Journal of Information Technology Management
ISSN-OnLine: 2980-7972
University of Tehran

Hybrid Deep Learning Model for IVF Outcome Prediction from Time-Series Hormonal Data

Shalini B N*

*Corresponding author, Research scholar, School of Computer Science and Engineering, REVA University, Bengaluru, Karnataka, India. E-mail: R23PCT09@reva.edu.in

Lithin Kumble

Associate Professor, School of Computer Science and Engineering, REVA University, Bengaluru, Karnataka, India. E-mail: lithin.k@reva.edu.in

Journal of Information Technology Management, 2025, Vol. 18, Issue 1, pp. 141-157
Published by the University of Tehran, College of Management
doi: https://doi.org/10.22059/jitm.2026.106258
Article Type: Research Paper
© Authors

Received: August 16, 2025
Received in revised form: September 07, 2025
Accepted: December 23, 2025
Published online: January 20, 2026

Abstract

Optimizing results in assisted reproduction requires tailoring the dosage of follicle-stimulating hormone (FSH) during controlled ovarian stimulation (COS), but this is still challenging because of patient variability. Real-time modifications during stimulation are not supported by the majority of current models, which are restricted to static forecasts of starting dosages. This study proposes an advanced AI-driven framework for forecasting hormone dynamics and improving reproductive outcome prediction in IVF cycles. The methodology integrates multi-source clinical data with high-resolution time-series hormone profiles collected from Day 1 to Day 30 of ovarian stimulation. Data pre-processing includes normalization of hormonal values, alignment of temporal and static clinical attributes, and creation of patient-level merged datasets. Feature engineering incorporates daily hormone variations, moving averages, peak detection, and wavelet-based temporal pattern extraction, alongside encoded and normalized clinical parameters. For hormone trend forecasting, a hybrid deep learning model is developed that combines Wavelet Transform for noise reduction with LSTM and Transformer layers for sequential representation learning. The architecture captures short-term hormone fluctuations and long-range temporal dependencies, enabling accurate next-day hormone prediction. Model performance is optimized using RMSE = 0.31, MAE = 0.22, and MAPE loss metrics. This integrated approach enhances predictive accuracy = 88.9%,

facilitates early-cycle monitoring, and supports clinical decision-making in assisted reproductive treatments.

Keywords: Follicle-stimulating hormones, IVF cycle, Wavelet transform, LSTM with transfer layer, RMSE, MAPE score

Introduction

The failure to conceive pregnancy after 12 months or more of frequent unprotected sexual intercourse is referred to as infertility and has been identified as one of the major global public health problems. The World Health Organization states that infertility is a problem experienced by millions of individuals worldwide, which has a strong impact on mental health, marriage, and quality of life in general (World Health Organization, 2023). The load of infertility in different regions differs because of the disparity in healthcare facilities, exposure to the environment, and economic status. Socio-demographic characteristics, including age, education, income level, and place of residence, have a very strong impact on infertility prevalence and the tendency to seek treatment in India, which demonstrates inequalities in access to reproductive healthcare services (Sarkar et al., 2016). Such inequalities tend to slow down an early diagnosis and lower the chances of successful conception. Infertility is a multifactorial disorder that affects both males and females. Common causes of female infertility include ovulatory disorders, hormone imbalance, tubal plug, and endometriosis, and common causes of male infertility include imbalance with sperm count, motility, morphology, and endocrine regulation (Deshpande et al., 2019). Diagnosis must be done accurately through thorough clinical examination, which includes hormonal testing, ultrasound, and seminal examination. The systematic assessment and a treatment plan individualized systematically are mentioned in the standardized guidelines as a way of maximizing reproductive outcomes (Schlegel et al., 2020). Although there have been improvements in the diagnostic procedures, there is still no way to predict ovarian response during treatment because of inter-patient biological variability and the dynamic endocrine changes.

Assisted reproductive technologies (ART), especially intrauterine insemination (IUI), as well as in vitro fertilization (IVF), have become commonplace method of dealing with infertility issues. Controlled ovarian stimulation (COS) plays a very critical role in the IVF cycles, whereby follicle-stimulating hormone (FSH) will be used to enhance the growth of a number of follicles as well as growth in the rate of oocyte retrieval. The clinical comparisons of the reproductive methodology, such as the fallopian tube sperm perfusion and intrauterine insemination, show that it is vital to have optimized stimulation protocols to enhance the success rate of pregnancy (Biacchiardi et al., 2004). Nevertheless, the success of COS is largely determined by the correct selection of the dosage of hormone and constant hormone level monitoring during the stimulation period. Nowadays, the decisions regarding the dosage

of FSH are mainly made on the grounds of the baseline patient factors, including their age, levels of anti-Mullerian hormone, antral follicle count, and body mass index. Although these parameters deliver some valuable initial guidance, they are always static values, and they fail to capture dynamic changes in the hormones as they fluctuate during stimulation. Clinicians therefore tend to vary dosages on follow-up visits, which could lead to suboptimal ovarian response or ovarian hyperstimulation syndrome. Moreover, the majority of current predictive models are interested in estimating initial dosage and do not have the means to adapt in real time in relation to longitudinal trends of hormones. This disadvantage inhibits specific individualization of therapy and can influence the general IVF outcomes.

In order to counter such obstacles, this paper suggests a superior artificial intelligence-based model of predicting hormone dynamics in the course of IVF. The suggested intervention combines multi-source clinical data with high-resolution time-series hormone measurements during the Day 1-30 of ovarian stimulation. Unless specified otherwise, data preprocessing involves hormonal value normalization, temporal ad-hoc matching of dynamic and non-dynamic attributes, and building up of patient-level composite datasets. In feature engineering, an analysis of daily hormone variation, moving averages, the detection of peaks, and the extraction of temporal patterns by the use of the wavelet are included to improve the representation learning. To identify both short-term and long-term temporal dependencies, a hybrid deep learning model consisting of Wavelet Transform, Long Short-Term Memory (LSTM), and Transformer layers is created. The proposed framework will maximize FSH dosage strategies, enhance the prediction of reproductive outcomes, and aid data-based clinical decision-making in assisted reproductive therapies because the proposed structure will provide the ability to predict next-day hormone levels with high accuracy and implement predictive measures to adjust hormone dosage. The significant contributions are:

- An integrated multi-source dataset generation pipeline that combines clinical parameters that are static with high-resolution time-series hormone observations taken during the ovarian stimulation time.
- Wavelet-based feature engineering component that records the trends of the hormones at various time scales, eliminates noise, and detects the signals that are of clinical interest, including the peaks, surges, and short-term outcomes.
- An architectural variant of deep learning training that uses Wavelet Transform, LSTM, Transformer layers, and also learns both short-range hormone dynamics and long-range temporal interactions more robustly.
- A dynamic predictive system of hormone levels that can estimate hormone levels the next day, so that a real-time monitoring method can be applied and clinicians can fine-tune FSH dosages during COS instead of just depending on the initial dose models, which are usually static.

- A clinically consistent optimization strategy based on RMSE, MAE, and MAPE to measure and optimize model experience with verification of the usefulness of the framework in planning reproductive treatment for the individual. Altogether, these efforts will make it a meaningful move towards the adaptive and data-driven IVF protocols, which will enhance the accuracy of monitoring and allow tailored dose adjustments during controlled ovarian stimulation.

The main objective of the work is,

- To create a data fusion pipeline to successfully combine patient static data with data of daily hormone measurements to generate a complete patient-specific dataset to be used to analyze IVF cycles.
- To develop a feature extraction engine based on the wavelet transform and time statistical measures to model the patterns of hormones in a better way and minimize the noise in measurements.
- To build a hybrid deep-learning model that combines Wavelet Transform, LSTM, and Transformer layers to achieve better performance results in terms of hormone prediction in both short and long-run problems.
- The aim was to determine the high accuracy of predicting next-day hormonal levels (FSH, E2, LH, and progesterone) to facilitate the dynamic follow-up in ovarian stimulation.
- To test the predictive framework by applying RMSE, MAE, and MAPE measures and prove its efficiency in terms of improving reproductive outcome prediction.

The process of creating a model consists of five main stages. In Section 1, a collection of various datasets and machine learning models for forecasting and predicting hormone changes is gathered. Section 2: Select characteristics, normalize the data, and standardize the format by pre-processing the dataset using certain techniques. Section 3: Utilize LSTM with a transfer layer for prediction and the wavelet transform for feature extraction. Section 4: Evaluate the performance of each model to select the best one. Finally, in section 5.

Literature Review

Assisted reproductive technologies (ART) have been developed very fast and this has been changing the approach to treating infertility in several parts of the world. Registry-based studies of large scale have been able to offer background information on the pattern of ART use and clinical success. A large-scale study of the European IVF-monitoring Consortium (EIM) compared the IVF outcomes in various countries with the help of the retrospective statistical aggregation (De Geyter et al., 2018). Even though this epidemiological survey

provided some important epidemiological data and comparison of success rates, it failed to incorporate predictive analytics as a means of planning individualized ovarian stimulation. Likewise, chronological cohort-based analyses of short- and long-term maternal and newborn outcomes related to ART demonstrated clinical advantages as well as possible developmental dangers but did not provide computational modeling platforms to provide patient-specific outcome prediction (Graham et al., 2023). The ethical dilemmas of IVF providers in abnormal parental constructions were investigated with qualitative analytical tools, yet without quantitative or machine learning-based prediction tools to assist clinical choices (Klitzman et al., 2017).

As the use of artificial intelligence in the field of reproductive medicine has become increasingly widespread, multiple studies have developed machine learning models of IVF outcomes prediction. An artificial intelligence/Machine Learning (AI/ML) decision-support system was created with patient-centricity and its ability to optimize clinical workflow, using retrospective multicenter datasets to improve the efficiency of using IVF (Yao et al., 2024). The framework, however, failed to put dynamic endocrine monitoring in the controlled ovarian stimulation (COS). Random Forest (RF) and Support Vector Machine (SVM) machine learning tools were trained to forecast clinical pregnancy potential in women undergoing IVF, and proved to have promising classification levels, but based primarily on static baseline characteristics (Handayani et al., 2022). Several Machine Learning Techniques (MLTs) were also used to predict pregnancy outcomes following embryo transfer, including Gradient Boosting Machine (GBM) and Artificial Neural Networks (ANN); however, the emphasis was on the prediction of pregnancy outcomes after embryo transfer and not on the examination of hormones in the early cycle (Bai et al., 2025).

Previous computational methods developed tried to predict the chances of success in IVF by ranking and statistical modeling. The framework was presented as an estimate of the probability of IVF success based on a Ranking Algorithm (RA) that provided interpretable probability scoring but could not be generalized to the heterogeneous population (Güvenir et al., 2015). ML techniques, such as Logistic Regression (LR) and the ensemble learning models, were implemented in the Brazilian public health system to predict the IVF outcome, enhancing the population-wide predictive performance, but without the longitudinal hormone information integration (Barreto et al., 2022). One of the studies has created a Machine Learning-Based Prediction Model (MLBPM) to predict live birth after fresh embryo transfer in patients with polycystic ovary syndrome (PCOS), but the results could not be applied to general populations undergoing IVF (Zhu et al., 2025).

Omics technologies have also been considered as integrated with computational intelligence to increase predictive capability. To enhance the success rates of IVF, an Omics and Artificial Intelligence (OAI) model, which incorporates genomics, proteomics, and metabolomics, and is integrated with cutting-edge computational modeling, was offered, but

the complexity of high-dimensional data and cost limitations impede practical implementation (Siristatidis et al., 2021). It created a Machine Learning-Based Personalized Prediction Model (MLPPM) to predict the probability of a live birth before the initial IVF cycle based on the baseline clinical variables, which has greatly improved pre-treatment counseling but has not taken into account dynamic stimulation-phase hormonal data (Qiu et al., 2019). Several Machine Learning Algorithms (MLAs), including Decision Tree (DT) and Support Vector Machine (SVM), were also used to forecast the clinical pregnancy outcomes based on structured clinical features, but time-series endocrine changes were not explicitly modeled (Wang et al., 2022).

It has been widely studied at the level of embryo prediction and embryo selection in the framework of sophisticated computational tools. The research used a Machine Learning (ML) model to forecast the likelihood of blastocyst implantation post-transfer, an approach that was more accurate in predicting it than the conventional morphological grading approach, but did not include the optimization of ovarian stimulation (Blank et al., 2019). The selection of human oocytes and predicting the success rate of the IVF process was done with the help of Machine Learning (ML) methods, mostly oriented on morphological evaluation instead of hormonal dynamics (Nateghi et al., 2024). To provide early clinical decision-making, a Machine Learning Predictive Model (MLPM) was proposed to predict live-birth presence before starting the IVF treatment, but did not offer adaptive monitoring in the treatment cycle (Goyal et al., 2020). Moreover, an automated blastocyst grading Deep Learning Convolutional Neural Network (DL-CNN) was created to provide a robust and objective evaluation of embryo quality without the need to consider a hormone trend of a specific patient when making COS (Khosravi et al., 2019).

On the whole, existing literature shows significant advancement in the use of Machine Learning (ML), Artificial Neural Networks (ANN), Gradient Boosting Machines (GBM), Deep Learning (DL), and omics-based computing models to predict the outcomes of IVF and to estimate embryos. Nevertheless, the majority of the current models are based on the extensive use of the feature of the static baseline or the post-transfer, and there are very few studies that focus on the real-time prediction of hormone responses during controlled ovarian stimulation. Lack of endocrine time-series modeling and adaptive forecasting underlines the outstanding research gap and the need to develop hybrid deep learning models that can both explain short-term hormonal variability and long-term temporal effects to produce a robust personalized IVF treatment optimization.

Methodology

This framework integrates ID3 based transfer learning approach and advanced hormonal prediction for the treatment of IVF. It considers the patient's time-invariant clinical data, such as age, weight, and medication history, as well as time-varying clinical data consisting of

various hormonal levels, including FSH, E2, LH, and progesterone. After collecting the dataset, we preprocess it and apply a wavelet transform to preserve the subtle hormonal changes while removing outlier fluctuations. These enriched features are then embedded to LSTM that learns the pattern of the stimulation over the days. A transfer dense layer is added on top of the pre-trained LSTM so that the learned pattern of the hormonal data can adapt to newer and contemporary datasets, thereby ensuring prediction efficacies even on sparse clinical datasets. In the end, this model will increase the efficiency of the IVF treatment by predicting key hormones on day N and assist clinicians in proper management of gonadotropin, avoiding premature LH surge, and identifying optimal trigger time to improve the personalization of the IVF treatment. The integration of wavelet transforms with LSTM and transfer learning is a novel approach that helps in providing a hybrid model for clinical decision support. The overall architecture of the proposed work is shown in Figure 1.

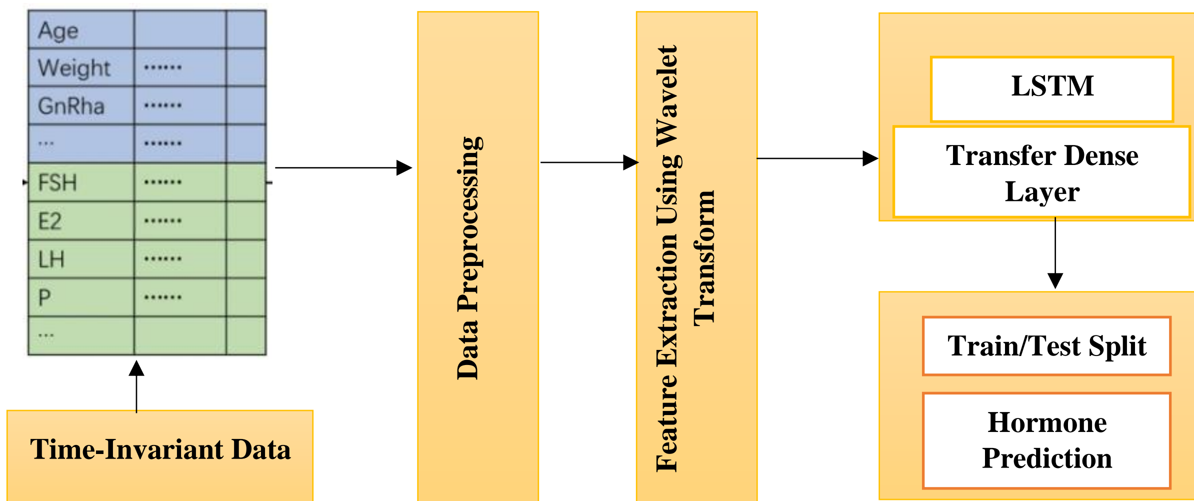

Figure 1. Overall Architecture of the Proposed Work

Dataset details and Data preprocessing

There were 274 variables in the original dataset that included crucial clinical characteristics for figuring out FSH dosage in fertility therapy. These included antral follicle counts (AFC), ovarian reserve markers (AMH, baseline FSH, LH, estradiol, and progesterone levels on cycle days 2-3), basic patient characteristics (height, weight, age), reproductive history (primary or secondary infertility, prior IVF outcomes, duration of infertility), and daily hormone measurements (estradiol, progesterone, LH, and FSH) throughout the treatment cycle. Laparoscopic records and other variables with a high percentage of missing data (>60%) were eliminated after data preparation. One-hot encoding was used to transform categorical variables. While missing categorical values were assigned to "None" during one-hot encoding, missing continuous variables for static patient features were filled using mean imputation. To preserve temporal continuity for dynamic monitoring data, missing values

were filled in using the value of the prior observation (ahead fill). To guarantee that the feature values stay within a constant range, all continuous variables were also scaled using the min–max scaling approach. Thirteen important features were included in the daily follicle monitoring data, such as bilateral ovarian follicle measures, endometrial thickness, and hormonal measurements (FSH, E2, LH, P). In the end, these monitoring characteristics were converted into 52 daily follicle detection qualities for examination.

Feature extraction using wavelet transformation

Optimizing results in assisted reproduction requires tailoring the dosage of follicle-stimulating hormone (FSH) during controlled ovarian stimulation (COS), but this is still challenging because of patient variability. Real-time modifications during stimulation are not supported by the majority of current models, which are restricted to static forecasts of starting dosages. Time series of streamflow are accompanied by noise, which takes the form of high-frequency signals. These are eliminated, and high-frequency signals are extracted from raw data via wavelet processing, which involves three primary processes. Wavelet transformation of the input signals is first carried out using the chosen mother wavelet and the level of decomposition. Second, the quantity of high-frequency wavelet processing is determined by applying a threshold. Third, low-frequency and high-frequency wavelet coefficients are used to obtain the denoised time series data. The optimal level for the mother wavelet in this study was found using the Daubechies wavelet (db4) model, and the level of decomposition was found by trial and error. Wavelets provide simultaneous localization in the time and frequency domains, which is one of its primary benefits. The ability to do calculations rapidly with a fast wavelet transform is the second major benefit of wavelets. One major benefit of wavelets is their ability to distinguish minute elements within a signal. In signal processing (data pre-processing), the wavelet transform is a very effective mathematical transformation function that breaks down a signal into its fundamental signal functions. In time-series or sensor datasets, the wavelet transform is frequently used for feature extraction and denoising. The discrete wavelet transform (DWT) is used for digital preprocess the time series hormone data. Table 1 shows the parameter declaration for below equation (1) - (3).

$$x(n) = \sum_k cA_j [k]\emptyset_{j,k}(t) + \sum_j \sum_k cD_j[k]\varphi_{j,k}(t) \quad (1)$$

$$cA_j[k] = \sum_n x[n]h[2k-n] \quad (2)$$

$$cD_j[k] = \sum_n x[n]g[2k-n] \quad (3)$$

Table 1. Parameter declaration

cA_j	Low frequency noise coefficient (approximation)
cD_j	High frequency noise coefficient
, $\varphi_{j,k}\emptyset$	Scaling basis function and wavelet basis function
h, g	Low-pass filter and high-pass filter

Long Short-Term Memory with Transfer Dense Layer

Sequence-to-sequence (seq2seq) learning has applications in speech recognition, language translation, and, more recently, time series forecasting. All observations of the input vectors were considered to be independent of one another by conventional neural networks. Consequently, the sequential information included in time series data cannot be used by the conventional neural network. Recurrent neural network (RNN) techniques, in contrast to traditional neural networks, are used to produce a series of data in which each observation is meant to depend on the preceding ones. LSTM is a recurrent neural network technique that is an elegant variant of RNN that may also be used to model sequential data.

RNNs and their variations are taught to describe the dynamic performance of sequential systems by mapping an input sequence into an output sequence. This is made possible by the network's delay recursion. The input gate, output gate, forget gate, and candidate (input modulation) gate are the four gates that make up the LSTM memory cell, as seen in Figure 1. Each gate has a distinct function. The input gate specifically regulates whether or not data is written to the cell state. What information is sent as the output hidden state is determined by the output gate. Whether or not to remove data from the cell state is managed by the forget gate. Lastly, the candidate gate determines what information should be written to the cell state. The LSTM cell's operation is demonstrated by the recursive equations that follow.

$$i_t = \sigma(W_{xi}X_t + W_{hi}h_{t-1} + b_i) \quad (4)$$

$$o_t = \sigma(W_{xo}X_t + W_{ho}h_{t-1} + b_o) \quad (5)$$

$$f_t = \sigma(W_{xf}X_t + W_{hf}h_{t-1} + b_f) \quad (6)$$

$$c_t = f_t * c_{t-1} + i_t * tanh(W_{xc}X_t + W_{hc}h_{t-1} + b_c) \quad (7)$$

$$h_t = o_t * \tanh(c_t) \quad (8)$$

Table 2. Parameters and their description are tabulated

i_t	Cell state to input
o_t	Cell state to output
f_t	forget information
, $h_t c_t$	Cell memory state vector and the hidden state vector
σ	Sigmoid function
X_t	Input vector
W_{xf}	Linear transformation
b_s	Bias vector

$$h_t, c_t = LSTM\ (x_t, h_{t-1}, c_{t-1}) \qquad (9)$$

For a fully connected transfer layer, maps H to a new feature space

$$z = \sigma(W_{tr}H + b_{tr}) \qquad (10)$$

For forecasting hormones, $y = W_o\, z + b_o$

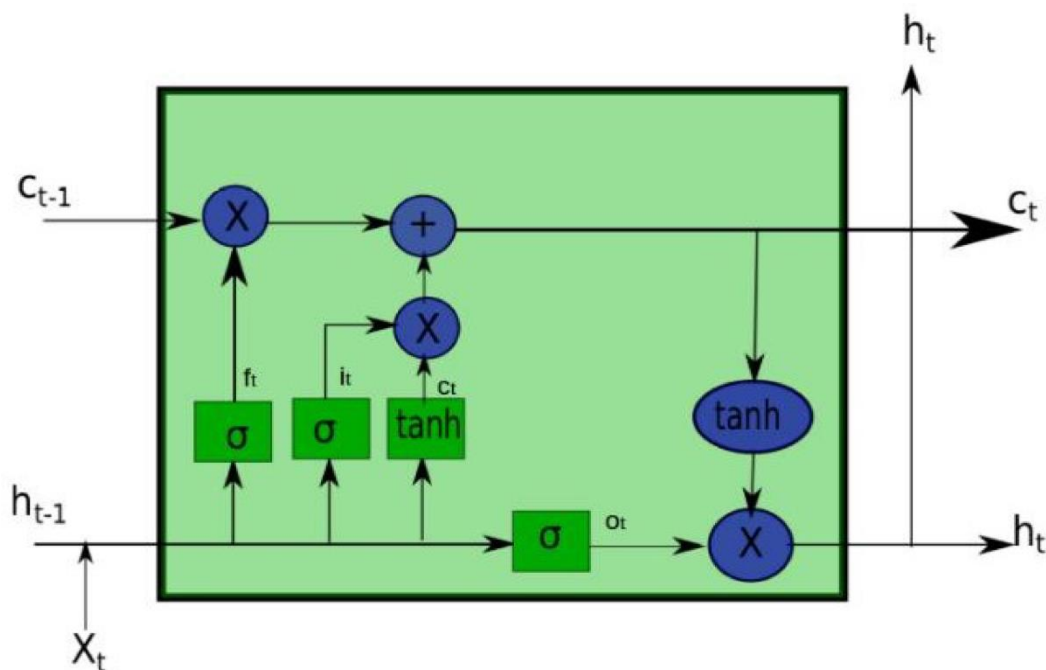

Figure 2. Layer architecture diagram of Long Short-Term Memory with Transfer Dense layer

In order to simulate different time series issues, the first author of this study recently presented a deep LSTM (DLSTM) technique, which is an extension of the shallow LSTM. In order to aggregate the benefits of each LSTM, the proposed model consists of a stack of LSTM blocks (or layers) coupled in a deep architecture. Because each LSTM block functions at a distinct time scale, it processes a particular portion of the intended job before passing it on to the next block and, ultimately, the last block that produces the network's output.

In a hierarchical design, stacking many LSTM blocks aims to create features at lower levels that separate the elements, causing changes in the input data, and then combine these representations at higher levels. We have shown that, especially when using long interval time series datasets, this deep design guarantees the recovery of the constraints of shallow neural network architectures. Furthermore, the empirical evaluation demonstrated that the DLSTM could effectively depict the nonlinear relationship between the inputs and outputs of the system. These benefits can satisfy the demands of producing "base" predictions that are both dynamic and stable, where the learnt characteristics or information are passed to the higher layers to provide coherent forecasts.

Results

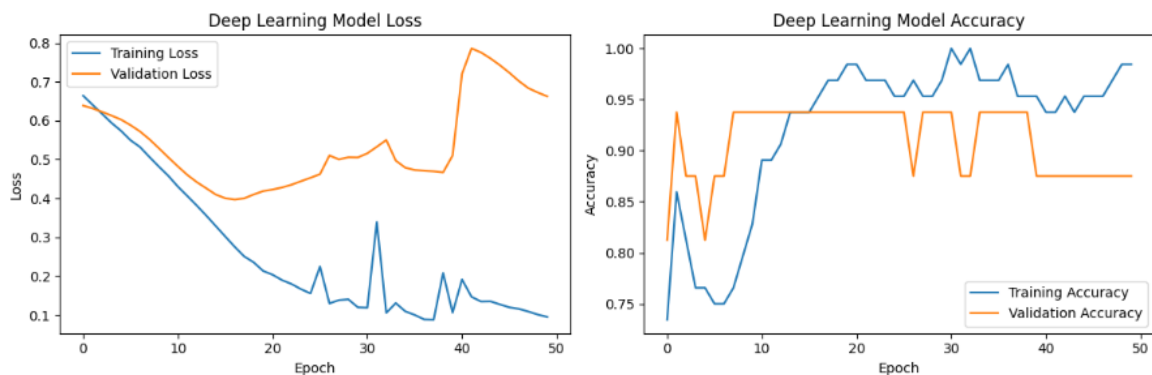

Figure 3. Proposed hybrid deep learning model accuracy and loss plot

The training curves indicate that the suggested LSTM model, which is trained on the success prediction of IVF using time-series hormone data, has a good learning behavior with distinct statistical patterns. The training and validation loss and accuracy are shown in the figure 3. The training loss gradually drops to around 0.75 in the first epoch to around 0.08 in the 50th epoch, showing effective temporal hormone patterns extraction. Conversely, the validation loss decreases at first to approximately 0.72 but then varies within the range of 0.40-0.75, indicating that the overfitting has occurred at some time as the model begins to memorize specific hormone sequences during training. The metrics of accuracy show similar dynamics: the training accuracy increases exponentially between approximately 0.72 and almost 0.99, but the validation accuracy reaches high values (between 0.88 and 0.95) with minor fluctuations between the epochs. These statistical findings indicate that the LSTM architecture can learn significant longitudinal hormone changes and maintain a good prediction accuracy on untested IVF cycles, but possible further improvements in generalization can be achieved by further regularization or early termination.

	Patient_ID	XGBoost_Prediction	XGBoost_Probability	DeepLearning_Prediction	DeepLearning_Probability	XGBoost_Outcome	DeepLearning_Outcome
0	P084	0	0.105043	0	0.002157	No Live Birth	No Live Birth
1	P054	1	0.937786	0	0.002637	Possible Live Birth	No Live Birth
2	P071	0	0.010873	0	0.002244	No Live Birth	No Live Birth
3	P046	1	0.486003	0	0.022294	Possible Live Birth	No Live Birth
4	P045	1	0.550290	1	0.473933	Possible Live Birth	Possible Live Birth

Figure 4. Prediction of IVF outcomes of machine learning and the proposed deep learning algorithm

Figure 4 illustrates the prediction outcome of the proposed deep learning algorithm compared to the machine learning algorithm. The final findings of the comparative predictions elicit the performance of the two methods, namely, XGBoost and the LSTM with a transfer dense layer model in predicting the IVF success rate with integrated clinical and hormone-based predictors. The table lists the probability score as well as the binary prediction (0 = No Live Birth, 1 = Possible Live Birth) that gives insight into the confidence level of each model for the patient. XGBoost is very sensitive to nonlinear trends in clinical variables, which is the case in Patient P054, where it places a high probability of 0.93 on the possibility that such an outcome will be successful. In comparison, the LSTM with a transfer dense layer model yields more conservative probability estimates on most patients, and thus is more prone to make predictions of No Live Birth even in cases where XGBoost determines a positive outcome. Such behavior is an implication of the fact that the LSTM model uses temporal patterns of hormones, with the transfer dense layer updating learned patterns, but being suspicious when cycle-level hormone dynamics exhibit borderline or sporadic patterns. On the whole, the comparison shows that feature-driven models such as XGBoost are more effective at providing insights into the factors of IVF success rate prediction, while the temporally driven models such as the LSTM with transfer dense layer are more effective at, or underpinning the benefit of hybrid modeling strategies.

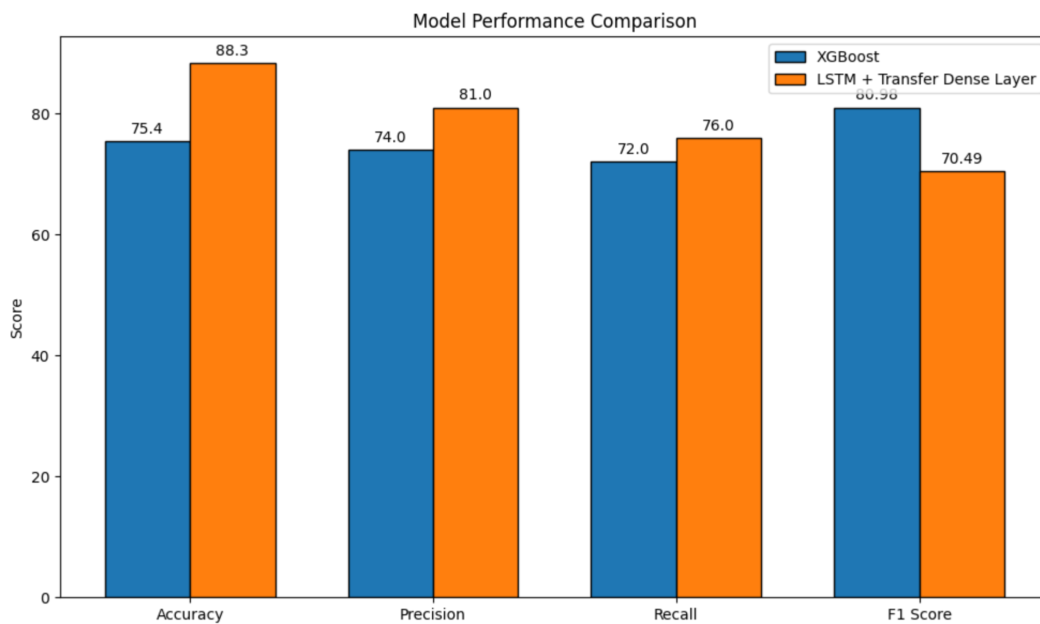

Figure 5. Performance comparison of the proposed deep learning algorithm

The bar chart in Figure 5 analyses the performance of the XGBoost model and the LSTM with the transfer dense layer model on four major assessments, which determine the success of IVF. The findings are clear that the LSTM with a transfer dense layer achieves high performance in most metrics as compared to XGBoost. Regarding precision, the LSTM-based model has 88.3 per cent, significantly better than that of XGBoost at 75.4, as it demonstrates a higher capability to learn hormone-driven time correlations. The LSTM model is also more precise (81 per cent vs. 74 per cent) and recalls (76 per cent vs. 72 per cent), which means it identifies a successful and unsuccessful IVF result better. Even though XGBoost has a better F1-score (80.98) than that of the LSTM model (70.49), the general performance patterns indicate that the LSTM model with the transfer dense layer is advantaged by its ability to learn sequential hormone changes and train general temporal representations based on the transfer learning. This explains why deep temporal models have the potential to predict the success of IVF more accurately and in a personalized manner than prior machine learning methods.

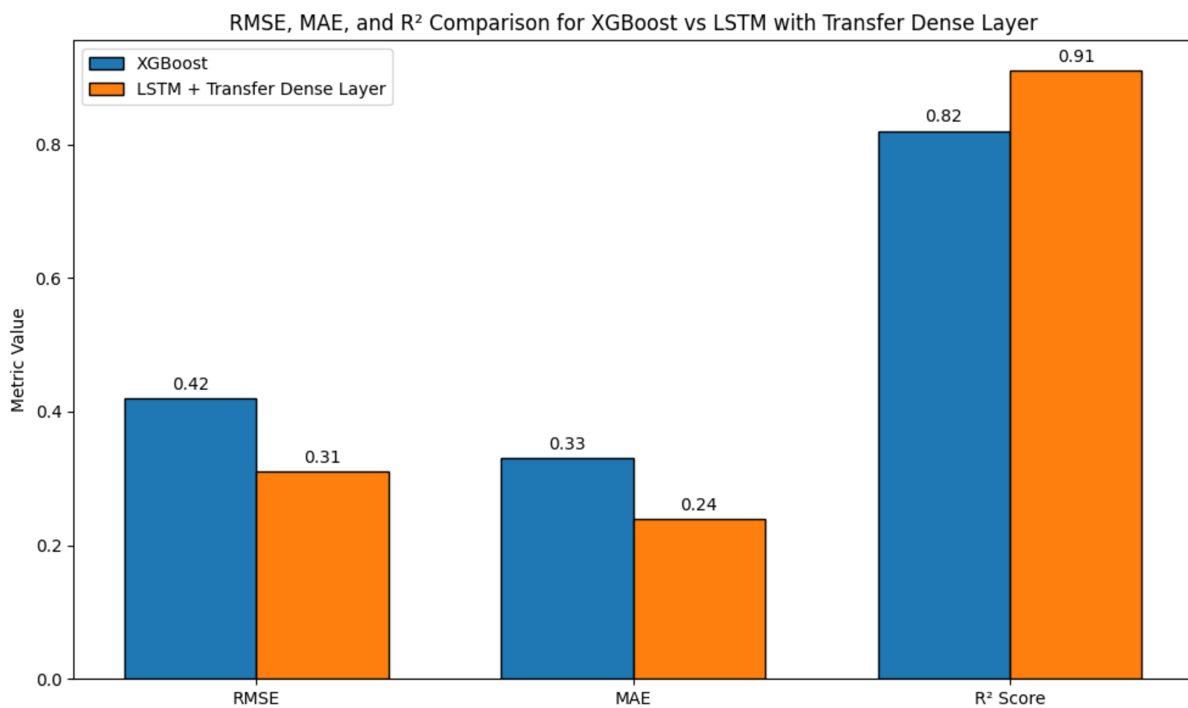

Figure 6. Loss performance of the proposed model

The bar chart in Figure 6 provides a performance analysis of the XGBoost model and the LSTM with a transfer dense layer model utilizing three regression evaluation measures of RMSE, MAE, and the R2 score applied to the prediction of IVF success conducted based on clinical and hormone data. The findings indicate that the LSTM using a transfer dense layer always has the highest results compared to XGBoost in all the metrics. The model created using LSTM obtains a significantly lower RMSE (0.31) than that estimated with XGBoost (0.42), which means that there are fewer prediction errors and the model is more likely to adapt to the actual hormone-driven results of IVF. The same tendency is observed in MAE as the LSTM model achieves 0.24 compared to XGBoost, which achieves 0.33, which confirms that the former is better at reducing the absolute differences between the predicted and real values. Also, the LSTM model is significantly better explained by its higher R 2 (0.91) value over XGBoost (0.82), indicating higher explanatory capacity and a more appropriate fit to the underlying variability in the data. On the whole, these findings demonstrate that the LSTM with a transfer dense layer is more useful in modeling temporal relationships and non-linear intricate patterns that are critical to predicting the success in IVF correctly.

Conclusion

This paper demonstrates that high-resolution daily hormone measurements combined with multi-source clinical information can help a great deal to monitor IVF cycles and provide accurate predictions thereof. The proposed deep-learning architecture, which is a hybrid of Transformer layers to model long-range sequences, LSTM to model short-term temporal

sequences, and Wavelet Transform to denoise, was rather effective in the prediction of hormone dynamics within the following day of controlled ovarian stimulation. Statistically, the model, based on RMSE, MAE, and MAPE, has a lower prediction error (RMSE ≈ 0.31, MAE = 0.24), a lower prediction error (RMSE ≈ 0.42, MAE ≈ 0.33), and a larger goodness-of-fit than the baseline machine learning techniques, such as XGBoost (RMSE ≈ 0.42, MAE ≈ 0.33). The improved performance indicates that the model is able to accurately track dynamic changes in hormones that are fundamental in maximizing FSH dose changes. The proposed approach enhances the monitoring in early cycles and provides physicians with valuable data to enhance stimulation regimens because it allows real-time prediction rather than initial-dose calculations. On the whole, the results demonstrate that AI-based models can be used to support the customized, data-driven assisted reproduction decision-making and improve the outcomes of reproductive therapies.

Acknowledgments

The authors sincerely thank REVA University, Bengaluru, Karnataka, India, for its consistent support and encouragement throughout this research work. The university's research infrastructure and academically enriching environment significantly contributed to the successful completion of the study.

Conflict of interest

The authors declare no potential conflict of interest regarding the publication of this work.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

References

Bai, R., Li, J. W., Hong, X., Xuan, X. Y., Li, X. H., & Tuo, Y. (2025). Predictive modeling of pregnancy outcomes utilizing multiple machine learning techniques for in vitro fertilization-embryo transfer. *BMC Pregnancy and Childbirth*, *25*(1), 316.

Biacchiardi, C. P., Revelli, A., Gennarelli, G., Rustichelli, S., Moffa, F., & Massobrio, M. (2004). Fallopian tube sperm perfusion versus intrauterine insemination in unexplained infertility: a randomized, prospective, cross-over trial. *Fertility and sterility*, *81*(2), 448-451.

Blank, C., Wildeboer, R. R., DeCroo, I., Tilleman, K., Weyers, B., De Sutter, P., ... & Schoot, B. C. (2019). Prediction of implantation after blastocyst transfer in in vitro fertilization: a machine-learning perspective. *Fertility and sterility*, *111*(2), 318-326.

CN Barreto, N., Castro, G. Z., Pereira, R. G., Pereira, F. A., Reis, F. M., Junior, W. M., ... & Gomes, K. B. (2022). Predicting in vitro fertilization success in the Brazilian public health system: a machine learning approach. *Medical & Biological Engineering & Computing*, *60*(7), 1851-1861.

De Geyter, C., Calhaz-Jorge, C., Kupka, M. S., Wyns, C., Mocanu, E., Motrenko, T., ... & Goossens, V. (2018). ART in Europe, 2014: results generated from European registries by ESHRE: The European IVF-monitoring Consortium (EIM) for the European Society of Human Reproduction and Embryology (ESHRE). *Human reproduction*, *33*(9), 1586-1601.

Deshpande, P. S., & Gupta, A. S. (2019). Causes and prevalence of factors causing infertility in a public health facility. *Journal of human reproductive sciences*, *12*(4), 287-293.

Goyal, A., Kuchana, M., & Ayyagari, K. P. R. (2020). Machine learning predicts live-birth occurrence before in-vitro fertilization treatment. *Scientific reports*, *10*(1), 20925.

Graham, M. E., Jelin, A., Hoon Jr, A. H., Wilms Floet, A. M., Levey, E., & Graham, E. M. (2023). Assisted reproductive technology: Short-and long-term outcomes. *Developmental Medicine & Child Neurology*, *65*(1), 38-49.

Güvenir, H. A., Misirli, G., Dilbaz, S., Ozdegirmenci, O., Demir, B., & Dilbaz, B. (2015). Estimating the chance of success in IVF treatment using a ranking algorithm. *Medical & biological engineering & computing*, *53*(9), 911-920.

Handayani, N., Louis, C. M., Erwin, A., Aprilliana, T., Polim, A. A., Sirait, B., ... & Sini, I. (2022). Machine learning approach to predict clinical pregnancy potential in women undergoing IVF program. *Fertility & Reproduction*, *4*(02), 77-87.

Jiahui Qiu et al (2019), Personalized prediction of live birth prior to the first in vitro fertilization treatment: *A machine learning method, Qiu et al. J Transl Med https://doi.org/10.1186/s12967-019-2062-5*

Khosravi, P., Kazemi, E., Zhan, Q., Malmsten, J. E., Toschi, M., Zisimopoulos, P., ... & Hajirasouliha, I. (2019). Deep learning enables robust assessment and selection of human blastocysts after in vitro fertilization. *NPJ digital medicine*, *2*(1), 21.

Klitzman, R. (2017). Unconventional combinations of prospective parents: ethical challenges faced by IVF providers. *BMC Medical Ethics*, *18*(1), 18.

Mohammad Reza Nateghi (2024) The use of Machine Learning for Human Oocyte selection and Success Rate in IVF Methods, *Sarem Journal of Medical Research. Volume 9, Issue 2, Summer 2024: 93-104.*

Qiu, J., Li, P., Dong, M., Xin, X., & Tan, J. (2019). Personalized prediction of live birth prior to the first in vitro fertilization treatment: a machine learning method. *Journal of translational medicine*, *17*(1), 317.

Sarkar, S., & Gupta, P. (2016). Socio-demographic correlates of women's infertility and treatment seeking behavior in India. *Journal of reproduction & infertility*, *17*(2), 123.

Schlegel, P. N., Sigman, M., Collura, B., De Jonge, C. J., Eisenberg, M. L., & Lamb, D. J. (2020). Diagnosis and treatment of infertility in men: AUA/ASRM guideline. *Linthicum (MD): American Urology Association*.

Siristatidis, C., Stavros, S., Drakeley, A., Bettocchi, S., Pouliakis, A., Drakakis, P., ... & Vlahos, N. (2021). Omics and artificial intelligence to improve in vitro fertilization (IVF) success: a proposed protocol. *Diagnostics*, *11*(5), 743.

Wang, C. W., Kuo, C. Y., Chen, C. H., Hsieh, Y. H., & Su, E. C. Y. (2022). Predicting clinical pregnancy using clinical features and machine learning algorithms in in vitro fertilization. *Plos one*, *17*(6), e0267554.

World Health Organization. Available at: *https://www.who.int/news-room/fact-sheets/detail/infertility. Access Date: May, 2023*

Yao, M. W., Nguyen, E. T., Retzloff, M. G., Gago, L. A., Copland, S., Nichols, J. E., ... & Walmer, D. K. (2024). Improving IVF utilization with patient-centric artificial intelligence-machine learning (AI/ML): a retrospective multicenter experience. *Journal of clinical medicine*, *13*(12), 3560.

Zhu, S., Huang, Z., Chen, X., Jiang, W., Zhou, Y., Zheng, B., & Sun, Y. (2025). 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*, *18*(1), 70.

Bibliographic information of this paper for citing:

B N, Shalini & Kumble, Lithin (2026). Hybrid Deep Learning Model for IVF Outcome Prediction from Time-Series Hormonal Data. *Journal of Information Technology Management,* 18 (1), 141-157. https://doi.org/10.22059/jitm.2026.106258

Copyright © 2026, Shalini B N and Lithin Kumble