International Journal of Engineering, Management and Humanities (IJEMH)
Volume 7, Issue 1, Jan.-Feb., 2026 pp: 01-25 ISSN: 2584-2145 www.ijemh.com

A Machine Learning Predictive Model for Estimating Live Birth Success in IVF Patients

Oluwatosin Andrew Odeyale[1], Waheed Azeez Ajani [2], Akinmuda Oluseye Ayobami[3].

[1,2,3]*Department of Computer Science, LeadCity University, Ibadan, Oyo State, Nigeria.*

Date of Submission: 01-01-2026 Date of Acceptance: 08-01-2026

Abstract

In the field of assisted reproductive technology (ART), accurately predicting in vitro fertilization (IVF) outcomes is critical for improving treatment success rates and personalizing patient care. Traditional methods of assessing IVF success often fail to account for the complex, multifactorial nature of infertility and the vast array of variables influencing treatment outcomes. This necessitates advanced solutions like machine learning (ML). Conventional predictive models struggle with non-linear and high-dimensional data, resulting in limited accuracy and applicability. This study aims to leverage advanced ML algorithms to enhance the accuracy of predicting IVF success or failure. Three supervised ML models Random Forest, Decision Tree, and Support Vector Machine (SVM) were trained and validated on a dataset encompassing patient demographics, medical history, treatment modalities, and clinical outcomes. The dataset was divided into training (70%) and testing (30%) subsets. Performance metrics such as accuracy, precision, recall, and area under the curve (AUC) were used to evaluate the models. The Random Forest and Decision Tree classifiers achieved identical accuracy scores of 63% and an AUC of 0.67, demonstrating balanced performance in predicting successful cases. SVM achieved slightly lower accuracy (62%) but exhibited higher recall (81%) for successful outcomes. Key findings indicate that patient age significantly influences IVF outcomes, with higher predictive accuracy for younger age groups (18–34 years). The study also highlighted improved outcomes with donor embryos, preimplantation genetic diagnosis (PGD), and tailored ovulation stimulation protocols. Challenges included reduced predictive accuracy for older patients and complex infertility cases. This research confirms that ML models offer enhanced predictive capabilities in IVF outcomes, particularly when incorporating diverse and high-dimensional data. Future work should focus on expanding datasets, integrating genomic and lifestyle variables, and applying advanced techniques like deep learning to refine predictive accuracy further.

Keywords: *Assisted Reproductive Technology, In Vitro Fertilization, Machine Learning (ML), Predictive Modeling, Preimplantation Genetic Diagnosis*

I. Introduction

Infertility is recognized as a significant global health issue, affecting approximately 15% of couples worldwide (Agarwal et al., 2015). It is defined as the inability to conceive after 12 months or more of regular, unprotected intercourse and represents a multifactorial condition that can impact both males and females (Friedrich & Tüttelmann, 2024). Although infertility was once viewed primarily as a female concern, contemporary evidence indicates that male and female factors contribute almost equally. Approximately 30–40% of infertility cases are attributed to male factors such as sperm abnormalities, hormonal imbalances, or genetic issues, while a similar proportion results from female factors like ovulatory disorders, tubal damage, or endometriosis (Shinde et al., 2021). An additional 30% of cases involve both partners, and around 10% remain idiopathic with no definitive cause identified (Liu et al., 2025). Globally, the burden of infertility has shown a steady increase in recent decades among reproductive-aged men and women (Zeng et al., 2025). The prevalence has grown significantly, especially in middle–socio-demographic index regions, reflecting demographic, lifestyle, and environmental influences. For instance, female infertility is most prevalent in the 30–39 age group, with an 84% increase in global cases between 1990 and 2021, while male infertility has also risen by nearly 77% in the same period (Huang et al., 2023; Liu et al., 2025).

In developed countries, the utilization of Assisted Reproductive Technologies (ART), including In Vitro Fertilization (IVF) and Intracytoplasmic Sperm Injection (ICSI), has

become increasingly common, with treatment rates ranging from 4% to 17% (Zeng et al., 2025). However, access to ART remains low in developing countries due to socioeconomic barriers and limited infrastructure. Despite the widespread acceptance of ART, its limitations such as cost, emotional strain, and varying success rates remain significant. IVF success rates typically range from 30% to 40% per cycle, depending on patient age, cause of infertility, and embryo quality (Shinde et al., 2021). Artificial intelligence (AI) has recently emerged as a transformative tool in reproductive medicine, offering new opportunities to enhance ART outcomes. Machine learning models are increasingly applied in embryo selection, sperm analysis, and ovarian stimulation protocols to improve treatment efficiency and success rates (Amini Mahabadi et al., 2024). One of the most notable applications of AI is in embryo selection. Traditionally, embryologists rely on visual morphology to assess embryo quality. However, deep learning algorithms can analyze time-lapse imaging data to predict implantation potential with higher accuracy, reducing subjectivity and human error (AlSaad et al., 2025). AI has also contributed to the personalization of ovarian stimulation protocols, optimizing hormone dosages and minimizing the risk of ovarian hyperstimulation. These predictive models use clinical data such as age, BMI, and hormone levels to tailor treatment strategies (Shen et al., 2024). In sperm selection for ICSI, AI-powered tools evaluate motility, morphology, and DNA integrity, enabling the selection of sperm with the highest fertilization potential (Qaderi et al., 2025). Nevertheless, the integration of AI into reproductive medicine faces challenges related to data heterogeneity, bias, and ethical concerns. Variations in data quality across clinics may limit model reliability and generalizability (Rolfes et al., 2023). Ethical issues such as data privacy, algorithmic transparency, and bias due to non-diverse training datasetsmust also be addressed. While AI shows great promise in enhancing clinical decision-making and improving success rates, it should serve as a complementary tool rather than a replacement for clinician expertise. Fertility specialists remain vital in providing holistic care that integrates medical, psychological, and ethical dimensions of treatment. Continued research, ethical oversight, and standardized multicenter datasets are crucial to ensuring the safe and effective integration of AI into reproductive health care.

1.2 Statement of Problem

In vitro fertilization (IVF) is a widely used assisted reproductive technology (ART), yet its success rates remain inconsistent, causing significant emotional and financial stress for patients. Predicting IVF outcomes is a complex task influenced by multiple factors such as patient age, infertility causes, embryo quality, and treatment protocols (Wang et al., 2022). Traditional predictive models often rely on statistical and clinical assessments, which lack the capacity to capture the intricate, non-linear relationships inherent in IVF datasets (Raef & Ferdousi, 2019). Consequently, these models frequently underperform, particularly in patients with complex infertility conditions, advanced maternal age, or unexplained infertility. Existing approaches also struggle with high-dimensional and interdependent data, leading to suboptimal accuracy and inefficient resource allocation in IVF treatment (Peng et al., 2024). Although machine learning (ML) applications in IVF are expanding, many models still lack sufficient generalizability, interpretability, and clinical validation to be widely adopted. To address these limitations, this study aims to develop and evaluate robust ML models capable of accurately predicting IVF outcomes, thereby improving treatment success rates and optimizing patient-specific care strategies.

1.3 Aim and Objectives of the Study

The study aims to develop a predictive models for estimating live birth success in In Vitro ertilization (IVF) patients using Multivariate clinial and lifestyle data. The specific objectives were to:

i. gather and preprocess an extensive dataset encompassing patient demographics, medical records, treatment specifics, and in vitro fertilisation (IVF) outcomes.
ii. identify relevant features and variables that may influence IVF success rates through exploratory data analysis and domain expertise.
iii. construct machine learning models employing three techniques, namely Decision Tree, Support Vector Machine (SVM), and Random Forest, with the purpose of forecasting the probability of in vitro fertilisation (IVF) outcomes, specifically success or failure.
iv. compare the performance of the models by evaluating their F1 score, accuracy, and precision.

II. Literature Review

2.1 In Vitro Fertilization (IVF)

Infertility is a widespread issue, with approximately 14% of couples globally experiencing difficulties conceiving (Wu et al., 2025). The global prevalence of infertility continues

to rise, underscoring the need for effective and accessible reproductive solutions (Siswanto et al., 2023). One of the most widely adopted methods for addressing infertility is in vitro fertilization (IVF), first successfully performed in 1979, which has since become a cornerstone of assisted reproductive technology (ART) (Bano et al., 2025). IVF involves fertilizing an egg outside the body, followed by the transfer of the resulting embryo into the uterus to achieve pregnancy (Siswanto et al., 2023). The procedure includes several intricate stages, beginning with oocyte and sperm collection, followed by fertilization either through conventional insemination or intracytoplasmic sperm injection (ICSI), in which a single sperm is directly injected into the egg to increase success rates (Charles et al., 2023). Once fertilization occurs, embryos are cultured under controlled laboratory conditions and later transferred into the uterus, with pregnancy confirmed after approximately two weeks (Wu et al., 2025).

Several factors influence the success rate of IVF, including patient age, hormonal levels, lifestyle factors, and embryo quality. Age is a major determinant; women under 35 years tend to have significantly higher success rates compared to older women, primarily due to better egg quality and ovarian reserve (Bano et al., 2025). Hormonal levels particularly luteinizing hormone (LH) and follicle-stimulating hormone (FSH) which play a vital role in follicular development and ovulation, with abnormal levels negatively impacting IVF outcomes (Nádasdi et al., 2025). Lifestyle factors such as smoking, excessive alcohol use, and obesity have been associated with reduced fertility and lower IVF success rates (Veselinović et al., 2025). Also, embryo quality, determined by morphology and chromosomal normality, is a critical predictor of implantation success and live birth outcomes (Wu et al., 2025).

The complexity of IVF treatment has led to the growing integration of machine learning (ML) and artificial intelligence (AI) in reproductive medicine. These technologies can process high-dimensional clinical data, identify hidden patterns, and enhance predictive accuracy for treatment outcomes (Dehghan et al., 2025). ML applications in IVF are being used to improve embryo selection, outcome prediction, and stimulation protocol optimization, offering clinicians data-driven insights to personalize and optimize care (Sadegh-Zadeh et al., 2024). These emerging approaches promise to enhance IVF success rates and reduce emotional, financial, and physical burdens on patients.

2.1.1 Conventional Approaches to Predicting IVF

In vitro fertilization (IVF) is a revolutionary advancement in reproductive medicine that offers hope to couples struggling with infertility. As one of the most commonly used assisted reproductive technologies (ART), IVF encompasses a series of complex medical and laboratory procedures designed to facilitate conception (Wang et al., 2022). Traditionally, predicting IVF success has relied on evaluating patient characteristics, treatment regimens, and laboratory parameters; however, these methods often fail to capture the multifactorial and nonlinear nature of reproductive outcomes (Dehghan et al., 2024). Conventional prediction models in IVF typically consider variables such as age, ovarian reserve, reproductive history, and infertility etiology (Vladimirsky et al., 2023). Age remains one of the strongest predictors of IVF success, as female fertility declines due to diminished oocyte quantity and quality. According to global data, the success rate for women under 35 averages 32%, declining to 25% for those aged 35–37 and dropping below 10% after age 42 (Sadegh-Zadeh et al., 2024). Ovarian reserve, assessed through biomarkers such as anti-Müllerian hormone (AMH), follicle-stimulating hormone (FSH), and antral follicle count (AFC), is another critical determinant higher AMH and AFC values are strongly associated with improved outcomes (Hua et al., 2022). Reproductive history also plays a key role in IVF prognosis. Women with prior successful pregnancies tend to exhibit higher implantation and live birth rates, whereas those with recurrent pregnancy loss may benefit from adjunctive assessments such as preimplantation genetic testing for aneuploidy (PGT-A) and endometrial receptivity analysis (ERA) (Ortiz et al., 2022). The cause of infertility including male factor, tubal occlusion, endometriosis, or polycystic ovary syndrome (PCOS) further influences outcomes (Wan et al., 2025). Male infertility, often related to sperm count, motility, or morphology, is managed using intracytoplasmic sperm injection (ICSI), while PCOS-related infertility may require tailored hormonal stimulation protocols (Zhu et al., 2025). Lifestyle and physiological factors also contribute significantly. Elevated or low body mass index (BMI), smoking, and alcohol use negatively affect reproductive outcomes. Studies confirm that maintaining a healthy BMI and balanced lifestyle can enhance implantation and live birth rates (Veselinović et al., 2025). Similarly, embryo and sperm quality assessed via morphology and motility are key determinants of fertilization success (Raef et al., 2020).

International Journal of Engineering, Management and Humanities (IJEMH)
Volume 7, Issue 1, Jan.-Feb., 2026 pp: 01-25 ISSN: 2584-2145 www.ijemh.com

Despite their clinical value, traditional IVF prediction models remain limited. They often rely on retrospective, static datasets that fail to reflect ongoing advances in IVF procedures and technologies (Liu et al., 2024). Moreover, interactions between biological, clinical, and environmental factors are nonlinear and complex, making traditional regression-based models insufficiently robust.

To overcome these limitations, recent studies have introduced machine learning (ML) as a powerful predictive approach. ML algorithms can process high-dimensional clinical and laboratory data, uncover hidden relationships, and generate highly accurate, personalized predictions. Hybrid and ensemble ML models such as random forests, support vector machines (SVMs), and extreme gradient boosting (XGBoost) have demonstrated superior predictive accuracy compared to traditional methods, achieving AUC scores exceeding 0.80 in IVF outcome forecasting (Sadegh-Zadeh et al., 2024; Dehghan et al., 2024). These models not only improve embryo selection and ovarian stimulation optimization but also enhance clinical decision-making, offering a transformative step toward data-driven fertility care.

2.1.2 Machine Learning

Machine learning (ML) is one of the most transformative fields of the 21st century, acting as a cornerstone of artificial intelligence (AI). It involves the use of algorithms and statistical models to enable machines to improve their performance on a task through experience, without being explicitly programmed[54]. ML has revolutionized multiple sectors, including healthcare, finance, transportation, and entertainment, by enabling systems to process large datasets, identify patterns, and make data-driven predictions or decisions

2.1.2.1 Types of Machine Learning

Machine learning is broadly classified into three main categories: supervised learning, unsupervised learning, and reinforcement learning (**Udousoro, 2020)**. Each type is tailored to specific types of tasks and datasets.

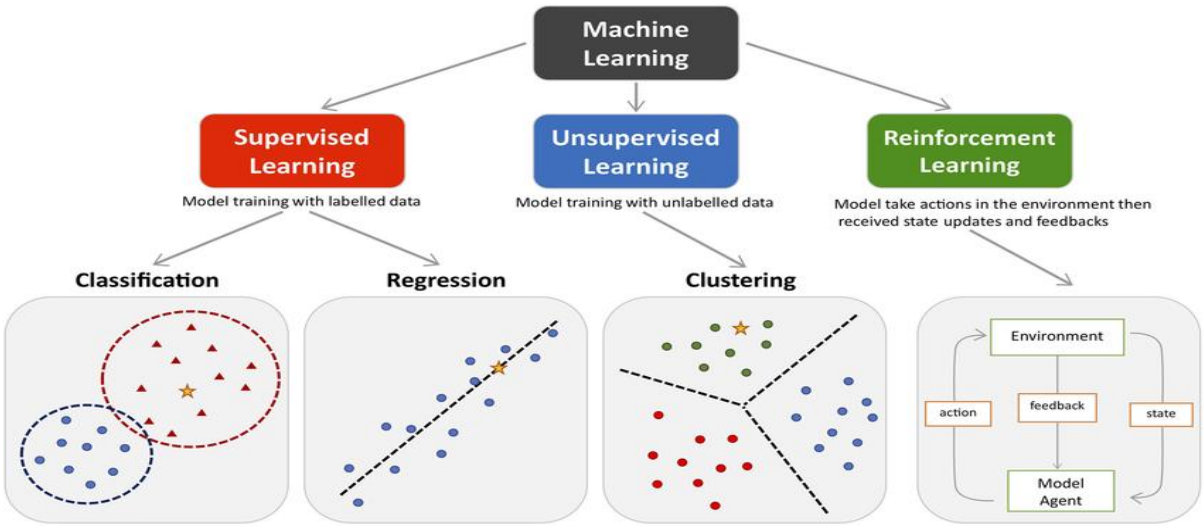

Figure 1: Types of Machine Learning (Mustapha et al., 2024**).**

Supervised Learning involves training a model on labeled data, where each input is paired with a known output (Doshi, 2024). The model learns to map inputs to outputs and is later tested on unseen data. Applications include image recognition, spam detection, and regression tasks like predicting house prices. Algorithms such as linear regression, decision trees, and support vector machines (SVMs) are commonly used for supervised learning. Unsupervised Learning deals with unlabeled data, where the model is tasked with finding hidden patterns or structures (Valkenborg et al., 2023). Clustering and dimensionality reduction are the two main types of unsupervised learning. Algorithms like k-means clustering and principal component analysis (PCA) are widely used. Applications include market segmentation, anomaly detection, and recommendation systems (Ijibadejo & Hashim, 2024). Reinforcement Learning is a reward-based learning approach where an agent interacts with an environment to maximize cumulative rewards (Canese et al., 2021). It is commonly used in

robotics, game development, and autonomous vehicles. Reinforcement learning algorithms, such as Q-learning and deep reinforcement learning, enable systems to make decisions in complex, dynamic environments (Akalin & Loutfi, 2021).

2.1.2.2 Supervised Learning Algorithm

Supervised Learning Algorithm is the primary emphasis of this study. With supervised learning, data from previously completed predictions are used to train a model for future predictions. Supervised learning can be broken down into two categories, regression and classification, depending on whether the output value is continuous or discrete (El Mrabe et al., 2021). Regression (both linear and logistic), Naive Bayes, K-nearest neighbour, random forest, support vector machines, decision trees, linear classifiers, and neural networks are all examples of supervised learning algorithms (Josephine et al., 2021). Predictive models for in vitro fertilisation (IVF) results can be built via a wide variety of Machine Learning algorithms. Logistic regression, decision trees, random forests, support vector machines (SVM), artificial neural networks (ANN), and gradient boosting techniques like XGBoost and LightGBM are all examples of popular Machine Learning algorithms (Kavzoglu & Teke, 2022). Decision Trees, Support Vector Machines, and the Random Forest Classifier are the supervised learning classification algorithms that will be compared and contrasted in this study.

2.1.3 Machine Learning in IVF Prediction

Machine learning (ML) is transforming reproductive medicine by enhancing the prediction of in vitro fertilization (IVF) outcomes. The inherent complexity of IVF treatments, which involve multiple interacting variables such as patient characteristics, embryo quality, and treatment protocols, makes ML a valuable tool for analyzing data and improving success rates.

1. Predicting IVF Success Probability

Accurately predicting the likelihood of IVF success is critical for informing patients and designing personalized treatment plans. ML models excel in analyzing diverse patient datasets, including age, hormone levels, medical history, and previous IVF cycles, to predict outcomes (Hariton, 2023). Data show that women under the age of 35 have a success rate of approximately 40%, while those aged 40–42 experience a success rate of around 15% (Barnett-Itzhaki et al., 2020). Traditional statistical models often fail to capture the nonlinear interactions among these factors, whereas machine learning (ML) algorithms such as **artificial neural networks (ANNs)** and **support vector machines (SVMs)** are particularly effective at identifying complex relationships within IVF datasets (Liu et al., 2024). Integrating multiple predictive variables further enhances model accuracy; for instance, combining **age** with hormone biomarkers such as **follicle-stimulating hormone (FSH)** and **anti-Müllerian hormone (AMH)** provides a more refined estimation of ovarian reserve and fertility potential (Ding et al., 2023).Studies have demonstrated that incorporating these biomarkers into ML models increases predictive accuracy by up to 20% compared to conventional statistical methods (Dehghan et al., 2024). These data-driven models provide actionable insights that support clinicians and patients in making informed, personalized decisions about whether to proceed with IVF, optimize treatment protocols, and improve clinical outcomes

2. Improving Embryo Selection

Selecting the healthiest embryos for transfer is one of the most crucial steps in in vitro fertilization (IVF). Traditionally, embryo selection relied on manual visual assessment of morphology, a process that is inherently subjective and prone to human variability. Recent advances in **machine learning (ML)** have transformed this process by integrating data on embryo morphology, chromosomal integrity, and time-lapse imaging to enhance predictive precision (Boucret et al., 2025). In large-scale analyses involving over 10,000 embryos, ML-based models demonstrated up to a **25% increase in implantation success rates** compared to conventional visual grading methods (Tran et al., 2019). **Time-lapse imaging**, which captures dynamic changes during embryonic development, has proven particularly valuable in improving prediction accuracy. These systems continuously monitor developmental milestones, allowing ML algorithms to identify subtle morphokinetic patterns associated with higher implantation potential (Silver et al., 2020). In addition, models such as the **Blastocyst Evaluation Learning Algorithm (BELA)** can detect **chromosomal abnormalities** non-invasively by analyzing imaging and genetic data, aiding in the selection of **euploid embryos** with the highest likelihood of leading to a healthy pregnancy (Rajendran et al., 2023). Moreover, **automated embryo scoring systems** powered by deep learning are now capable of standardizing evaluations and reducing subjectivity. These

International Journal of Engineering, Management and Humanities (IJEMH)
Volume 7, Issue 1, Jan.-Feb., 2026 pp: 01-25 ISSN: 2584-2145 www.ijemh.com

systems combine image recognition and developmental kinetics to provide objective, reproducible embryo assessments.

3. Predicting Ovarian Response to Stimulation

Predicting ovarian response to stimulation is a critical step in in vitro fertilization (IVF), as it determines the number and quality of oocytes retrieved, directly influencing treatment success. Machine learning (ML) models have demonstrated strong potential in forecasting ovarian response by analyzing patient-specific data such as age, body mass index (BMI), hormone levels, and ultrasound measurements. Studies using algorithms like random forests and gradient boosting machines achieved prediction accuracies of up to **85%**, outperforming traditional regression models (Yan et al., 2021). Key predictors identified include **anti-Müllerian hormone (AMH)** and **antral follicle count (AFC)**, both highly correlated with ovarian reserve. Patients with AMH levels below 1.0 ng/mL were found to be three times more likely to exhibit a poor response. Incorporating ultrasound-derived features, such as follicle count and ovarian volume, further enhanced prediction accuracy (Ferrand et al., 2023). ML-based models like LightGBM have also been used to optimize gonadotropin dosing, achieving robust predictive performance ($R^2 = 0.89$) and supporting personalized stimulation strategies (Fan et al., 2025). These data-driven approaches enable clinicians to tailor treatments, minimize ovarian hyperstimulation syndrome (OHSS) risk, and improve IVF efficiency marking a significant advancement toward personalized reproductive medicine.

2.2 Methodological Framework

In this section, the supervised classification algorithm used in this study was presented

1. Decision Tree

Decision trees operate as sequential models that utilize a series of logical tests to evaluate data, making them both intuitive and powerful for various analytical scenarios (Charbuty & Abdulazeez, 2021). A decision tree is structured like a flowchart, comprising nodes and branches that facilitate the step-by-step breakdown of data. The nodes represent specific features or variables within a dataset, while the branches indicate possible values or conditions for those features. Each internal node in the tree represents a decision point based on a particular feature, where a numerical property is compared to a predetermined threshold. The tree's structure allows it to divide datasets into subsets, making it a popular choice for classification tasks. For example, in image processing, a decision tree might classify pixels based on color intensity thresholds to distinguish objects within an image. Similarly, in pattern detection, a DT can evaluate features of the data to group similar patterns into categories (Azam et al., 2023).

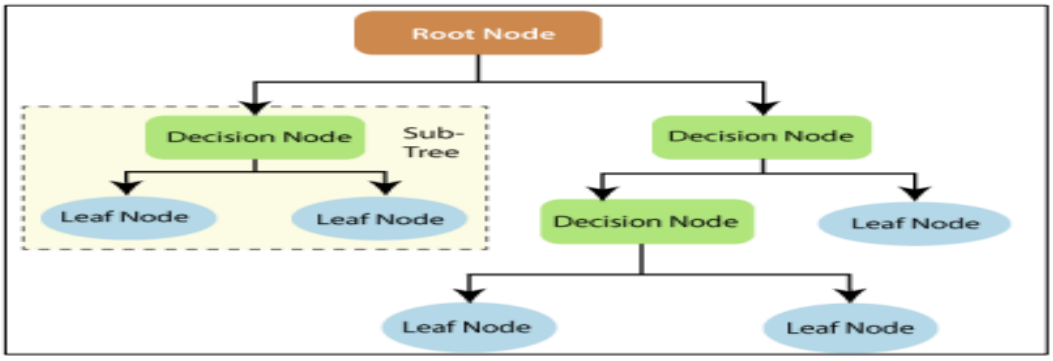

Figure 2: Decision Tree Flow Chart (Ajala et al., 2025).

Decision algorithms frequently make use of measures of entropy and information gain to characterise the degree of randomness or contamination in a dataset. The value of entropy is restricted to be between zero and one. Zero is the best possible value, and any deviation from that is undesirable. That is to say, a lower figure indicates superior performance.

$$\text{Entropy } (S) = \sum_{i=1}^{c} P_i \log 2^{p_i} \quad (2.1)$$

Where P_i is the ratio of the sample number of the subset and i-th attribute value.

2. Support Vector Regression (SVR)

Support Vector Machines (SVMs) are a widely utilized machine learning approach, particularly in

International Journal of Engineering, Management and Humanities (IJEMH)
Volume 7, Issue 1, Jan.-Feb., 2026 pp: 01-25 ISSN: 2584-2145 www.ijemh.com

classification and regression problems (Zhang, 2021). SVMs operate by identifying a hyperplane that separates data points into distinct classes while maximizing the margin between these classes (Sadewo, 2020). This ensures that the classification is not only accurate but also generalizes well to new, unseen data. The foundation of SVMs lies in representing data in an n-dimensional space, where each coordinate corresponds to a value of a specific characteristic or feature of the data point. In a two-dimensional space, data points can be visualized as scattered dots, and the task of the SVM is to find the optimal line (or hyperplane) that separates these points into their respective classes (Geng et al., 2024). For higher-dimensional data, the hyperplane becomes a plane or manifold capable of dividing the data effectively. Through the use of the kernel trick and convex optimization, SVMs transform complex problems into manageable tasks, ensuring robust performance across diverse applications. While challenges such as computational cost and interpretability exist, ongoing advancements in the field continue to enhance the capabilities and applicability of SVMs, solidifying their role as a vital component of machine learning.

$$E[\Pr(error)] \leq \frac{E[number\ of\ support\ vectors]}{[number\ of\ training\ vectors]} \quad (2.2)$$

The decision function will be given as

$$D(x) = w\emptyset(x) + b \quad (2.3)$$

which is the best line that integrates the training data, w and b are parameters of the SVM, and $\emptyset(x)$ is the function which transforms the data into the new M dimension

$$\frac{D(x)}{||w||} \quad (2.4)$$

represents the line, which is the distance betIen item x and the hyperplane.

The parameters of the linear decision function that will maximize M are:

$$w = \sum_k a_k y_k x_k \quad (2.5)$$

$$b = (y_k - w^* x_k) \quad (2.6)$$

3. Random Forest (RF) Algorithms

Random Forest is a supervised ensemble learning method widely used for classification and regression tasks. It leverages the power of multiple decision trees to create robust predictive models (Sekulić et al., 2020). The method operates by generating a large number of individual decision trees, each trained on a random subset of the data, and combining their outputs for prediction. This ensemble approach ensures that the model benefits from the collective intelligence of multiple classifiers, improving both accuracy and generalization. In a Random Forest, each decision tree acts as an independent classifier, with its own set of rules for categorizing or predicting input data (Priyanka & Kumar, 2020).

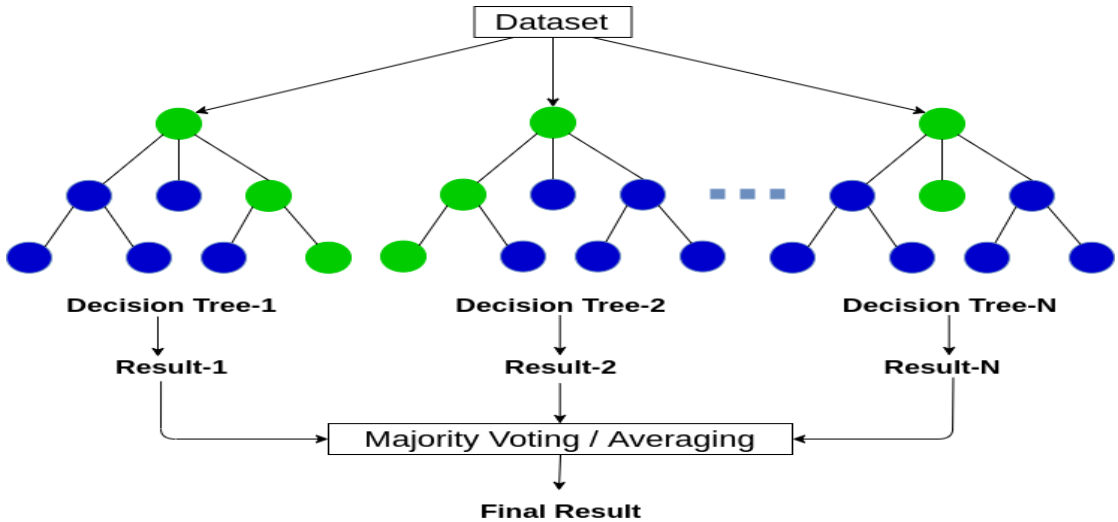

Figure 3: Random Forest Flow Chart (Sarker, 2021).

International Journal of Engineering, Management and Humanities (IJEMH)
Volume 7, Issue 1, Jan.-Feb., 2026 pp: 01-25 ISSN: 2584-2145 www.ijemh.com

The RF algorithm is very efficient, as it handles datasets that contain continuous variables, as categorical variables robustly. An RF classifier contains subsets of various tree classifiers $\{h(x,\Theta_k), k=1,2,...\}$ where the Θ_k are independently and identically distributed random vectors, with each tree being able to specify the modal class at input x.. The performance index, which solely approximates the confidence interval (CI) of the RF model is given as

$$mg(x,y)=av_kI(h_k(x,\Theta_k)=y)-\max_{j\neq y} av_kI(h_k(x,\Theta_k)=j) \quad (2.7)$$

where $I(.)$ denotes an indicator function, and $av(.)$, the average value. It is observed that as the margin increases, the confidence level also increases.

2.3 Review of Relates Studies

Wan et al. (2025) developed a machine learning (ML)-based model to predict clinical pregnancy in endometriosis (EM) patients undergoing fresh embryo transfer (IVF/ICSI). The study retrospectively analyzed 1,752 EM patients treated between 2014 and 2024, evaluating 24 clinical and embryonic characteristics as predictors. Six ML algorithms Naïve Bayes, Logistic Regression, Random Forest, k-Nearest Neighbors, Neural Network, and XGBoost were tested. Recursive feature elimination and tenfold cross-validation were used for feature selection and validation. Male age, fertilization count, and embryo number significantly predicted pregnancy outcomes. The XGBoost model achieved the highest performance with an AUC of 0.764 (training) and 0.622 (testing). Using SHapley Additive Explanations (SHAP), the authors enhanced model interpretability. The study concluded that XGBoost effectively predicts pregnancy likelihood in EM patients after IVF/ICSI, offering a clinically interpretable and reliable decision-support tool.

Sadegh-Zadeh et al. (2024) explored advanced ML paradigms to improve IVF success rate prediction. The study integrated multiple datasets (2010–2018) and tested several algorithms including Logistic Regression, Gaussian Naïve Bayes, Support Vector Machines (SVM), Multi-Layer Perceptron (MLP), k-Nearest Neighbors (KNN), and ensemble models such as Random Forest, AdaBoost, LogitBoost, and RUSBoost. Among these, LogitBoost demonstrated the best accuracy at 96.35%, outperforming other models. Important predictors included patient demographics, infertility causes, and treatment protocols. The study emphasized the advantages of ensemble learning in handling heterogeneous data and improving predictive precision. Results highlighted the importance of collaboration between data scientists and gynecologists in enhancing IVF personalization and treatment planning.

Dehghan et al. (2024) compared multiple ML algorithms for predicting IVF success and introduced Genetic Algorithm (GA) feature selection to improve model robustness. The study evaluated five ML models Random Forest, Artificial Neural Network (ANN), Support Vector Machine (SVM), Recursive Partitioning and Regression Trees (RPART), and AdaBoost—using clinical datasets. GA was applied to identify key predictors, including female age, AMH levels, endometrial thickness, sperm count, and embryo quality. The AdaBoost + GA hybrid model achieved the highest accuracy (89.8%), followed by Random Forest (87.4%). The inclusion of GA improved model performance across all classifiers by optimizing relevant feature sets.

Nguyen et al. (2024) compared machine learning center-specific (MLCS) models with the US national SART model to predict IVF live birth probabilities across six fertility centers. Using 4,645 first-cycle IVF records (2013–2022), the study validated two MLCS models (MLCS1 and MLCS2) against SART benchmarks. MLCS2 achieved superior results, with a median AUC of 0.75 and an F1 score of 0.74 at a 50% probability threshold, compared to 0.71 for SART. MLCS2 identified 9% more patients likely to achieve live birth without overprediction bias. The authors noted that incorporating center-specific data enhanced model calibration and predictive performance. The study demonstrated that MLCS models could personalize patient counseling and improve IVF cost-success transparency, marking a step toward more localized AI-assisted reproductive care.

Ferrand et al. (2023) developed a Light Gradient Boosting Machine (LightGBM) model to predict the number of oocytes retrieved after ovarian stimulation in IVF. Using clinical data from 11,000 cycles, the model incorporated AMH, AFC, baseline FSH, and BMI as main predictors. The LightGBM algorithm outperformed linear and logistic regression models, achieving an R^2 of 0.89 and excellent calibration. Feature importance analysis identified AMH and AFC as the most significant predictors of ovarian response. The study demonstrated that ML-driven dose optimization can guide clinicians in tailoring gonadotropin protocols,

improving treatment efficiency while minimizing ovarian hyperstimulation risk. This approach supports the clinical application of ML models for personalized ovarian stimulation strategies.

Bar et al. (2025) examined the role of vaginal microbiota and inflammation in predicting IVF success using ML models. This pilot study involved 28 participants undergoing IVF, integrating microbiome sequencing and immune marker data at three cycle time points. Using supervised ML algorithms, the model achieved the highest prediction accuracy at mid-cycle (time point 2). Results revealed that pregnant participants had lower microbial diversity and inflammation, with variations based on infertility type. The ML model identified combined microbial and inflammatory signatures predictive of implantation success.

Yang (2025) developed a nomogram-based ML model to predict live birth after single embryo transfer (SET) following IVF/ICSI. The study analyzed 21,385 IVF/ICSI cases (2020–2023), dividing them into training (70%) and validation (30%) cohorts. Using LASSO regression and multivariable logistic modeling, 11 key predictors were identified, including female age, AMH, BMI, endometrial thickness, sperm morphology, and embryo Gardner score. The model achieved an AUC of 0.897 (training) and 0.866 (validation), indicating strong predictive accuracy. Calibration and decision curve analyses confirmed excellent model performance. The study introduced a clinically applicable nomogram tool for individualized IVF prognosis and treatment planning, demonstrating high potential for integration into clinical workflows.

III. Methodology

3.1 Research Approach

This chapter provides a summary of the methodology used in this research to estimate the likelihood of a successful IVF. Data collection, preprocessing, model development, training, validation, and evaluation are all covered in this chapter's review of the research design strategy. The study used a supervised learning strategy, wherein a classification model is trained on a labelled dataset to predict the efficacy of in vitro fertilisation. Specifically, the Decision Tree model, the Random Forest model, and the Support Vector Machine model are used to classify the data. Separating the dataset into training and testing sets and employing cross-validation allows for an accurate assessment of the models' performance

3.2 Research Design

1. Data Collection

This dataset used for this study was sourced from the Human Fertilisation and Embryology Authority (hfea.gov.uk/about-us/data-research). The dataset includes anonymised data from 2015 and 2016. Live birth occurrence is the target variable in the dataset with other details on the treatment cycle in 2015 and 2016. There are two possible outcomes for this variable: successful and unsuccessful. The dataset contain 96 columns and 158520 rows with detailed information related to fertility treatments, such as IVF (in vitro fertilization) and DI (donor insemination). Patient information includes age categories for patients undergoing treatment (e.g., 18-34, 35-37, etc.). Treatment history contains the number of previous cycles and treatments for both IVF and DI, indicating the extent of prior fertility interventions. Pregnancy and birth outcomes includes details on previous pregnancies, birth outcomes, birth weight, weeks of gestation, and congenital abnormalities. Also, multiple birth details information is present for multiple births (e.g., "Heart One," "Heart Two," etc.), tracking various birth outcomes and related data.

2. Data Preprocessing

The initial phase of the analysis involved setting up the necessary Python environment by importing the key libraries: pandas, numpy, matplotlib, seaborn, and scikit-learn. Each of these libraries played a specific role in the data analysis workflow. Pandas was essential for data manipulation, allowing for easy reading, cleaning, and organizing of the data. Numpy facilitated numerical operations and array manipulations, which are fundamental for data processing. For visualization, `matplotlib` and `seaborn` were used to create informative charts and graphs, aiding in the understanding of patterns within the data. Lastly, `scikit-learn` was included to support the machine learning aspects of the analysis, such as data modeling and prediction.

3. Feature Engineering

One of the initial tasks in feature engineering was creating a list of traits termed "infertility_causes_features. The dataset contained both categorical and numerical variables. To make the data suitable for modeling, categorical variables such as "Type of Ovulation Induction" and "Donated Embryo" needed to be transformed into a numerical format.

International Journal of Engineering, Management and Humanities (IJEMH)
Volume 7, Issue 1, Jan.-Feb., 2026 pp: 01-25 ISSN: 2584-2145 www.ijemh.com

Label encoding was employed for these categorical features, converting the categories into numeric values. This step was necessary because most machine learning algorithms require numerical inputs. Label encoding preserved the information about categorical distinctions while making the data compatible with the algorithms.

Normalization was another important step in feature engineering[6]. Given the varying scales of numerical features in the dataset, it was essential to standardize these variables to avoid biasing the model toward features with larger scales. The MinMaxScaler was used to normalize the numerical features, rescaling the values to a range between 0 and 1.

4. Handling Class Imbalance

The RandomUnderSampler was utilized to address class imbalance in the dataset, specifically between the "Live Birth" and "No Live Birth" outcomes. In many medical datasets, such as those related to fertility treatments, the occurrence of one class (e.g., "Live Birth") may be significantly lower than the other, leading to an imbalance. This imbalance can negatively impact the performance of machine learning models, as they tend to be biased toward the majority class. To mitigate this issue, the RandomUnderSampler reduces the number of samples in the majority class ("No Live Birth") to match the size of the minority class ("Live Birth"), resulting in a balanced class distribution

3.3 Model Building and Evaluation

The machine learning process involves training three models: the Random Forest Classifier, Decision Tree Classifier, and Support Vector Machine (SVM) using data from the Human Fertilisation and Embryology Authority. Each model follows a specific pipeline where numerical features are standardized using `StandardScaler`, ensuring that all variables are on a similar scale. This step is crucial for models sensitive to feature scaling, such as SVM, as it prevents size differences from affecting the model's efficiency. The data preprocessing begins by splitting the dataset into features (X) and the target variable (y), with the target representing "Live Birth" occurrences. Label encoding is applied to categorical features as needed. The dataset is then divided into training (80%) and testing sets (20%) to evaluate the model's generalization ability. The scaled training data is used to fit the models, while the testing data assesses their predictive performance.

IV. Result

4.1 Result on Exploratory Data Analysis (EDA)

Extensive exploratory data analysis (EDA) was conducted to gain a deeper understanding of the dataset and uncover patterns that could influence the success of in vitro fertilization (IVF) outcomes. Visual representations were employed to analyze key variables such as "Patient Age at Treatment," "Infertility Types," "Treatment Types," and their relationships with "Live Birth Occurrence." Bar plots and other graphical methods highlighted the age distribution of patients, categorized infertility types, and showcased how various treatment methods correlated with live birth rates

	Patient Age at Treatment	Date patient started trying to become pregnant OR date of last pregnancy	Total Number of Previous cycles, Both IVF and DI	Total Number of Previous treatments, Both IVF and DI at clinic	Total Number of Previous IVF cycles	Total Number of Previous DI cycles	Total number of previous pregnancies, Both IVF and DI	Total number of IVF pregnancies	Total number of DI pregnancies	Total number of live births - conceived through IVF or DI
0	40-42	NaN	2	2	2	0	0	0	0	0
1	45-50	NaN	3	3	3	0	0	0	0	0
2	35-37	NaN	0	0	0	0	0	0	0	0
3	18 - 34	NaN	2	2	2	0	0	0	0	0

Figure 4: Sample Dataset

International Journal of Engineering, Management and Humanities (IJEMH)
Volume 7, Issue 1, Jan.-Feb., 2026 pp: 01-25 ISSN: 2584-2145 www.ijemh.com

1. Distribution of Patient Age at Treatment

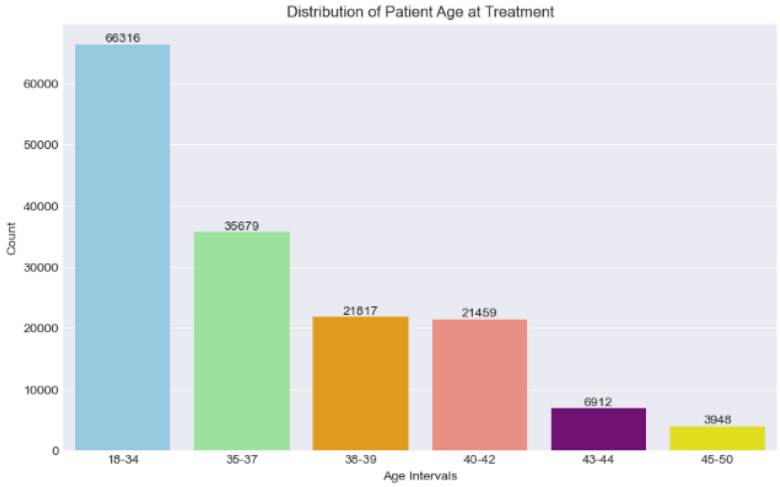

Figure 5: Distribution of Patient Age at Treatment

The bar plot titled Distribution of Patient Age at Treatment illustrates the number of patients undergoing fertility treatments across age intervals. The 15–34 age group shows the highest participation (65,316), reflecting greater fertility potential and IVF success rates among younger women. Patient numbers decline steadily with age: 36–37 years (30,679), 38–39 years (21,817), and 40–42 years (21,459). A sharp decrease occurs after age 43, with only 6,912 in the 43–44 range and 3,483 in the 45–50 group. This trend underscores age-related fertility decline and reduced treatment success, emphasizing the importance of early reproductive planning and timely fertility intervention.

2. The Distribution of Infertility Types

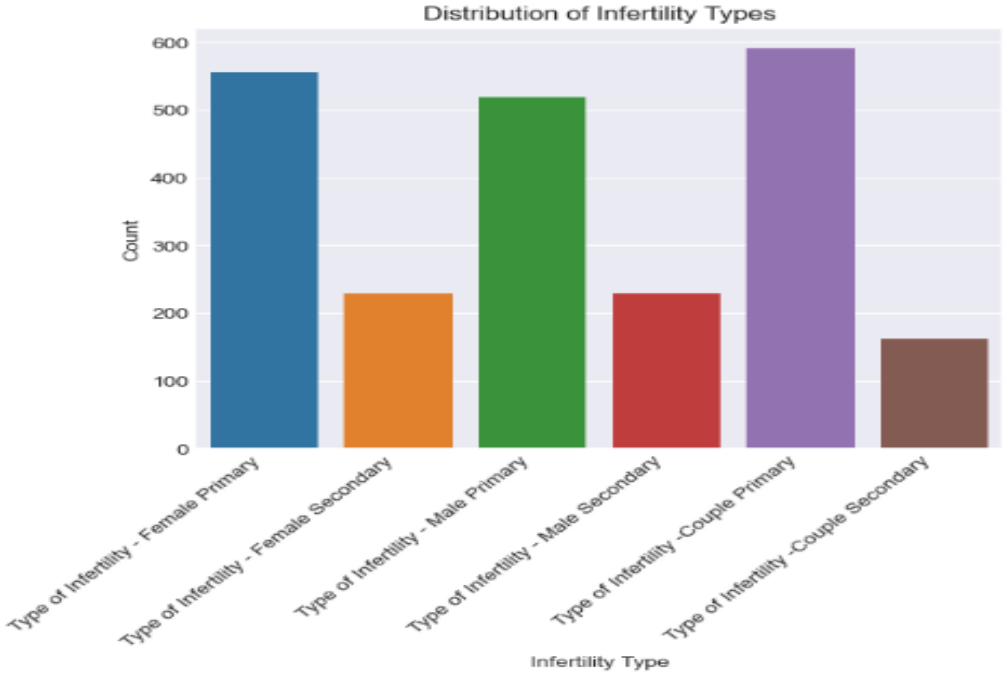

Figure 6: The Distribution of Infertility Types

International Journal of Engineering, Management and Humanities (IJEMH)
Volume 7, Issue 1, Jan.-Feb., 2026 pp: 01-25 ISSN: 2584-2145 www.ijemh.com

The bar plot titled Distribution of Infertility Types illustrates the frequency of different infertility categories among patients. Couple primary infertility is the most prevalent, exceeding 550 cases, indicating that many couples struggle to conceive for the first time. Female primary infertility follows closely at around 525, while male primary infertility is slightly lower, highlighting notable reproductive challenges for both genders. Female and male secondary infertility show reduced counts, near 200, and couple secondary infertility is the least common, below 150. Overall, the data suggests that primary infertility remains a greater concern, emphasizing the importance of early fertility assessment and intervention.

3. The Distribution of Infertility Cause

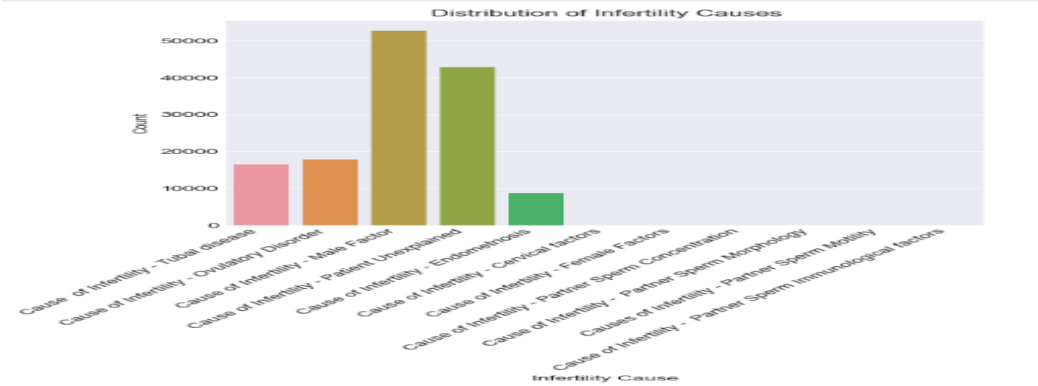

Figure 7: The Distribution of Infertility Cause

The bar plot titled "Distribution of Infertility Causes" highlights the primary reasons for infertility among patients. The most significant cause is male factors, with the highest count exceeding 50,000. This indicates that male infertility plays a major role in reproductive challenges, likely involving issues such as low sperm count, poor motility, or abnormal morphology. The second leading cause is ovulatory disorders, with a count just above 45,000. This reflects a large number of women experiencing irregular or absent ovulation, which is a major barrier to conception. Tubal issues and unexplained infertility follow, each with approximately 15,000 cases, suggesting that blocked or damaged fallopian tubes, as well as undetermined factors, are significant contributors. Lastly, other causes like uterine, cervical, and combined factors show relatively lower frequencies, with counts below 10,000. The plot demonstrates that male factors and ovulatory disorders dominate the infertility spectrum, underscoring the importance of targeted diagnosis and treatment strategies.

4. Live Birth Ocurrence by Treatment Type

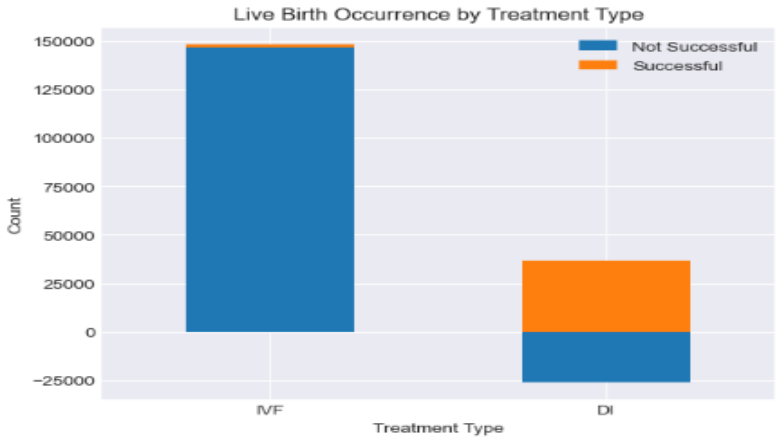

Figure 8: Live Birth Ocurrence by Treatment Type

International Journal of Engineering, Management and Humanities (IJEMH)
Volume 7, Issue 1, Jan.-Feb., 2026 pp: 01-25 ISSN: 2584-2145 www.ijemh.com

The bar chart titled illustrates the success and failure rates for two types of fertility treatments: IVF (In Vitro Fertilization) and DI (Donor Insemination). The chart shows that for IVF, a significantly large number of treatments were not successful, as indicated by the substantial blue portion of the bar. While there is a noticeable count of successful outcomes, the overall success rate remains proportionally lower compared to the total number of treatments. In contrast, DI treatments show a smaller total number of cases, but the proportion of successful outcomes (orange segment) is higher relative to unsuccessful attempts. This indicates that while DI is less frequently performed compared to IVF, it demonstrates a better success ratio. The chart highlights the greater overall reliance on IVF for achieving live births but also underscores its lower success rate. Conversely, DI offers a relatively promising success rate despite fewer recorded treatments. This analysis suggests the need for ongoing optimization of IVF treatments to improve outcomes.

5. Impact of Stimulation Used on Live birth Occurrence

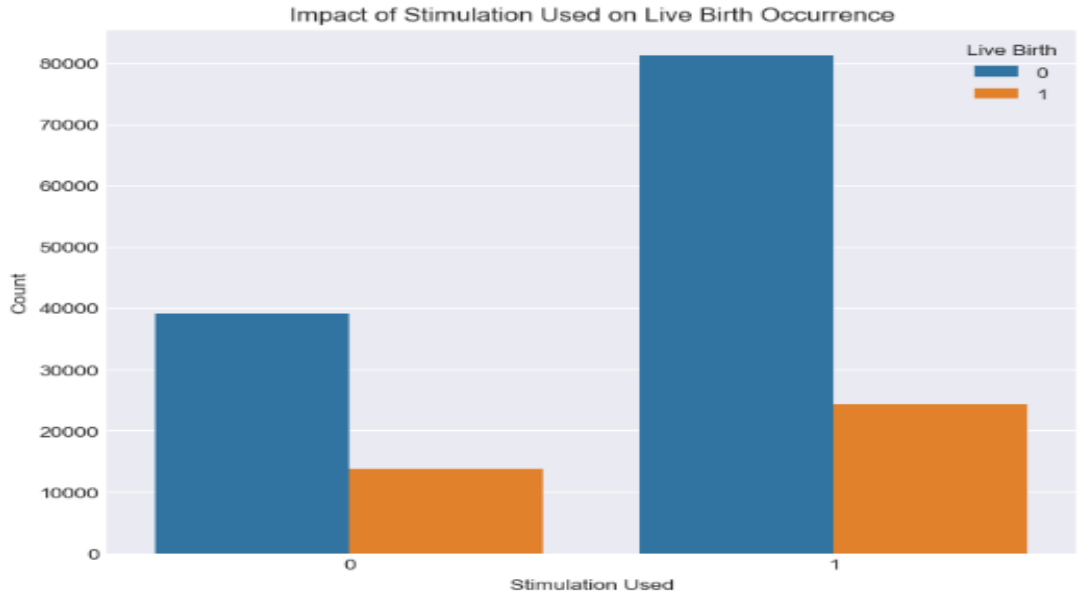

Figure 9 : Impact of Simulation Used on Live Birth Occurrence

The bar chart shows the relationship between the use of stimulation and live birth outcomes. The two categories of stimulation are labeled 0 (No Stimulation) and 1 (Stimulation Used). Live birth outcomes are represented by two colors: blue for unsuccessful (0) and orange for successful (1). For cases where no stimulation was used (0), the chart indicates that the majority of treatments did not result in live births, as seen in the large blue segment. A smaller portion of the treatments, represented by the orange bar, resulted in live births. In contrast, where stimulation was used (1), the total count of treatments is significantly higher. However, while unsuccessful outcomes (blue) still dominate, the proportion of successful live births (orange) is also visibly higher compared to non-stimulated cases.

International Journal of Engineering, Management and Humanities (IJEMH)
Volume 7, Issue 1, Jan.-Feb., 2026 pp: 01-25 ISSN: 2584-2145 www.ijemh.com

6. Examination of Donor Embryos on Live Birth Occurrence

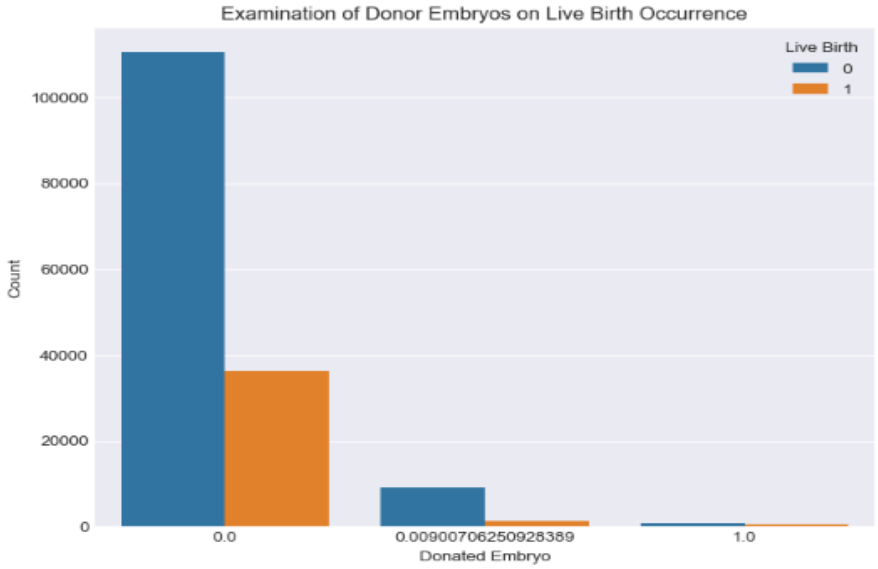

Figure 10: Examination of Donor Embryos on Live Birth Occurrence

The bar chart illustrates the relationship between the use of donor embryos and live birth outcomes. The x-axis categorizes data into 0.0 (no donor embryos), 0.009… (anomalous data or extremely low usage), and 1.0 (donor embryos used). The live birth outcomes are divided into 0 (Not Successful), represented by blue bars, and 1 (Successful), represented by orange bars. For cases where no donor embryos were used (0.0), the chart shows a very high count of unsuccessful outcomes, as indicated by the large blue bar. While a significant number of successful live births are also observed, the proportion of successful cases remains relatively low compared to failures. In the category where donor embryos were used (1.0), the count of treatments is significantly lower, yet the successful live birth outcomes are almost equal to the unsuccessful ones. This suggests that while donor embryo usage is less frequent, its success rate appears comparatively promising, highlighting its potential effectiveness. The middle category (0.009…) appears to reflect outlier or incomplete data, contributing minimally to the analysis.

7. Live birth based on Egg Source

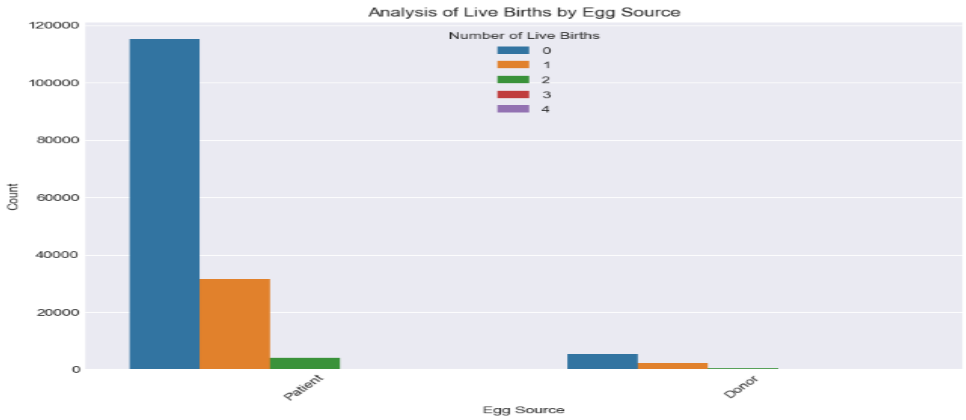

Figure 11: Analysis of Live Births by Egg Source

The bar chart explores the relationship between egg source (patient's own eggs versus donor eggs) and the number of live births. The x-axis represents the two egg sources Patient and

International Journal of Engineering, Management and Humanities (IJEMH)
Volume 7, Issue 1, Jan.-Feb., 2026 pp: 01-25 ISSN: 2584-2145 www.ijemh.com

Donor, while the y-axis shows the count of live births categorized into numbers ranging from 0 to 4. For cases where the patient's own eggs were used, the vast majority of treatments were unsuccessful, as evidenced by the large blue bar exceeding 120,000. A smaller proportion of patients achieved one live birth (orange), with the counts dropping further for two or more live births. This highlights a significant challenge in achieving success when using the patient's own eggs, especially for older individuals or those with diminished ovarian reserves. In contrast, the use of donor eggs shows a significantly smaller total count, but the proportion of successful live births is higher relative to unsuccessful cases. While unsuccessful outcomes still dominate, the orange bar (one live birth) is more prominent compared to the patient's own egg results. This trend underscores the effectiveness of donor eggs in improving live birth outcomes, particularly in cases where the patient's eggs are less viable.

4.2 Model Building

This section will discuss the outcomes of the machine learning models for estimating the likelihood of IVF success and failure. Three models were used the random forest classifier, the decision tree classifier, and the Support Vector Machine. Using precision, recall, and F1-score, the efficacy of the models are analysed. The Random Forest Classifier, Decision Tree Classifier, and Support Vector Machine were all trained with data from the Human Fertilisation and Embryology Authority. The data used in this study can be downloaded for free.

1. Random Forest Classifier

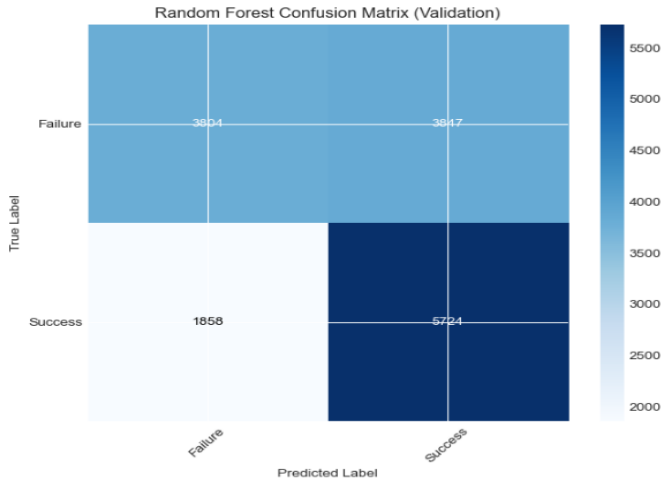

Figure 11: Confusion Matrix for Random Forest Classifier

In the top-left quadrant, 3,854 instances were correctly predicted as Failure (True Negative), indicating the model's ability to identify unsuccessful outcomes accurately. The top-right quadrant shows 3,617 instances where the model incorrectly predicted Success when the actual result was Failure (False Positive), reflecting a notable error rate in overestimating successful outcomes. The bottom-left quadrant contains 1,858 instances where the model incorrectly predicted Failure when the actual outcome was Success (False Negative). This reflects missed opportunities where successful outcomes were not correctly identified. Lastly, the bottom-right quadrant displays 5,734 instances where the model correctly predicted Success (True Positive), showcasing the model's ability to accurately classify successful outcomes.

International Journal of Engineering, Management and Humanities (IJEMH)
Volume 7, Issue 1, Jan.-Feb., 2026 pp: 01-25 ISSN: 2584-2145 www.ijemh.com

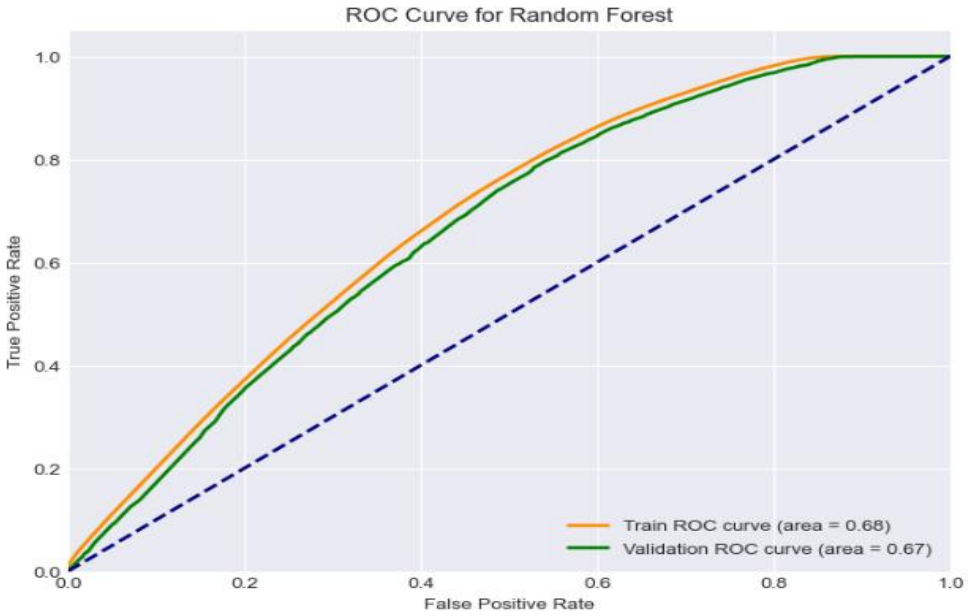

Figure 12: ROC Curve for Random Forest Classifier ROC Curve Training (AUC = 0.68)

The ROC Curve for Random Forest evaluates the model's performance in distinguishing between positive and negative classes, with the True Positive Rate (TPR) plotted against the False Positive Rate (FPR). Two ROC curves are presented: one for the training set (orange line) and one for the validation set (green line). The area under the curve (AUC) for the training ROC curve is 0.68, while the AUC for the validation ROC curve is 0.67. These values indicate that the Random Forest model has moderate predictive power. The current AUC scores suggest the model can perform better than random guessing but still falls short of optimal performance. Both curves follow a similar trajectory, suggesting that the model generalizes reasonably well without significant overfitting. The slight drop in AUC for the validation set (0.67) compared to the training set (0.68) indicates minor overfitting. The ROC curve's gradual upward slope suggests that the model struggles to achieve a high True Positive Rate without also increasing the False Positive Rate.

Classification Report for Random Forest Classifier

Classification Report for Random Forest (Validation):

	precision	recall	f1-score	support
0	0.67	0.50	0.57	7651
1	0.60	0.75	0.67	7582
accuracy			0.63	15233
macro avg	0.63	0.63	0.62	15233
weighted avg	0.64	0.63	0.62	15233

Figure 13: Classification Report for Random Forest

The overall accuracy of the model is 63%, meaning that 63% of predictions were correct. The macro average precision, recall, and F1-score are all 0.63, highlighting balanced performance across both classes

International Journal of Engineering, Management and Humanities (IJEMH)
Volume 7, Issue 1, Jan.-Feb., 2026 pp: 01-25 ISSN: 2584-2145 www.ijemh.com

2. Decision Tree Classifier

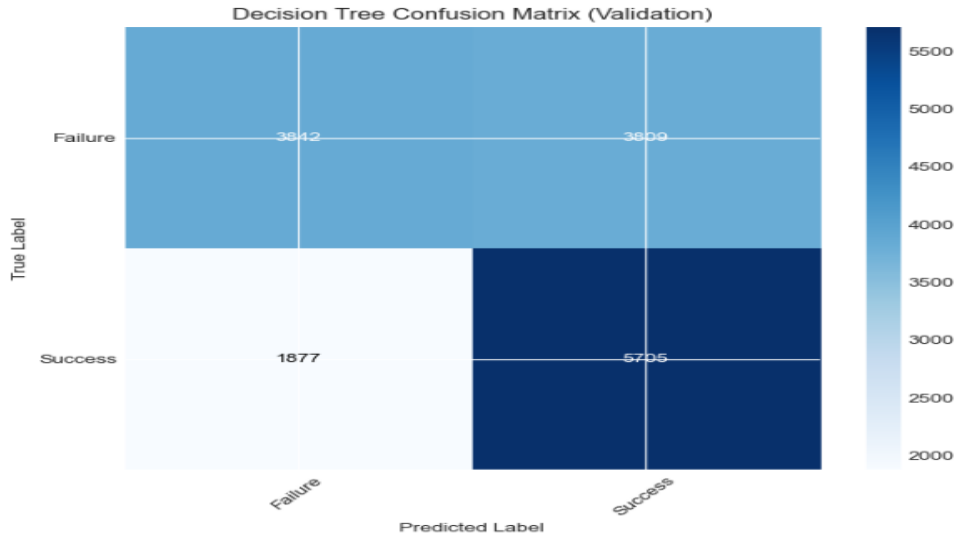

Figure 14: Confusion Matrix For Decision Tree Classifier

Figure 14 presents the performance of a Decision Tree Classifier in predicting two classes: Failure and Success. The top-left quadrant represents the True Negatives, where 3,842 failure cases were correctly predicted as "Failure." The top-right quadrant reflects False Positives, with 3,869 failure cases misclassified as "Success." This indicates a moderate error rate where the model incorrectly predicts success when the actual outcome is failure. The bottom-left quadrant shows False Negatives, where 1,877 success cases were misclassified as "Failure." These are instances where the model fails to identify actual successes. Lastly, the bottom-right quadrant represents the True Positives, where 5,785 success cases were correctly predicted as "Success." This demonstrates the model's ability to identify successful outcomes effectively. The distribution shows that the model performs better at predicting Success (True Positives) compared to predicting Failure (True Negatives), although misclassifications remain significant.

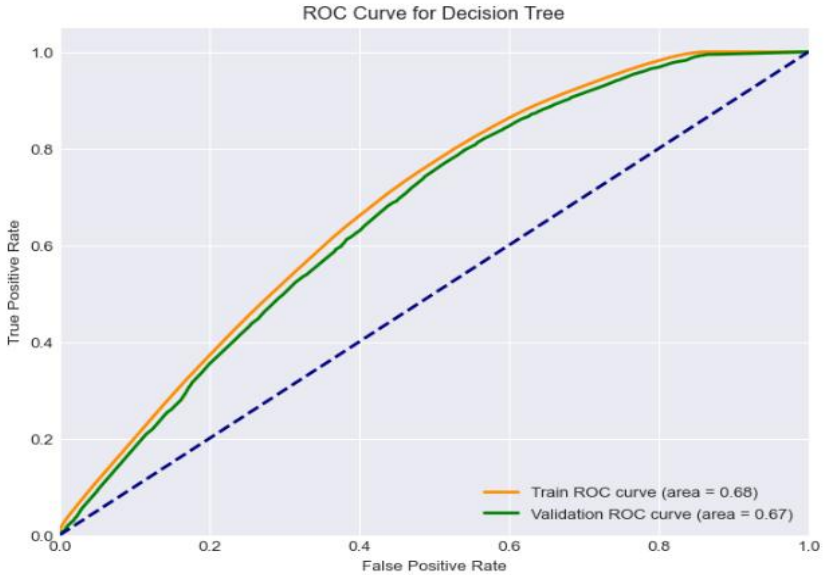

Figure 15: ROC Curve for the Decision Tree Classifier with the Provided AUC Values ROC Curve Training (AUC = 0.68)

International Journal of Engineering, Management and Humanities (IJEMH)
Volume 7, Issue 1, Jan.-Feb., 2026 pp: 01-25 ISSN: 2584-2145 www.ijemh.com

The orange curve represents the ROC curve for the training set, with an AUC (Area Under the Curve) of 0.68, while the green curve corresponds to the validation set, with an AUC of 0.67. These AUC values suggest that the Decision Tree model has moderate discriminative ability. An AUC value of 0.5 indicates random performance, while a value closer to 1.0 reflects perfect classification. The curves for both training and validation sets follow a similar pattern, indicating that the model generalizes relatively well across unseen data without significant overfitting. However, the curve's gradual rise and the gap from the optimal top-left corner suggest that the model struggles to achieve high TPR without increasing the FPR. This limitation reflects the trade-offs in correctly predicting positive outcomes while minimizing false positives. The performance, as measured by the AUC, indicates that while the Decision Tree model is capable of identifying patterns in the data, its ability to separate successful outcomes from failures remains moderate.

Classification Table for Decision Tree Classifier

Classification Report for Decision Tree (Validation):

	precision	recall	f1-score	support
0	0.67	0.50	0.57	7651
1	0.60	0.75	0.67	7582
accuracy			0.63	15233
macro avg	0.64	0.63	0.62	15233
weighted avg	0.64	0.63	0.62	15233

Figure 16: Classification Table for Decision Tree Classifier

The model's overall accuracy is 63%, reflecting that 63% of predictions were correct. The macro average for precision, recall, and F1-score is 0.64, 0.63, and 0.62, respectively, indicating balanced performance across both classes.

3. Support Vector Machine (SVM) Classifier

Confusion Matrix for Support Vector Machine

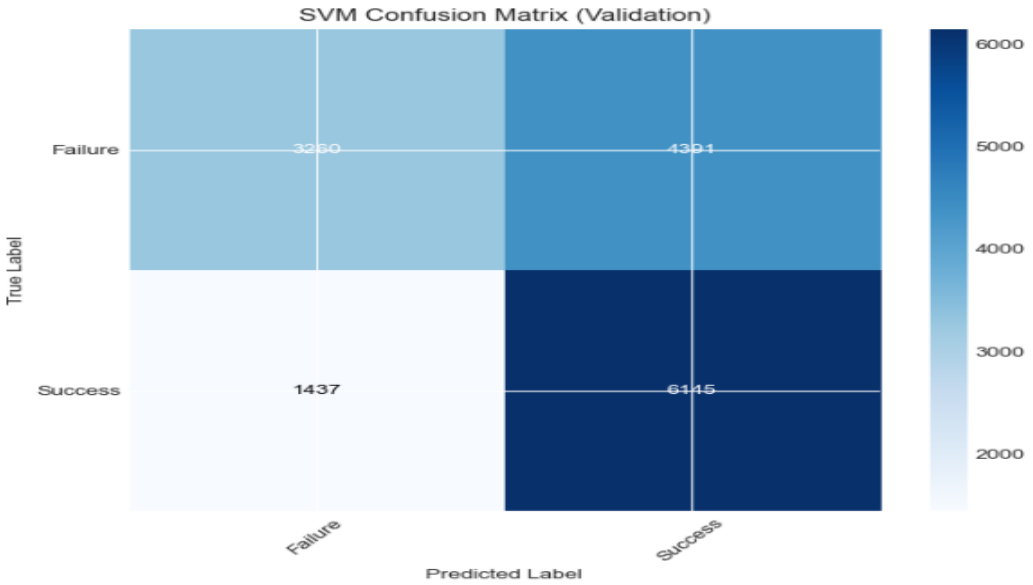

Figure 17: Confusion Matrix for Support Vector Machine

International Journal of Engineering, Management and Humanities (IJEMH)
Volume 7, Issue 1, Jan.-Feb., 2026 pp: 01-25 ISSN: 2584-2145 www.ijemh.com

The matrix contains four sections that show the model's true predictions and errors. In the top-left quadrant, there are 3,850 True Negatives, representing failure cases that were correctly predicted as "Failure." The top-right quadrant contains 4,331 False Positives, where the model incorrectly predicted "Success" for actual failure cases. This indicates that the SVM has a tendency to overestimate success, resulting in misclassifications.

In the bottom-left quadrant, there are 1,437 False Negatives, where successful cases were incorrectly predicted as "Failure." This error represents missed opportunities where the model failed to identify true successes. The bottom-right quadrant includes 6,145 True Positives, showing that the model correctly predicted "Success" for a significant number of cases. The confusion matrix reveals that the SVM classifier performs reasonably well, as evidenced by the large number of True Positives and True Negatives. However, the presence of 4,331 False Positives suggests that the model struggles to distinguish between actual failures and successes, potentially lowering its precision for predicting success. Similarly, the 1,437 False Negatives highlight areas where the model fails to identify successful outcomes accurately.

ROC Curve for Support Vector Machine

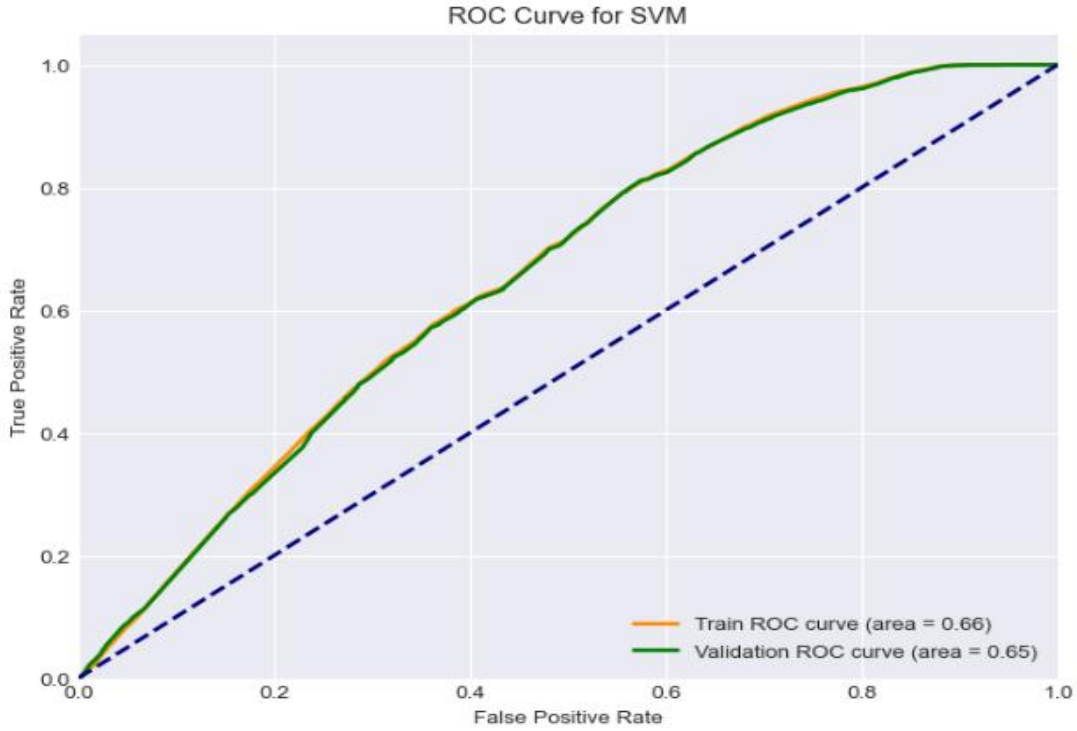

Figure 18: ROC Curve for Support Vector Machine

From figure 18, the orange line represents the ROC curve for the training set, with an AUC (Area Under the Curve) value of 0.66, while the green line represents the ROC curve for the validation set, with an AUC value of 0.65. These values suggest moderate discriminative ability of the SVM model. An AUC of 0.5 indicates random guessing, while values closer to 1.0 reflect a highly accurate model. Both ROC curves for training and validation follow a similar trajectory, indicating that the model generalizes reasonably well and does not suffer from significant overfitting. However, the model struggles to achieve a high True Positive Rate without also increasing the False Positive Rate. This gradual and steady rise of the ROC curves shows that the SVM model has difficulty in confidently distinguishing between positive and negative classes. The ROC curve analysis reveals that the SVM classifier has only moderate performance, with AUC values of 0.66 and 0.65 for the training and validation sets, respectively.

International Journal of Engineering, Management and Humanities (IJEMH)
Volume 7, Issue 1, Jan.-Feb., 2026 pp: 01-25 ISSN: 2584-2145 www.ijemh.com

Classification Table for Support Vector Machine

Classification Report for SVM (Validation):

	precision	recall	f1-score	support
0	0.69	0.43	0.53	7651
1	0.58	0.81	0.68	7582
accuracy			0.62	15233
macro avg	0.64	0.62	0.60	15233
weighted avg	0.64	0.62	0.60	15233

Figure 19: Classification Table for Support Vector Machine

The overall accuracy of the model is 62%, meaning that 62% of all predictions were correct. The macro average for precision, recall, and F1-score is 0.64, 0.62, and 0.60, respectively, indicating balanced performance across the two classes.

4.3 Model Comparison

Table 4.1: Model Comparison

Model	Macro Avg F1-Score	Precision (Class 0)	Recall (Class 0)	Precision (Class 1)	Recall (Class 1)
Random Forest	0.62	0.67	0.50	0.60	0.75
Decision Tree	0.62	0.67	0.50	0.60	0.75
SVM	0.60	0.69	0.43	0.58	0.81

The Model Comparison table summarizes the performance of the three machine learning models Random Forest, Decision Tree, and Support Vector Machine (SVM) using key metrics: Macro Average F1-Score, Precision, and Recall for both classes (0 and 1). The Random Forest and Decision Tree models have identical performance metrics. Both achieve a Macro Average F1-Score of 0.62, indicating balanced performance across both classes. For class 0 (Failure), they produce a Precision of 0.67 and Recall of 0.50, meaning they correctly classify 67% of predicted failures but only capture 50% of actual failures. For class 1 (Success), both models perform better, with a Precision of 0.60 and a significantly higher Recall of 0.75, reflecting their stronger ability to identify successful outcomes. The SVM model has a slightly lower Macro Average F1-Score of 0.60, showing slightly weaker overall performance. However, it achieves the highest Precision for Class 0 (0.69), indicating that when SVM predicts a failure, it is more often correct compared to the other models. Despite this, its Recall for Class 0 (0.43) is the lowest, meaning it misses a large portion of actual failures. In contrast, SVM excels in identifying successes (Class 1), with a Recall of 0.81, the highest among the three models. This means SVM captures a greater proportion of true successes but at the cost of misclassifying failures. The accuracy estimates for the various models are all close to 0.62, but the Decision Tree model seems to perform evenly across both classes. Decision Trees are naturally interpretable, making it simple to visualise decision rules and feature importance, which may be helpful when trying to make sense of the model's predictions.

4.4 Discussion of Findings

From the results, the comparative evaluation of the three models reveals notable differences in predictive performance. The Decision Tree and Random Forest classifiers achieved a Macro Average F1-Score of 0.62, while the SVM attained a slightly lower score of 0.60. These metrics indicate that the models offer moderate accuracy in predicting IVF outcomes. Both the Decision Tree and Random Forest demonstrated balanced performance between precision and recall, effectively identifying successful outcomes (class 1) with a recall of 0.75. In contrast, the SVM excelled in recall for success cases (0.81) but struggled with failure prediction, capturing only 43% of actual failures. The decision tree's ability to provide transparent decision rules makes it particularly useful in clinical settings, where understanding predictions is as important as achieving accuracy.

The result shows patient characteristics, particularly age, as a critical factor in IVF success prediction. The Decision Tree model performed exceptionally well for younger patients aged 18-34, reflecting biological trends where fertility is higher due to robust ovarian reserves and better egg quality. Balanced precision and recall were also observed in the 35-37 and 38-39 age ranges, further reinforcing the relationship between age and IVF success. For older age groups, such as 43-44 and 45-50, the model struggled to predict success, likely due to the

inherent biological limitations. These findings emphasize the need for early intervention and tailored fertility treatments for older patients, including options like egg donation to enhance success rates.

Further, the findings also highlight the complexity of predicting outcomes based on the nature and causes of infertility. The Decision Tree model revealed poorer predictive accuracy for cases involving Female Primary Infertility and certain causes, such as tubal diseases and ovulatory disorders. These conditions often involve multifactorial challenges that are harder to address with standard models, suggesting the need for additional clinical data or refined features. Interestingly, the model demonstrated better predictive performance for cases classified under Male Factor Infertility or Unexplained Infertility. This may indicate that outcomes for these cases are more predictable, likely due to fewer confounding variables or the model's reliance on influential features like sperm motility, morphology, or unexplained diagnosis categories. The result further illustrates that treatment protocols, particularly ovulation stimulation, play a pivotal role in IVF success. When cases were categorized under "Type of Ovulation Induction = Yes but not recorded" or "Specific Treatment Type = IVI," the model provided higher success predictions. This suggests that stimulation protocols and intrauterine insemination (IVI) contribute significantly to positive outcomes, likely by optimizing ovarian response and enhancing the likelihood of fertilization. In contrast, the model struggled to predict failure for specific conditions, such as tubal disorders or ovulatory issues, indicating that treatment optimization is crucial for these categories. The findings underscore the importance of personalized treatment strategies, where stimulation and protocols are tailored to individual patient profiles for better outcomes.

The findings related to donor embryos and preimplantation genetic diagnosis (PGD) highlight their value in improving IVF success rates. The model demonstrated higher predictive accuracy when "Donated Embryo = 0.0", suggesting that patients' own embryos, when viable, offer promising results. However, donor embryos still showed relatively high success probabilities, particularly for patients with diminished ovarian reserves or advanced maternal age. The correlation between "PGD = 0.016…" and higher success rates further highlights the importance of embryo selection. PGD, which involves screening embryos for genetic abnormalities, improves implantation success by ensuring the transfer of healthy embryos. Additionally, the comparative analysis of models reveals that while SVM achieves superior recall for predicting successes, it struggles with false positives and fails to identify actual failures effectively. In contrast, the Decision Tree offers balanced precision and recall across both classes, making it a suitable choice for clinical application. Its interpretability allows clinicians to understand and trust the model's predictions, an essential factor when making decisions that impact patient care. The findings demonstrate that predictive models can play a critical role in assessing IVF success likelihood, helping clinicians identify favorable cases, optimize treatment plans, and manage patient expectations. By integrating patient characteristics, infertility causes, and treatment strategies, these models provide a data-driven framework for improving outcomes in reproductive medicine.

V. Conclusion and Recommendation

The findings of this study underscore the potential of machine learning models to predict the likelihood of success or failure in in vitro fertilisation (IVF) procedures. Among the models tested, the Decision Tree and Random Forest classifiers demonstrated balanced performance, achieving an accuracy of 63% and an AUC of 0.67. These models proved effective in predicting successful outcomes, particularly for younger patients and cases with specific treatment protocols. The Support Vector Machine (SVM) also performed reasonably well, with a slightly lower accuracy of 62%, but it excelled in identifying successful cases with a recall of 81%. However, its limitations in predicting failures highlight the need for model improvements. Key factors influencing IVF outcomes were identified, including patient age, the nature and causes of infertility, and treatment modalities. Younger patients, particularly those in the 18-34 age group, showed the highest likelihood of success, reflecting the biological advantages associated with this cohort. Cases involving male factor infertility and unexplained infertility were more successfully predicted, whereas conditions like female primary infertility, tubal disease, and ovulatory disorders posed challenges for the models. The use of ovulation stimulation and specific protocols, such as intrauterine insemination (IVI), were shown to positively impact success rates, emphasizing the importance of tailored treatment strategies. The study also highlighted the effectiveness of advanced reproductive techniques, such as the use of donor embryos and preimplantation genetic diagnosis (PGD), in

improving outcomes. These methods provided enhanced predictive reliability, particularly in cases involving complex infertility challenges. Based on the findings, it is recommended to incorporate additional clinical features such as genetic profiles, metabolic markers, and lifestyle factors which could improve the models' ability to predict outcomes, particularly for challenging cases like female primary infertility or unexplained infertility.

References

[1]. Agarwal, Aditi A. Mulgund, A. Hamada and M. Chyatte. A unique view on male infertility around the globe. Reproductive Biology and Endocrinology : RB&E, 13 (2015). https://doi.org/10.1186/s12958-015-0032-1.

[2]. Ajala, T. B., ND-Asogu, V. S., & Fatokun, J. (2025). Investigating Vulnerabilities in Modern Network Hardware in Firewalls and Intrusion Detection Systems. 10.5281/zenodo.17235246.

[3]. Akalin, N., & Loutfi, A. (2021). Reinforcement learning approaches in social robotics. *Sensors*, *21*(4), 1292.

[4]. AlSaad, R., Abusarhan, L., Odeh, N., Abd-Alrazaq, A., Choucair, F., Zegour, R., ... & Sheikh, J. (2025). Deep learning applications for human embryo assessment using time-lapse imaging: scoping review. *Frontiers in Reproductive Health*, *7*, 1549642.

[5]. Amini Mahabadi, J., Enderami, S. E., Nikzad, H., & Hassani Bafrani, H. (2024). The Use of Machine Learning for Human Sperm and Oocyte Selection and Success Rate in IVF Methods. *Andrologia*, *2024*(1), 8165541.

[6]. Azam, Z., Islam, M. M., & Huda, M. N. (2023). Comparative analysis of intrusion detection systems and machine learning-based model analysis through decision tree. *Ieee Access*, *11*, 80348-80391.

[7]. Bano, T., Shaheen, M., Shabbir, N., Jadoon, A., Ameen, S., Atta, M., Noor, S., Khattak, M., & S, A. (2025). Clinical Outcomes of In Vitro Fertilization (IVF) Versus Intrauterine Insemination (IUI) in Infertile Patients: Toward Evidence-Based Fertility Planning. Cureus, 17. https://doi.org/10.7759/cureus.92511.

[8]. Bar, O., Vagios, S., Barkai, O., Elshirbini, J., Souter, I., Xu, J., James, K., Bormann, C., Mitsunami, M., Chavarro, J., Foessleitner, P., Kwon, D., Yassour, M., & Mitchell, C. (2025). Harnessing vaginal inflammation and microbiome: a machine learning model for predicting IVF success. NPJ Biofilms and Microbiomes, 11. https://doi.org/10.1038/s41522-025-00732-8.

[9]. Barnett-Itzhaki, Z., Elbaz, M., Butterman, R., Amar, D., Amitay, M., Racowsky, C., Orvieto, R., Hauser, R., Baccarelli, A., & Machtinger, R. (2020). Machine learning vs. classic statistics for the prediction of IVF outcomes. Journal of Assisted Reproduction and Genetics, 37, 2405 - 2412. https://doi.org/10.1007/s10815-020-01908-1.

[10]. Boucret, L., Chabrun, F., Boguenet, M., Reynier, P., Bouet, P., & May-Panloup, P. (2025). Deep-learning model for embryo selection using time-lapse imaging of matched high-quality embryos. Scientific Reports, 15. https://doi.org/10.1038/s41598-025-10531-y.

[11]. Canese, L., Cardarilli, G. C., Di Nunzio, L., Fazzolari, R., Giardino, D., Re, M., & Spanò, S. (2021). Multi-agent reinforcement learning: A review of challenges and applications. *Applied Sciences*, *11*(11), 4948.

[12]. Charbuty, B., & Abdulazeez, A. (2021). Classification based on decision tree algorithm for machine learning. *Journal of applied science and technology trends*, *2*(01), 20-28.

[13]. Charles, D., Lange, M., Ortiz, N., Purcell, S., & Smith, R. (2023). A narrative review of sperm selection technology for assisted reproduction techniques. Translational Andrology and Urology, 13, 2119 - 2133. https://doi.org/10.21037/tau-24-195.

[14]. Dehghan, S., Moghaddasi, H., Rabiei, R., Choobineh, H., Maghooli, K., & Vahidi-Asl, M. (2025). Machine learning in predicting infertility treatment success: A systematic literature review of techniques. Journal of Education and Health Promotion, 14. https://doi.org/10.4103/jehp.jehp_1798_23.

[15]. Ding, T., Ren, W., Wang, T., Han, Y., , W., Wang, M., Fu, F., Li, Y., & Wang, S. (2023). Assessment and quantification of ovarian reserve on the basis of machine learning models. Frontiers in Endocrinology, 14. https://doi.org/10.3389/fendo.2023.1087429.

[16]. Doshi, R. (2024). *Machine Learning: Master supervised and unsupervised learning algorithms with real example*. BPB Publications.

[17]. El Mrabet, M. A., El Makkaoui, K., & Faize, A. (2021, December). Supervised machine learning: A survey. In *2021 4th International conference on advanced communication*

technologies and networking (CommNet) (pp. 1-10). IEEE.

[18]. Fan W, Ye T, Du L, Tian L, Kong H. Development and externally validated prediction model of individualization of FSH starting dose in the depot GnRH agonist protocol for the early follicular phase. Frontiers in Endocrinology. 2025 Mar 27;16:1542736.

[19]. Ferrand T, Boulant J, He C, Chambost J, Jacques C, Pena CA, Hickman C, Reignier A, Fréour T. Predicting the number of oocytes retrieved from controlled ovarian hyperstimulation with machine learning. Human Reproduction. 2023 Oct 1;38(10):1918-26. https://doi.org/10.1093/humrep/dead163

[20]. Friedrich, C., & Tüttelmann, F. (2024). Genetics of female and male infertility. Medizinische Genetik, 36, 161 - 170. https://doi.org/10.1515/medgen-2024-2040.

[21]. Geng, Y., Li, Q., Yang, G., & Qiu, W. (2024). Support Vector Machine. In *Practical Machine Learning Illustrated with KNIME* (pp. 159-184). Singapore: Springer Nature Singapore.

[22]. Hariton, E., Pavlovic, Z., Fanton, M., & Jiang, V. S. (2023). Applications of artificial intelligence in ovarian stimulation: a tool for improving efficiency and outcomes. *Fertility and sterility*, *120*(1), 8-16.

[23]. https://www.hfea.gov.uk/about-us/data-research/

[24]. Hua, L., Zhe, Y., Jing, Y., Fujin, S., Jiao, C., & Liu, L. (2022). Prediction model of gonadotropin starting dose and its clinical application in controlled ovarian stimulation. *BMC Pregnancy and Childbirth*, *22*(1), 810.https://doi.org/10.1186/s12884-022-05152-6

[25]. Huang, B., Wang, Z., Kong, Y., Jin, M., & , L. (2023). Global, regional and national burden of male infertility in 204 countries and territories between 1990 and 2019: an analysis of global burden of disease study. BMC Public Health, 23. https://doi.org/10.1186/s12889-023-16793-3.

[26]. Ijibadejo, O. W., & Hashim, S. M. (2024). Principal Component Analysis for Database Scan Using Hierarchical KNN Model for Telecom Customer Segmentation.

[27]. Josephine, P. K., Prakash, V. S., & Divya, K. S. (2021). Supervised learning algorithms: A comparison. *Kristu Jayanti Journal of Computational Sciences (KJCS)*, 01-12.

[28]. Kavzoglu, T., & Teke, A. (2022). Predictive performances of ensemble machine learning algorithms in landslide susceptibility mapping using random forest, extreme gradient boosting (XGBoost) and natural gradient boosting (NGBoost). *Arabian Journal for Science and Engineering*, *47*(6), 7367-7385.

[29]. Liu, J., Qin, Y., Liu, H., Liu, Y., Yang, Y., Ning, Y., & Ye, H. (2025). Global, regional, and national burden of female infertility and trends from 1990 to 2021 with projections to 2050 based on the GBD 2021 analysis. Scientific Reports, 15. https://doi.org/10.1038/s41598-025-01498-x.

[30]. Liu, L., Liang, H., Yang, J., Shen, F., Chen, J., & Ao, L. (2024). Clinical data-based modeling of IVF live birth outcome and its application. *Reproductive biology and endocrinology*, *22*(1), 76.https://doi.org/10.1186/s12958-024-01253-3

[31]. Liu, R., Bai, S., Jiang, X., Luo, L., Tong, X., Zheng, S., ... & Xu, B. (2021). Multifactor prediction of embryo transfer outcomes based on a machine learning algorithm. *Frontiers in Endocrinology*, *12*, 745039.

[32]. Mustapha, A. A., Udeh, A. S., Ashi, T. A., Sobowale, O. S., Akinwande, M. J., & Oteniara, A. O. (2024). Comprehensive review of machine learning models for sql injection detection in e-commerce. World Journal of Advanced Research and Reviews. 23. 10.30574/wjarr.2024.23.1.2004.

[33]. Nádasdi, B., Vedelek, V., Bereczki, K., Bukva, M., Kozinszky, Z., Sinka, R., Zádori, J., & Vágvölgyi, A. (2025). The Influence of Pre-IVF Day 2 TSH Levels on Treatment Success and Obstetric Outcomes: A Retrospective Single-Center Analysis with Machine Learning-Based Data Evaluation. Journal of Clinical Medicine, 14. https://doi.org/10.3390/jcm14134407.

[34]. Nguyen, E., B, M., Gago, L., Nichols, J., Payne, J., Ripps, B., F, M., Groll, J., F, R., Nowak, L., B, G., B, J., Swanson, T., MSc, C., & MyleneW.M.YaoMDa, .. (2024). Predicting IVF live birth probabilities using machine learning, center-specific and national registry-based models. . https://doi.org/10.1101/2024.06.20.24308970.

[35]. Ortiz, J. A., Morales, R., Lledo, B., Vicente, J. A., Gonzalez, J., García-Hernandez, E. M., ... & Bernabeu, R. (2022). Application of

machine learning to predict aneuploidy and mosaicism in embryos from in vitro fertilization cycles. *AJOG Global Reports*, *2*(4), 100103.

[36]. Peng, J., Geng, X., Zhao, Y., Hou, Z., Tian, X., Liu, X., Xiao, Y., & Liu, Y. (2024). Machine learning algorithms in constructing prediction models for assisted reproductive technology (ART) related live birth outcomes. Scientific Reports, 14. https://doi.org/10.1038/s41598-024-83781-x.

[37]. Priyanka, & Kumar, D. (2020). Decision tree classifier: a detailed survey. *International Journal of Information and Decision Sciences*, *12*(3), 246-269.

[38]. Qaderi, K., Sharifipour, F., Dabir, M., Shams, R., & Behmanesh, A. (2025). Artificial intelligence (AI) approaches to male infertility in IVF: a mapping review. *European Journal of Medical Research*, *30*(1), 246.

[39]. Raef, B., & Ferdousi, R. (2019). A Review of Machine Learning Approaches in Assisted Reproductive Technologies. Acta Informatica Medica, 27, 205 - 211. https://doi.org/10.5455/aim.2019.27.205-211.

[40]. Raef, B., Maleki, M., & Ferdousi, R. (2020). Computational prediction of implantation outcome after embryo transfer. *Health informatics journal*, *26*(3), 1810-1826.doi:10.1177/1460458219892138

[41]. Rajendran, S., Brendel, M., Barnes, J., Zhan, Q., Malmsten, J., Zisimopoulos, P., Sigaras, A., Ofori-Atta, K., Meseguer, M., Miller, K., Hoffman, D., Rosenwaks, Z., Elemento, O., Zaninovic, N., & Hajirasouliha, I. (2023). Automatic Ploidy Prediction and Quality Assessment of Human Blastocyst Using Time-Lapse Imaging. bioRxiv. https://doi.org/10.1101/2023.08.31.555741.

[42]. Rolfes, V., Bittner, U., Gerhards, H., Krüssel, J. S., Fehm, T., Ranisch, R., & Fangerau, H. (2023). Artificial intelligence in reproductive medicine–an ethical perspective. *Geburtshilfe und Frauenheilkunde*, *83*(01), 106-115.

[43]. Sadegh-Zadeh, S., Khanjani, S., Javanmardi, S., Bayat, B., Naderi, Z., & Hajiyavand, A. (2024). Catalyzing IVF outcome prediction: exploring advanced machine learning paradigms for enhanced success rate prognostication. Frontiers in Artificial Intelligence, 7. https://doi.org/10.3389/frai.2024.1392611.

[44]. Sadewo, W., Rustam, Z., Hamidah, H., & Chusmarsyah, A. R. (2020). Pancreatic cancer early detection using twin support vector machine based on kernel. *Symmetry*, *12*(4), 667.

[45]. Sarker, I. H. (2021). Machine learning: Algorithms, real-world applications and research directions. *SN computer science*, *2*(3), 160.

[46]. Sekulić, A., Kilibarda, M., Heuvelink, G. B., Nikolić, M., & Bajat, B. (2020). Random forest spatial interpolation. *Remote Sensing*, *12*(10), 1687.

[47]. Shinde, K., Krovi, N., & Dhutraj, P. (2021). Study of factors contributing to female infertility. International Journal of Reproduction, Contraception, Obstetrics and Gynecology. https://doi.org/10.18203/2320-1770.ijrcog20212340.

[48]. Silver, D., Feder, M., Gold-Zamir, Y., Polsky, A., Rosentraub, S., Shachor, E., Weinberger, A., Mazur, P., Zukin, V., & Bronstein, A. (2020). Data-Driven Prediction of Embryo Implantation Probability Using IVF Time-lapse Imaging. ArXiv, abs/2006.01035.

[49]. Siswanto, J., Faizah, Z., & Dwiningsih, S. (2023). Unraveling the Complexity of Infertility: A Comprehensive Review of Semen Abnormalities, Etiological Factors, and Cutting-edge Approaches in Reproductive Medicine. International Journal of Research Publications. https://doi.org/10.47119/ijrp100140112024599 3.

[50]. Tran, D., Cooke, S., Illingworth, P. J., & Gardner, D. K. (2019). Deep learning as a predictive tool for fetal heart pregnancy following time-lapse incubation and blastocyst transfer. *Human reproduction*, *34*(6), 1011-1018.

[51]. Udousoro, I. C. (2020). Machine learning: a review. *Semiconductor Science and Information Devices*, *2*(2).

[52]. Valkenborg, D., Rousseau, A. J., Geubbelmans, M., & Burzykowski, T. (2023). Unsupervised learning. *American Journal of Orthodontics and Dentofacial Orthopedics*, *163*(6), 877-882.

[53]. Veselinović, A., Kotur-Stevuljević, J., Stojsavljević, A., Bojović-Jović, D., Arsić, A., Paunović, M., & Vučić, V. (2025). Factorial analysis of zinc serum levels, fatty acids, oxidative stress parameters and supplementation on assisted reproductive technology outcome. Journal of Medical Biochemistry, 44, 840 - 853.

https://doi.org/10.5937/jomb0-53112.

[54]. Vladimirsky, G. M., Zhuravleva, M. A., Dashieva, A. E., Korneeva, I. E., & Nazarenko, T. A. (2023). PREDICTING THE OUTCOMES OF IN VITRO FERTILIZATION PROGRAMS USING A RANDOM FOREST MACHINE LEARNING MODEL. *Bulletin of Russian State Medical University*, (6), 64-70.

[55]. Wan, X., Yu, M., Wu, X., Huang, Z., & Tan, J. (2025). Machine learning prediction of clinical pregnancy in endometriosis patients following fresh IVF/ICSI-ET. *European Journal of Medical Research*, *30*(1), 838.

[56]. Wang, C., Kuo, C., Chen, C., Hsieh, Y., & Su, E. (2022). Predicting clinical pregnancy using clinical features and machine learning algorithms in in vitro fertilization. PLoS ONE, 17. https://doi.org/10.1371/journal.pone.0267554.

[57]. Wu, S., Wang, X., Liu, Y., Ren, Y., Zhao, M., Song, H., Shen, H., Wu, Y., Wei, Z., Lu, H., & Li, K. (2025). Predictive models for live birth outcomes following fresh embryo transfer in assisted reproductive technologies using machine learning. Journal of Translational Medicine, 23. https://doi.org/10.1186/s12967-025-07045-6.

[58]. Yang, P. (2025). P-419 Prediction Model for Live Birth After Single Embryo Transfer Following IVF and ICSI Based on Nomogram. Human Reproduction. https://doi.org/10.1093/humrep/deaf097.725.

[59]. Zeng, G., Liu, L., Wang, Y., Yu, J., Wang, H., & Li, F. (2025). Global, regional, and national burden and trends of reproductive-aged male and female infertility from 1990–2021. Frontiers in Endocrinology, 16. https://doi.org/10.3389/fendo.2025.1506229.

[60]. Zhang, D. (2021). Support vector machine. In *Fundamentals of Image Data Mining: Analysis, Features, Classification and Retrieval* (pp. 201-228). Cham: Springer International Publishing.

[61]. Zhu S, Huang Z, Chen X, Jiang W, Zhou Y, Zheng B, Sun Y. Construction and evaluation of machine learning-based prediction model for live birth following fresh embryo transfer in IVF/ICSI patients with polycystic ovary syndrome. Journal of Ovarian Research. 2025 Apr 4;18(1):70.