Jain, K. S. Vaisla, O. Castillo, and J. Kacprzyk, "A comprehensive review on type 2 fuzzy logic applications: Past, present and future," *Engineering Applications of Artificial Intelligence*, vol. 95, pp. 1-12, 2020, doi: 10.1016/j.engappai.2020.103916.

[26] O. Cordón, P. Kazienko, and B. Trawiński, "Special issue on hybrid and ensemble methods in machine learning," *New Generation Computing*, vol. 29, no. 3, pp. 241–244, 2011, doi: 10.1007/s00354-011-0300-3.

[27] R. S. Guh, T. C. J. Wu, and S. P. Weng, "Integrating genetic algorithm and decision tree learning for assistance in predicting in vitro fertilization outcomes," *Expert Systems with Applications*, vol. 38, no. 4, pp. 4437–4449, 2011, doi: 10.1016/j.eswa.2010.09.112.

[28] T. B. Mesen, J. E. Mersereau, J. B. Kane, and A. Z. Steiner, "Optimal timing for elective egg freezing," *Fertility and Sterility*, vol. 103, no. 6, pp. 1551-1556, 2015, doi: 10.1016/j.fertnstert.2015.03.002.

[29] R. Devjak, T. B. Papler, I. Verdenik, K. F. Tacer, and E. V. Bokal, "Embryo quality predictive models based on cumulus cells gene expression," *Balkan Journal of Medical Genetics*, vol. 19, no. 1, pp. 5–12, 2016, doi: 10.1515/bjmg-2016-0001.

[30] B. Carrasco *et al.*, "Selecting embryos with the highest implantation potential using data mining and decision tree based on classical embryo morphology and morphokinetics," *Journal of Assisted Reproduction and Genetics*, vol. 34, no. 8, pp. 983–990, 2017, doi: 10.1007/s10815-017-0955-x.

[31] L. L. V. Loendersloot *et al.*, "Cost-effectiveness of single versus double embryo transfer in IVF in relation to female age," *European Journal of Obstetrics and Gynecology and Reproductive Biology*, vol. 214, pp. 25–30, 2017, doi: 10.1016/j.ejogrb.2017.04.031.

[32] P. Hafiz, M. Nematollahi, R. Boostani, and B. N. Jahromi, "Predicting implantation outcome of in vitro fertilization and intracytoplasmic sperm injection using data mining techniques," *International Journal of Fertility and Sterility*, vol. 11, no. 3, pp. 184–190, 2017, doi: 10.22074/ifs.2017.4882.

[33] R. K. K. Lee *et al.*, "Sperm morphology analysis using strict criteria as a prognostic factor in intrauterine insemination," *International Journal of Andrology*, vol. 25, no. 5, pp. 277–280, 2002, doi: 10.1046/j.1365-2605.2002.00355.x.

[34] J. Auger, "Assessing human sperm morphology: Top models, underdogs or biometrics?," *Asian Journal of Andrology*, vol. 12, no. 1, pp. 36–46, 2010, doi: 10.1038/aja.2009.8.

[35] S. G. Goodson *et al.*, "CASAnova: A multiclass support vector machine model for the classification of human sperm motility patterns," *Biology of Reproduction*, vol. 97, no. 5, pp. 698–708, 2017, doi: 10.1093/biolre/iox120.

[36] E. S. Filho, J. A. Noble, M. Poli, T. Griffiths, G. Emerson, and D. Wells, "A method for semi-automatic grading of human blastocyst microscope images," *Human Reproduction*, vol. 27, no. 9, pp. 2641–2648, 2012, doi: 10.1093/humrep/des219.

[37] K. K. Tseng, Y. Li, C. Y. Hsu, H. N. Huang, M. Zhao, and M. Ding, "Computer-assisted system with multiple feature fused support vector machine for sperm morphology diagnosis," *BioMed Research International*, vol. 2013, no. 1, 2013, doi: 10.1155/2013/687607.

[38] A. J. Sahoo and Y. Kumar, "Seminal quality prediction using data mining methods," *Technology and Health Care*, vol. 22, no. 4, pp. 531–545, 2014, doi: 10.3233/THC-140816.

[39] S. K. Mirsky, I. Barnea, M. Levi, H. Greenspan, and N. T. Shaked, "Automated analysis of individual sperm cells using stain-free interferometric phase microscopy and machine learning," *Cytometry Part A*, vol. 91, no. 9, pp. 893–900, 2017, doi: 10.1002/cyto.a.23189.

[40] D. A. Morales *et al.*, "Bayesian classification for the selection of in vitro human embryos using morphological and clinical data," *Computer Methods and Programs in Biomedicine*, vol. 90, no. 2, pp. 104–116, 2008, doi: 10.1016/j.cmpb.2007.11.018.

[41] A. Uyar, A. Bener, and H. N. Ciray, "Predictive modeling of implantation outcome in an in vitro fertilization setting," *Medical Decision Making*, vol. 35, no. 6, pp. 714–725, 2015, doi: 10.1177/0272989X14535984.

[42] S. R. Dash and R. Ray, "Predicting seminal quality and its dependence on life style factors through ensemble learning," *International Journal of E-Health and Medical Communications*, vol. 11, no. 2, pp. 78–95, 2020, doi: 10.4018/IJEHMC.2020040105.

[43] S. Koç, L. Tomak, and E. Karabulut, "A predictive model for the risk of infertility in men using machine learning," *Journal of Urological Surgery*, vol. 9, no. 4, pp. 265-271, 2022, doi: 10.4274/jus.galenos.2022.2021.0134.

[44] C. Krittanawong, H. J. Zhang, Z. Wang, M. Aydar, and T. Kitai, "Artificial intelligence in precision cardiovascular medicine," *Journal of the American College of Cardiology*, vol. 69, no. 21, pp. 2657–2664, 2017, doi: 10.1016/j.jacc.2017.03.571.

[45] M. L. Hoang and N. Delmonte, "K-centroid convergence clustering identification in one-label per type for disease prediction," *IAES International Journal of Artificial Intelligence*, vol. 13, no. 1, pp. 11149-1159, 2024, doi: 10.11591/ijai.v13.i1.pp1149-1159.

[46] R. Miotto, F. Wang, S. Wang, X. Jiang, and J. T. Dudley, "Deep learning for healthcare: Review, opportunities and challenges," *Briefings in Bioinformatics*, vol. 19, no. 6, pp. 1236–1246, 2017, doi: 10.1093/bib/bbx044.

[47] N. Muthukrishnan, F. Maleki, K. Ovens, C. Reinhold, B. Forghani, and R. Forghani, "Brief history of artificial intelligence," *Neuroimaging Clinics of North America*, vol. 30, no. 4, pp. 393–399, 2020, doi: 10.1016/j.nic.2020.07.004.

[48] M. Warpechowski, J. J. Warpechowski, M. Milewski, A. Zańko, and R. Milewski, "The use of the kohonen neural network for comparing the declared and actual state of knowledge regarding reproductive health and the impact of selected lifestyle components on reproductive health," *Studies in Logic, Grammar and Rhetoric*, vol. 66, no. 3, pp. 573–586, 2021, doi: 10.2478/slgr-2021-0033.

[49] L. C. P. F. D. S. Vieira, P. M. D. S. R. Rizol, and L. F. C. Nascimento, "Fuzzy logic and hospital admission due to respiratory diseases using estimated values by mathematical model," *Ciencia e Saude Coletiva*, vol. 24, no. 3, pp. 1083–1090, 2019, doi: 10.1590/1413-81232018243.08172017.

[50] A. Asemi, S. S. B. Salim, S. R. Shahamiri, A. Asemi, and N. Houshangi, "Adaptive neuro-fuzzy inference system for evaluating dysarthric automatic speech recognition (ASR) systems: a case study on MVML-based ASR," *Soft Computing*, vol. 23, no. 10, pp. 3529–3544, 2019, doi: 10.1007/s00500-018-3013-4.

[51] S. Ardabili, A. Mosavi, A. Mahmoudi, T. M. Gundoshmian, S. Nosratabadi, and A. R. V. -Kóczy, "Modelling temperature variation of mushroom growing hall using artificial neural networks," in *Engineering for Sustainable Future*, vol. 101, pp. 33–45, 2020, doi: 10.1007/978-3-030-36841-8_3.

[52] I. Dragović, N. Turajlić, D. Pilčević, B. Petrović, and D. Radojević, "A boolean consistent fuzzy inference system for diagnosing diseases and its application for determining peritonitis likelihood," *Computational and Mathematical Methods in Medicine*, vol. 2015, 2015, doi: 10.1155/2015/147947.

[53] M. A. Ibrisimovic, G. Karli, H. E. Balkaya, M. Ibrisimovic, and MirsadaHukic, "A fuzzy model to predict risk of urinary tract infection," *IFMBE Proceedings*, vol. 62, pp. 289–293, 2017, doi: 10.1007/978-981-10-4166-2_43.

[54] I. Traulsen and J. Krieter, "Assessing airborne transmission of foot and mouth disease using fuzzy logic," *Expert Systems with Applications*, vol. 39, no. 5, pp. 5071–5077, 2012, doi: 10.1016/j.eswa.2011.11.032.

[55] R. Kothandaraman, S. Andavar, and R. S. P. Raj, "A hybrid feature ranking algorithm for assisted reproductive technology outcome prediction," *Brazilian Archives of Biology and Technology*, vol. 65, 2022, doi: 10.1590/1678-4324-2022210605.

[56] J. T. -Niño, A. R. -González, R. C. -Palacios, E. J. -Domingo, and G. A. -Hernandez, "Improving accuracy of decision trees using clustering techniques," *Journal of Universal Computer Science*, vol. 19, no. 4, pp. 484–501, 2013.

[57] V. S. Jiang *et al.*, "The use of voting ensembles and patient characteristics to improve the accuracy of deep neural networks as a non-invasive method to classify embryo ploidy status," *Fertility and Sterility*, vol. 116, no. 3, pp. e155–e156, 2021, doi: 10.1016/j.fertnstert.2021.07.421.

[58] S. K. Tadepalli and P. V. Lakshmi, "An entropy enabled random forest neural network algorithm to grade the reproductive system for efficient early detection of infertility," *5th IEEE International Conference on Cybernetics, Cognition and Machine Learning Applications, ICCCMLA 2023*, pp. 95–100, 2023, doi: 10.1109/ICCCMLA58983.2023.10346771.

[59] S. J. Dyer and M. Patel, "The economic impact of infertility on women in developing countries - a systematic review," *Facts, Views & Vision in ObGyn*, vol. 4, no. 2, pp. 102–109, 2012.

**BIOGRAPHIES OF AUTHORS**

**Ariawan Adimoelja** holds a medical degree from Udayana University, Indonesia in 1994. He also received his Master of Clinical Embryology from National University of Singapore, Singapore in 2000. He is currently as Obstetrician and Gynaecologist in St. Vincentius a Paulo Catholic Hospital, Surabaya. He is also as Social Infertility Consultant in Obstetrics and Gynaecology since 2018. He interests in medical artificial intelligence and applicable digital social medicine. He is also as social medicine doctotal student at Universitas Brawijaya. He can be contacted at email: ariawan_a@student.ub.ac.id.

**Wayan Firdaus Mahmudy** obtained a Bachelor of Science degree from the Department of Mathematics, Brawijaya University in 1995. His Master in Informatics Engineering degree was obtained from the Sepuluh Nopember Institute of Technology, Surabaya in 1999 while a Ph.D. in Manufacturing Engineering was obtained from the University of South Australia in 2014. He is a Professor at Department of Informatics Engineering, Brawijaya University, Indonesia. His research interests include optimization of combinatorial problems and machine learning. He can be contacted at email: wayanfm@ub.ac.id.

**Diva Kurnianingtyas** holds a Bachelor of Computer Science in Informatics Engineering and Doctor in Industrial Engineering, besides several professional certificates and skills. She is currently lecturing with the Department of Informatics Engineering at Universitas Brawijaya, Malang, Indonesia. Her research areas of interest include optimization, artificial intelligence, modeling and simulation, policy analysis and management, and healthcare system. She can be contacted at email: divaku@ub.ac.id.