SAKARYA UNIVERSITY JOURNAL OF COMPUTER AND INFORMATION SCIENCES
http://saucis.sakarya.edu.tr/
e-ISSN: 2636-8129
Publisher: Sakarya University
Vol. 8, No. 2, 301-311, 2025
DOI: 10.35377/saucis..1612668

RESEARCH ARTICLE

The Development of Digital Twin Baby Incubators for Fault Detection and Performance Analysis

Hatice Kabaoğlu[1*], Fecir Duran[2], Emine Uçar[3]

[1] Gazi University, Faculty of Technology, Department of Computer Engineering, Ankara, Türkiye, ror.org/054xkpr46

[2] Gazi University, Faculty of Technology, Department of Computer Engineering, Ankara, Türkiye, ror.org/054xkpr46

[3] İzmir Bakırçay University, Faculty of Economics and Administrative Sciences, Department of Management Information Systems, İzmir, Turkey, ror.org/017v96566

Corresponding author:
Hatice Kabaoglu, Gazi University
haticetrasoglu@gmail.com

Article History:
Received: 03.02.2025
Revised: 28.05.2025
Accepted: 29.05.2025
Published Online: 23.06.2025

ABSTRACT
This study focuses on developing a digital twin for baby incubators in neonatal intensive care units to enhance monitoring and care for premature infants. The digital twin employs a hybrid model integrating Long Short-Term Memory (LSTM) and Random Forest (RF) algorithms to predict potential errors and alarms. The LSTM algorithm was trained using sensor data provided by a health technology company to predict future measurements. Subsequently, the RF algorithm classifies these predictions into specific error conditions. The hybrid model demonstrates success with mean squared error and mean absolute error values of 1540533.6 and 160.8 for the LSTM model and an 86.44% accuracy rate for the RF model. The study's key findings emphasize the effectiveness of the hybrid model in predicting future sensor values and classifying errors, representing a significant step towards improving premature baby care. Integrating LSTM and RF algorithms offers an innovative approach to error prediction, minimizing risks and improving premature infant health outcomes. In summary, this study successfully develops a digital twin for baby incubators, offering a promising solution for advancing newborn healthcare services and providing a foundation for future research.

Keywords: Baby Incubator, Digital Twin, Decision Tree, Health Monitoring, Machine Learning Classifier

1. Introduction

The rapid advancement of technology has led to the emergence of innovative solutions in various sectors, including the healthcare industry. Critical areas in healthcare, especially those requiring special attention, such as the care of premature infants, have seen significant developments. Premature birth occurs when delivery takes place before the 37th week of pregnancy. Challenges arising from the weak thermal regulation systems of premature infants highlight the sensitivity of this condition [1]. These infants are susceptible to high heat loss due to their large body surface area relative to their weight and underdeveloped nervous systems [2], [3]. Baby incubators are essential devices that provide crucial support for the survival of prematurely born infants. Their low weight, insufficient organ development, and inability to regulate body temperature can lead to various complications and an increased risk of death.

To support the growth and development of premature infants, they need to be placed in incubators. These incubators maintain the body temperature of infants and protect them from infections and external factors that could negatively impact their condition by providing a regulated environment. Additionally, incubators allow monitoring of vital signs such as heart rate and respiration, enabling timely medical interventions to maximize premature infants' health and survival chances. In summary, baby incubators create a controlled environment, preserving appropriate temperature, humidity, and oxygen levels to support prematurely born infants' survival and growth. However, specific situations can render incubators dangerous. Equipment malfunctions, human errors, or environmental factors can disrupt the controlled environment, leading to fluctuations in temperature, humidity, and oxygen levels, posing risks to the delicate and underdeveloped bodies of premature infants. Such fluctuations can result in complications like hypothermia, infection, and brain damage, posing a vital threat to the infant. Therefore, this study aims to create a digital twin of a baby incubator to prevent potential damage during the observation and care of baby incubators.

The digital twin of baby incubators is an innovative concept that has gained significant importance in the healthcare sector with today's advancing technology. This concept involves transferring physical baby incubators into a fully digital environment, enabling real-time monitoring, analysis, and simulation. The digital twin is an accurate and precise reflection of real baby incubators, allowing the simulation of their behaviors, functions, and responses with complete accuracy.

Cite as: H. Kabaoğlu, F. Duran & E. Uçar. (2025). The development of digital twin baby incubators for fault detection and performance analysis. *Sakarya University Journal of Computer and Information Sciences*. vol. 8, no. 2, pp. 301-311, 2025 doi: 10.35377/saucis...1612668

Moreover, this technology is used to understand better, predict, and optimize the care of premature infants. The environmental conditions within the incubators, sensor data, and parameters like temperature, humidity, and pressure are continuously monitored and analyzed in real time, providing healthcare professionals with the opportunity for real-time tracking, prediction, and intervention. This technology can assist healthcare professionals in developing early warning systems, minimizing risks, and achieving better health outcomes. Furthermore, the digital twin of baby incubators can be utilized for testing innovations, creating educational simulations, and enhancing the effectiveness of baby care. In conclusion, the digital twin of baby incubators exemplifies the integration of technology and healthcare services in the health sector, potentially making premature infant care smarter, safer, and more effective while serving as a foresight for future health applications.

In transforming a device into a digital twin, various functions are required for data and its transformation into information. In this study, the baby incubator was modeled using the LSTM and RF methods. Sensor data collected from the OKUMAN company was adapted for use in the developed digital twin model. The developed model was used to train and create simulations for baby incubators' use, operation, or maintenance. This study aims to understand the causes of problems in baby incubators, monitor real-world situations in real time, and generate solutions.

The main contributions of our developed digital twin to the literature in this area are as follows: Firstly, this technology provides healthcare professionals with the opportunity for more precise and effective interventions by offering real-time monitoring and analysis in the care of premature infants. Secondly, continuously monitoring environmental parameters within incubators contributes to developing early warning systems, thereby assisting in minimizing risks. Thirdly, the digital twin of baby incubators can be used as a platform for creating educational simulations and testing innovations, facilitating continuous improvement in the care of premature infants. These contributions offer significant potential in premature infant health and care and can serve as a guiding force for future health applications.

The remainder of the study is organized as follows: The second section involves a review of relevant literature and summarizing key findings. The third section explains the methods, data collection processes, and analysis techniques. The fourth section covers experimental studies, obtained findings, and a discussion section, while the fifth and final section highlights the significant results of the study.

2. Related Works

The increasing demand for healthcare services, aging populations, the prevalence of chronic diseases, and limited financial resources make adopting digital technologies crucial in the health field. Although digital twins and "hyper-automation" solutions have emerged as significant technological trends in recent years, they have not been fully utilized in the medical field [4].

The use of digital twin technology in areas such as disease prevention, preparedness for medical crises, and patient counseling stands out as a response to the increasing demand for healthcare services and the need to better respond to challenging conditions [5]–[6]. For instance, a study by Haleem and his team emphasized the importance of digital twin technology in providing reliable medical advice based on patient health data, personalizing treatment, and enhancing the efficiency of hospital operations [7]. Similarly, Peshkova and her team explored the potential of digital twins in cancer treatment, examining their ability to predict disease dynamics [8]. Additionally, Han and his team developed a framework to optimize hospital operations using digital twins in the context of smart hospitals, highlighting real-time data analysis as a critical factor in improving hospital operations [9].

Addressing the future of prediction and health management strategies in the manufacturing industry, Toothman and his team developed a digital twin-based framework standardizing health monitoring modeling, emphasizing how this approach could contribute to the health monitoring strategies of industrial equipment [10]. A study on emergency hospital services by Aluvalu and his team made a significant contribution by optimizing the medical history and treatment processes of anonymous patients to ensure rapid and effective treatment [11]. These studies indicate the increasing importance of digital twin technology in healthcare.

Digital twins conduct data collection processes with architectures involving sensor and measurement technology, the Internet of Things (IoT), and Machine Learning (ML) [12]. Machine Learning, an artificial intelligence and computer science subfield, aims to mimic human learning abilities using data and algorithms. Digital twins, especially by employing machine learning models to solve specific tasks, can gain experience. This enables modeling real-world events in a virtual environment and integrating ML methods to predict future situations. Digital twins can make data-driven decisions using the analytical capabilities of ML, leading to more effective strategies [13]–[15]. Projects involving digital twins of human organs, such as the heart, have been initiated by several companies, including Dassault Systems, for use in drug discovery and healthcare [16]. Manocha and his team enriched an intelligent healthcare solutions framework using advanced techniques such as IoT, deep learning, and Blockchain, demonstrating the effectiveness of smart healthcare solutions in improving medical services [17].

Given that the COVID-19 pandemic has accelerated digital transformation in the health sector, the role of technology in healthcare has become even more prominent [18]. César and his team significantly contributed to modeling the COVID-19 pandemic in the health field. Their study demonstrated that epidemiological models could be more effectively and less expensively predicted using advanced machine learning methods such as Long Short-Term Memory (LSTM). This approach highlights how data analytics and machine learning techniques can contribute valuable healthcare contributions [19]. Chen and his colleagues investigated using artificial intelligence algorithms, especially digital twins, to predict and control rapidly spreading situations like COVID-19. They emphasized the usability of digital twin technology for real-time monitoring and prediction of epidemiological prevention and control situations. Accurate prediction of data trends over time is crucial in controlling such situations. Researchers focused on the usability of LSTM, a recurrent neural network capable of effectively modeling long-term dependencies. This study highlights the potential of digital twins and LSTM technologies in smart cities' pandemic prevention and control processes, aiming to enhance information security and improve epidemiological prediction accuracy [20]. Lv and his team developed a digital twin-based human-robot collaboration assembly approach to meet the increased demand for medical equipment production in the post-COVID-19 period. This study offers an important perspective on how digital twins can be used in healthcare [21]. Neog and his team designed a remote health monitoring system using IoT and ML techniques. The study compared sensor data with COVID-19 risk using unsupervised ML algorithms, particularly the LSTM algorithm and Markov Model, to determine COVID-19 risk. The LSTM algorithm provided better results [22].

The health and comfort of a baby are primary concerns in baby incubator environments. In this context, accurately predicting critical data such as temperature and humidity inside the incubator is crucial to improving the baby incubator experience. Traditional predictive methods often lack accuracy for complex time-series data. In contrast, data-driven approaches such as machine learning and deep learning have become increasingly prominent. Notably, the Long Short-Term Memory (LSTM) architecture—a variant of recurrent neural networks—has proven highly effective for sequential prediction tasks. LSTM stands out because it can effectively model long-term time dependencies and resilience against the vanishing gradient problem. The vanishing gradient problem is a challenge encountered during the training process of artificial neural networks. Artificial neural networks are typically trained by propagating gradients backward. As these gradients propagate backward, they can decrease. The vanishing gradient problem weakens the network's learning ability, especially in deep artificial neural networks. In this situation, insufficient updates can be made to the training data in the initial layers, leading to the network not learning as desired. Types of recurrent neural networks, such as LSTM, are designed to overcome this problem [23], [24]. While there seems to be no specific study conducted in the context of baby incubators, the success of studies in other areas, such as COVID-19, demonstrates the potential use of this method in the baby incubator environment [25]–[27]. This study aims to highlight the potential use of digital twins in optimizing baby incubators and improving care processes.

Baby incubators support prematurely born infants' survival and healthy development. Previous research has provided in-depth knowledge about incubators' design, function, and effectiveness. Earlier studies addressed fundamental objectives such as regulating environmental conditions inside incubators, maintaining the body temperature of infants, and reducing infection risks. For example, Yeler and Koseoglu developed a mathematical model to predict the performance of a baby incubator used to care for premature infants [28]. Cuervo and his team designed and tested a low-cost newborn incubator to reduce newborn deaths in developing countries [29]. Kapen and other researchers developed an automatic newborn incubator to improve health [30]. Hannouch and her colleagues analyzed babies' thermal comfort and losses by examining heat and mass transfer in baby incubators [31]. In another study, Puyana-Romero and her team aimed to measure the echo time in incubators by suggesting that high sound levels in newborn incubators were enhanced by incubator materials [32]. Chandrasekaran and his team investigated whether a low-cost cardboard incubator could maintain the thermal regulation of premature babies [33]. In another study, Fraguela and his colleagues mathematically modeled heat exchange and energy balance in a closed incubator to ensure the thermal stability of newborns [34]. However, there is no research on the digital twin of baby incubators in the literature.

This study aims to highlight the potential use of digital twins in optimizing baby incubators and improving care processes. This approach was implemented using the LSTM algorithm, which effectively models complex time-series data for vulnerable patient groups, such as premature infants. Additionally, the Random Forest algorithm—known for highlighting the importance of features in datasets—provided valuable insights to support clinical decision-making.While underscoring the importance of baby incubators and digital twins in the healthcare sector, this study carries the potential to guide future research.

3. Material and Methods

This study aims to develop a digital twin model for monitoring incubators critical for premature infants. In this context, a dataset obtained from OKUMAN Health Company was utilized to predict future sensor values using the LSTM algorithm. Subsequently, the predicted values were classified into specific error scenarios using the RF algorithm. This novel approach aims to anticipate potential errors by predicting future sensor values in the incubator. The overall structure of the proposed model is illustrated in Figure 1.

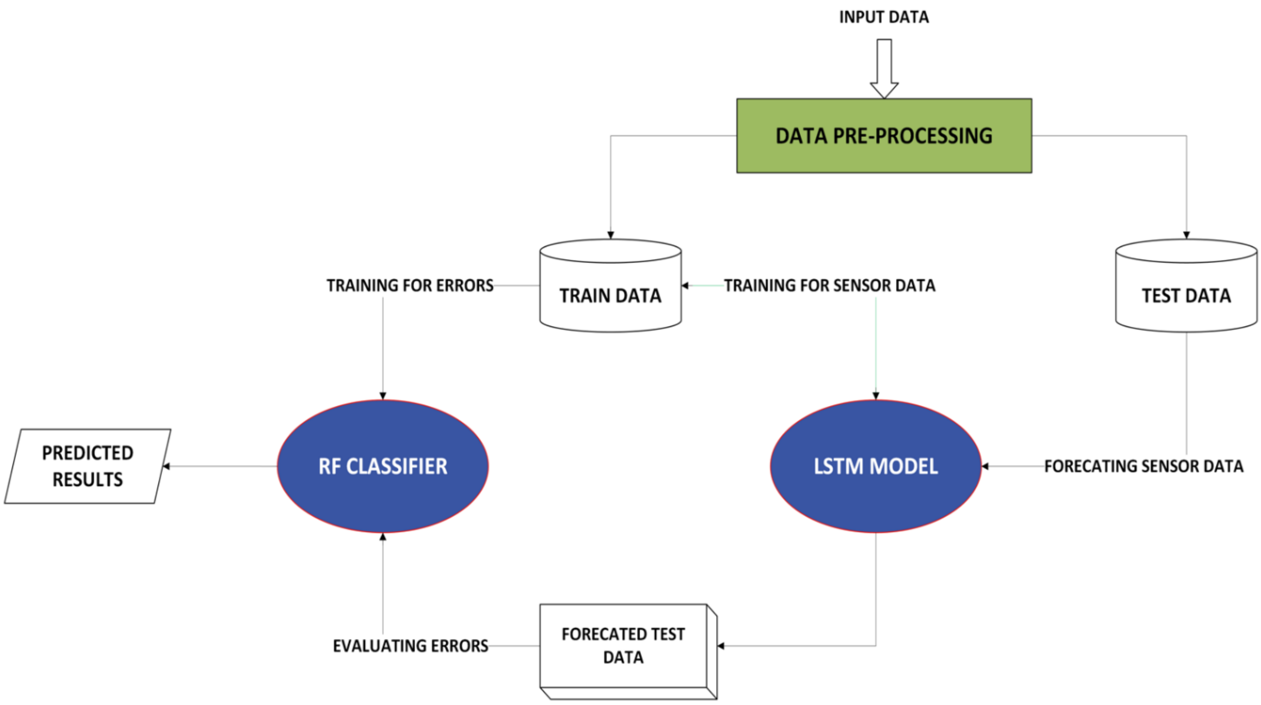

Figure 1. Architecture of the Proposed Forecasted Model

3.1 Datasets

The data used in the study were obtained from the OKUMAN company. The dataset comprises real sensor readings obtained from neonatal incubators developed and maintained by the company. Although the dataset is proprietary, it can be made available for academic use upon reasonable request to the authors or the company. This research, 13 variables were used as input parameters for the prediction algorithm, excluding date and alarm output values. As seen in Table 1, The dataset used contains 279,671 data points. The data were acquired from a series of sensors in the incubator, as well as information related to the system state.

Table 1. Incubator Sensor Data Properties

Serial No	Property Name	Description
1	Set Temperature	Indicates the set temperature of the incubator for the baby. The temperature range is determined to maintain the baby's comfort and health.
2	Air Temperature 1	Data was obtained from a sensor measuring the air temperature inside the incubator.
3	Air Temperature 2	Data from a second sensor measuring the air temperature inside the incubator.
4	Air 1 Temperature Data (Raw)	Raw data from the first air temperature sensor.
5	Air 2 Temperature Data (Raw)	Raw data from the second air temperature sensor.
6	Skin Temperature 1	First, skin temperature is measured from the baby's skin.
7	Skin Temperature 2	Second, skin temperature is measured on the baby's skin.
8	Heater Temperature	Indicates the temperature of the heating element inside the incubator.
9	Heater Power Percentage	Percentage value at which the heater power is set.
10	Heater Current	Electric current passes through the heater to generate heat.
11	Oxygen Percentage	Data was obtained from a sensor measuring the oxygen level inside the incubator.
12	Fan Current	Electric current passes through the fan motor inside the incubator.
13	Fan Speed	Indicates the rotational speed of the fan inside the incubator.
14	Date-Time	Date and time information when the data were recorded.
15	System Alarms	Alarm information indicating possible errors or deviations related to the incubator system.

3.2 Data Preparation

The dataset was prepared sequentially, undergoing data cleaning, splitting, and saving the data to make it suitable for analysis. In the data cleaning stage, unnecessary or irrelevant data that does not contribute to the prediction process was initially removed. The dataset was then divided into input and output data using specific input and output step parameters, ensuring the data was appropriately prepared for the model's training. Additionally, each month's dataset segments were merged to create a single dataset, mitigating the impact of gaps and missing data. The prepared datasets were saved for future use, allowing for the data's reuse and enhancing the results' reproducibility.

To minimize the impact of time and prevent negative effects on predictions, the dataset was arranged chronologically and grouped on a time basis. While 80% of the dataset was used for training, the remaining 20% was reserved for testing.

3.3 Long Short-Term Memory (LSTM) Algorithm

LSTM (Long Short-Term Memory) is one of the fundamental deep learning algorithms used in this study to predict the future health status of premature babies. LSTM is known for its resistance to the vanishing and exploding gradient problems that arise in traditional Recurrent Neural Network (RNN) models. This feature makes it ideal for modeling long-term dependencies, making it suitable for analyzing complex relationships over time, such as time series. LSTM cells comprise four main components: the forget gate, input gate, memory cell, and output gate. These components are essential for the network to forget past information, incorporate new information, and generate output. Below are the key equations describing the behavior of these gates:

Forget Gate (ft):

$$\boldsymbol{ft = \sigma(Wf \cdot [ht - 1, xt] + bf)} \tag{1}$$

Input Gate (it):

$$\boldsymbol{it = \sigma(Wi \cdot [ht - 1, xt] + bi)} \tag{2}$$

Updated Memory (C_t):

$$\boldsymbol{Ct = ft \cdot Ct - 1 + it \cdot \tanh(WC \cdot [ht - 1, xt] + bC)} \tag{3}$$

Output Gate (o_t):

$$\boldsymbol{ot = \sigma(Wo \cdot [ht - 1, xt] + bo)} \tag{4}$$

Cell Output (h_t):

$$\boldsymbol{ht = ot \cdot \tanh(Ct)} \tag{5}$$

Here, x_t represents the input data , h_t is the cell's output, C_t is the memory cell, W_f , W_i , W_C ,W_o and , b_f , b_i , b_C , b_0 represent the learned weights and biases. σ denotes the sigmoid function, and tanh represents the hyperbolic tangent function. As seen in Equation 1, the forget gate determines which information is discarded from the cell state. The input gate's function, as seen in Equation 2, involves determining which values from the input should be updated to the cell state. Equations 3 and 4 describe how the memory cell is updated and how the output is calculated.

The LSTM algorithm has been customized to analyze the dataset's time series and sensor data. These analyses rely on data obtained from various sensors in the baby incubator and information related to the system's status. The model has been trained on an 80% data slice and tested on the remaining 20%. The predictive capabilities of LSTM are aimed at accurately forecasting future alarm situations based on sensor data within a specific time interval. This algorithm provides a valuable tool for monitoring the health of premature babies and intervening when necessary.

3.4 The Random Forest (RF) Algorithm

The Random Forest (RF) algorithm is another significant machine-learning technique employed in this study to predict the health status of premature babies.

Random Forest is a widely used learning algorithm in machine learning, and it does not have a specific general formula; its fundamental structure is based on decision trees. RF creates multiple decision trees by randomly sampling from the dataset. Through this sampling method, each tree is trained on a different subset. A decision tree is created for each random sample. These trees assign data points to specific classes or values using the features in the dataset, learning complex relationships in the data. For classification problems, the class is determined by a voting process of predictions from all generated decision trees. For regression problems, the final prediction is made by averaging the predictions from the trees [59]. Although the performance of the random forest classifier surpasses individual decision trees, it heavily depends on the structure of the dataset [60], [61]. Nevertheless, the classifier requires minimal configuration and can make reasonable predictions across a broad range of data.

This study uses the RF algorithm to analyze sensor data and time series. It focuses on predicting future alarm situations within a specific time frame by processing information from different sensors in the dataset.

3.5 Performance Metrics

To evaluate the performance success of the LSTM model in the study, the following metrics have been employed: MSE (Mean Squared Error), MAE (Mean Absolute Error), and MAPE (Mean Absolute Percentage Error).

Mean Squared Error (MSE):

$$MSE = \frac{1}{n}\sum_{n=1}^{n}(y^i - \hat{y}^i)^2 \tag{6}$$

MSE assesses predictions made in a single step. It is calculated by subtracting the observed value from the predicted value, squaring this difference, summing all squared values, and then dividing by the number of observations. As seen in Equation 6, this metric evaluates the average of the squares of the errors, indicating how close a regression line is to a set of points.

Mean Absolute Error (MAE):

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|y^i - \hat{y}^i\right| \tag{7}$$

MAE measures the average magnitude of the errors in a set of predictions without considering their direction. As seen in Equation 7, it averages the absolute differences between observed and predicted values.

MAPE (Mean Absolute Percentage Error):

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{\left|y^i - \hat{y}^i\right|}{|y^i|}\right)X100 \tag{8}$$

MAPE expresses accuracy as a percentage, which is calculated by taking the average of the absolute percentage errors by the number of observations. As seen in Equation 8, this metric is useful for understanding the accuracy of the prediction in percentage.

The accuracy value was used to evaluate the performance of the Random Forest algorithm. Accuracy is a metric that measures the accuracy rate of a classification model. Generally, it expresses the ratio of correctly predicted instances to the total instances, as seen in Equation 9. "Correct Predictions" represents the number of instances correctly classified by the model, and "Total Number of Samples" indicates the total number of evaluated instances. Thus, accuracy shows the ratio of correctly predicted instances to the total number of instances and is usually expressed as a percentage (%). For example, 80% accuracy indicates that the model correctly classified 80% of the instances. Accuracy is a good performance measure for balanced classes or equal class distributions. This metric is used to assess the overall performance of the model.

$$\textbf{Accuracy} = \frac{\textbf{Correct Predictions}}{\textbf{Total number of samples}} \tag{9}$$

4. Experiments and Results

All models used in this study were compiled on the Google Colab platform with GPU support. The open-source deep learning library Keras, a high-level Python language API, was used for all codes. Data processing was conducted using the Pandas and NumPy libraries.

4.1 Data Preparation for the Model

Initially, the dataset is grouped according to a specific column in the study. The original dataset is grouped monthly, and monthly sub-data frames are obtained from these groups. This is a common practice when working with time series data, aiming to analyze and model the data based on a specific period. Training data are obtained from a comprehensive dataset consisting of 13 different sensor parameters. These sensors include measurements taken from the incubator, as detailed in Section 3.1. The dataset consists of time series representing periods of 20 seconds each.

4.2 Training the LSTM Model and Error Prediction with the RF Algorithm

The input sequence determined for the LSTM model, a sequential series where each period is 20 seconds, has predicted states at the 21st and 22nd seconds. In other words, the shape of the input sequence is set as (20, 2). Sequential data processing models like LSTM can analyze such sequences and be used in tasks such as predicting future values or predicting the next period in a sequential series.

The labels for the Random Forest (RF) classifier were obtained from the "System Alarms" field within the dataset, which records predefined error conditions generated by the incubator system. A total of 11 distinct error classes were identified and utilized as output labels, including conditions such as high temperature, low temperature, sensor malfunction, heater failure, and oxygen imbalance. Each alarm type was encoded as a categorical variable to facilitate multi-class classification. Due to the imbalanced nature of the label distribution, appropriate measures were considered during the performance evaluation of the classification model.

During the initial prediction data acquisition, the LSTM layer structure contains 300 neurons, and the 'relu' activation function is used. The LSTM model is compiled with the 'Adam' optimizer and mean squared error (MSE) loss. Figure 2 provides a convergence chart showing the mean squared error (MSE) values on the training and validation datasets during the training of this LSTM model. The chart monitors the model's performance on the training and validation datasets during training. If the validation error rises while the training error drops, it may indicate overfitting, suggesting a decrease in the model's generalization ability. A reasonable balance between training and validation errors is sought. Therefore, the early stopping technique is applied in model training in this study, and training is stopped if the validation error does not decrease throughout 10 epochs.

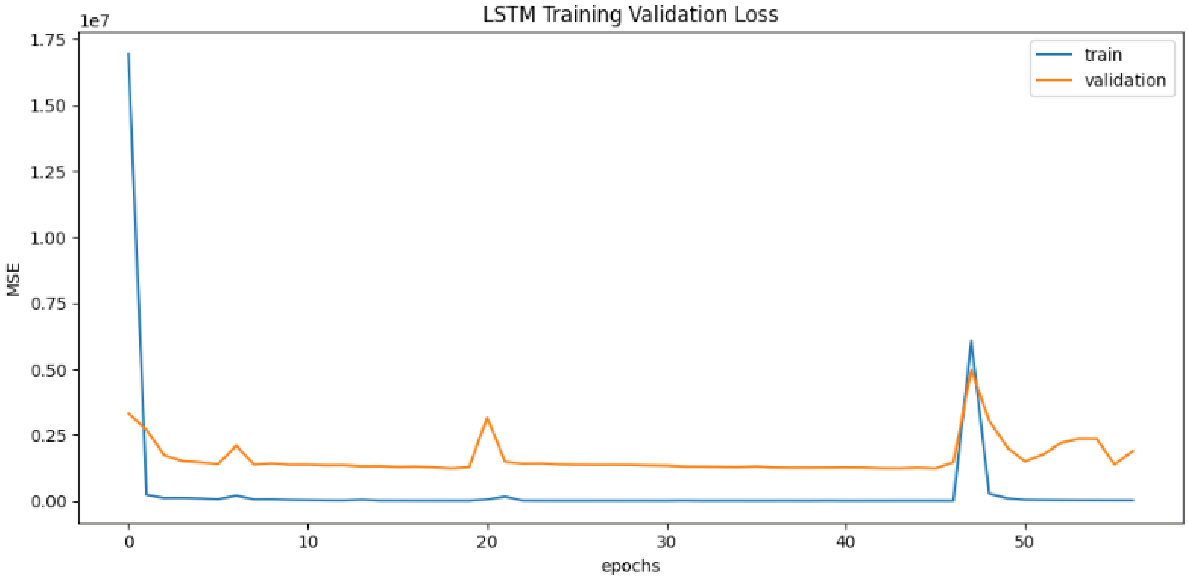

Figure 2. LSTM Training Graph

Later, experiments were repeated by changing the hyperparameters of the LSTM model to create the best-performing training conditions. Table 2 shows the hyperparameters of the LSTM model and their respective value ranges.

Table 2. LSTM Model Hyperparameters and Search Ranges

Model Hyperparameter Name	**Search Range for Optimal Hyperparameter**
Number of Epoch	[50, 75, 100]
Activation Function	[relu, sigmoid, tanh, softmax, linear]
Optimization Algorithm	[adam, RMSProp, AdaDelta]
Loss Function	[mse, binary_crossentropy]
Batch Size	[256]
Number of LSTM Units	[200, 250, 300, 400]
Input Number for Prediction	[10, 20, 40]

Then, the classification process with the RF algorithm was performed using the data predicted by the LSTM model. The classifier's predictions were compared with the errors of the values that the LSTM model should provide, and the model's success was determined. The parameter values used in the experimental studies and the performance results of the models are shown in Table 3. For example, in Table 3, it can be observed that an LSTM model trained with 256 units, 'relu' activation function, 100 epochs, and 'adam' optimizer has an MAE value of 163.2, MAPE value of 6.10 x 10^15, and MSE value of 11640.2. At the same time, the RF algorithm achieves 76.81% accuracy. The table includes important parameters such as the number of epochs, optimizer type, input sequence step, output step, and metrics used (MAE, MAPE, MSE).

Table 3. Comparison of LSTM and RF Model Performance with Various Hyperparameter Configurations

									LSTM			RF
Test	**Batch size**	**units**	**activation**	**epochs**	**optimizer**	**loss**	**In**	**Out**	**MAE**	**MAPE**	**MSE**	**Accuracy**
1	256	200	relu	100	adam	mse	20	2	163.2	6.10 x 10^15	11640.2	%76,81
2	256	200	relu	50	adam	mse	20	2	241.3	9.16 x 10^15	15830.5	%85,16
3	256	200	relu	75	adam	mse	20	2	114.0	7.03 x 10^15	94030.4	%80,79
4	256	200	relu	120	adam	mse	20	2	140.82	1.40 x 10^15	39590.8	%78.14
5	256	200	relu	50	RMSProp	mse	20	2	210.0	6.20 x 10^15	11910.9	%46.12
6	256	200	relu	50	AdaDelta	mse	20	2	5004.8	1.16 x 10^15	68040.0	%38.61
7	128	200	relu	50	AdaDelta	mse	20	2	4076.2	4.49 x 10^15	44840.0	%37.95
8	128	200	relu	50	adam	mse	20	2	153.8	5.99 x 10^15	13210.4	%70.13
9	256	200	sigmoid	50	adam	mse	20	2	6167.7	1.06 x 10^15	35059.0	%50.0
10	256	200	tanh	50	adam	mse	20	2	6206.9	1.07 x 10^15	34764.0	%11.16
11	256	200	Softmax	50	adam	mse	20	2	5206.9	1.06 x 10^15	24764.0	%20.16
12	256	200	linear	50	adam	mse	20	2	516.7	4.45 x 10^15	1858793.5	%21.84
13	256	200	relu	50	adam	b_cr	20	2	4182.2	9.47 x 10^15	24360.0	%11.16
14	256	300	relu	50	adam	mse	20	2	**160.8**	**9.01 x 10^15**	**1540533.6**	**%86.44**
15	256	400	relu	50	adam	mse	20	2	147.4	6.18 x 10^15	846971.94	%72.56
16	256	350	relu	50	adam	mse	20	2	225.5	4.52 x 10^15	1157109.9	%56.29
17	256	250	relu	50	adam	mse	20	2	181.9	7.44 x 10^15	1421834.6	%56.93
18	256	300	relu	50	adam	mse	40	2	-4.61	8.95 x 10^15	3214528.8	%69.36
19	256	300	relu	50	adam	mse	20	4	230.6	2.25 x 10^15	1362678.9	%63.5
20	256	300	relu	50	adam	mse	10	2	121.8	5.11 x 10^15	1228995.9	%51.3

By comparing the performance of models trained with different hyperparameters, the results of experiments shed light on determining the most effective hyperparameter combinations to optimize the prediction of the incubator conditions. As seen from Table 3, the RF model achieved the best performance with 86.44% accuracy when trained with an LSTM model with 256 units, 'relu' activation function, 50 epochs, 'adam' optimizer, and 'mse' loss, yielding MAE of 160.8, MAPE of 9.01 x 10^15, and MSE of 1540533.6.'

The LSTM model attempts to understand patterns in time series data using learned features and relationships. After processing and learning from the data, this model can predict the results at each step. Following the predictions of the LSTM model, an

RF classifier model comes into play. This model takes the outputs of the LSTM and classifies the states at each time step into specific alarm conditions. In other words, based on the outputs of the LSTM, it classifies the state at each time and predicts alarm conditions.

Finally, a confusion matrix is used to evaluate the performance of the RF model. The confusion matrix is a matrix that contains the numbers of correct and incorrect classifications by the model. As seen in Figure 3, the matrix displays the true and predicted values for alarm conditions within the test dataset. This matrix is visualized with a heatmap, allowing a visual understanding of which alarm conditions the model predicts better or worse. The heatmap is a colored matrix representation, providing insights to understand and improve the model's performance.

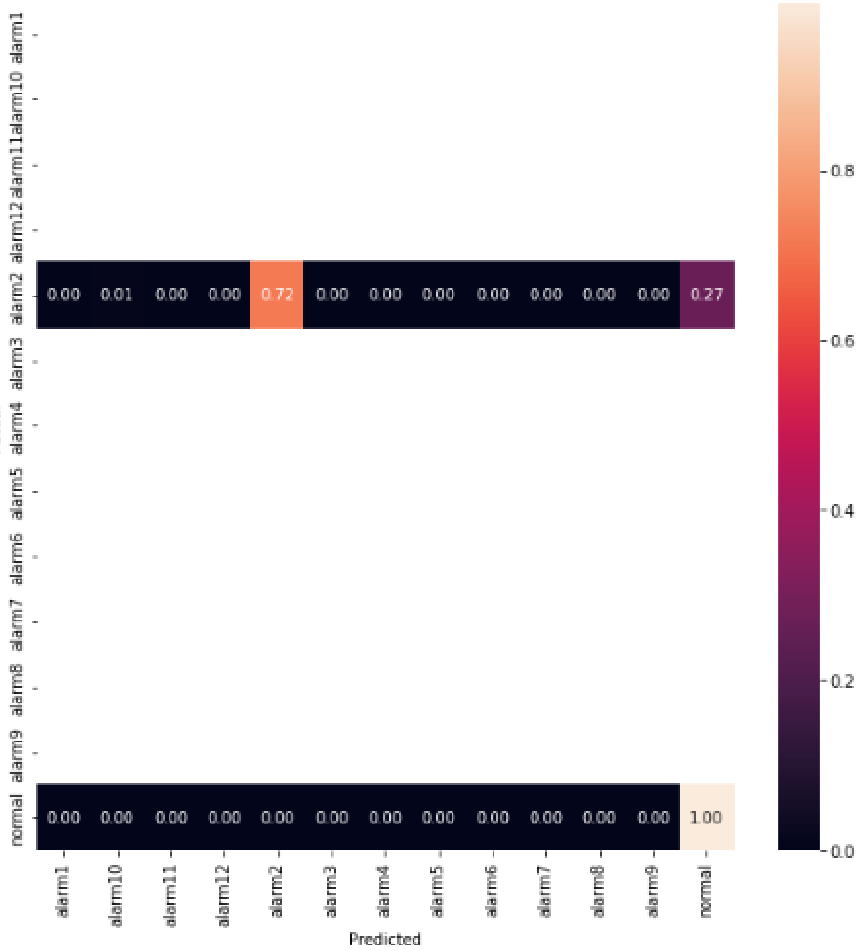

Figure 3. Confusion Matrix

5. Results and Recommendations

This study highlights the significant potential of artificial intelligence and deep learning techniques in neonatal

care. The results demonstrate that the LSTM-RF hybrid model is an effective tool for assessing the health status of premature infants. The algorithm successfully identified important patterns that could be associated with the health conditions of infants using sensor data.

The successful performance metrics of the RF algorithm, such as accuracy, sensitivity, and specificity, suggest that this model could be reliably used in neonatal care. However, it's essential to note that these successes were typically achieved on limited datasets, and further research is needed to explore the generalizability potential in different hospital environments.

Among the study's limitations are the limited scope of the dataset and the lack of information on infant sensor data. This limitation might restrict the model's application to a broader context. However, future studies incorporating larger datasets and additional sensor data could enhance the model's reliability and generalizability.

In conclusion, this study has shown that the accurate configuration of the model significantly influences the effectiveness of artificial intelligence models in neonatal care. Future research conducted on larger datasets and in different hospital environments will help us better understand how effective these models are in clinical applications.

References

[1] Goldenberg, R. L., & Rouse, D. J. (1998). Prevention of premature birth. *New England Journal of Medicine*, *339*(5), 313-320.

[2] Kumar, V., Shearer, J. C., Kumar, A., & Darmstadt, G. L. (2009). Neonatal hypothermia in low resource settings: a review. *Journal of perinatology*, *29*(6), 401-412

[3] Lunze, K., & Hamer, D. H. (2012). Thermal protection of the newborn in resource-limited environments. *Journal of Perinatology*, *32*(5), 317-324.

[4] H. Elayan, M. Aloqaily, and M. Guizani, "Digital Twin for Intelligent Context-Aware IoT Healthcare Systems," *IEEE Internet of Things Journal*, vol. 8, no. 23, pp. 16749–16757, Dec. 2021, doi: 10.1109/JIOT.2021.3051158.

[5] Y. Liu *et al*., "A Novel Cloud-Based Framework for the Elderly Healthcare Services Using Digital Twin," in *IEEE Access*, vol. 7, pp. 49088-49101, 2019, doi: 10.1109/ACCESS.2019.2909828

[6] Erol, T., Mendi, A. F., & Doğan, D. (2020, October). The digital twin revolution in healthcare. In *2020 4th international symposium on multidisciplinary studies and innovative technologies (ISMSIT)* (pp. 1-7). IEEE.

[7] A. Haleem, M. Javaid, R. Pratap Singh, and R. Suman, "Exploring the revolution in healthcare systems through the applications of digital twin technology," *Biomedical Technology*, vol. 4, pp. 28–38, Dec. 2023, doi: 10.1016/j.bmt.2023.02.001.

[8] M. Peshkova, V. Yumasheva, E. Rudenko, N. Kretova, P. Timashev, and T. Demura, "Digital twin concept: Healthcare, education, research," *Journal of Pathology Informatics*, vol. 14, p. 100313, Jan. 2023, doi: 10.1016/j.jpi.2023.100313.

[9] Y. Han, Y. Li, Y. Li, B. Yang, and L. Cao, "Digital twinning for smart hospital operations: Framework and proof of concept," *Technology in Society*, vol. 74, p. 102317, Aug. 2023, doi: 10.1016/j.techsoc.2023.102317.

[10] "A digital twin framework for prognostics and health management," *Computers in Industry*, vol. 150, p. 103948, Sep. 2023, doi: 10.1016/j.compind.2023.103948.

[11] R. Aluvalu, S. Mudrakola, U. M. V, A. C. Kaladevi, M. V. S. Sandhya, and C. R. Bhat, "The novel emergency hospital services for patients using digital twins," *Microprocessors and Microsystems*, vol. 98, p. 104794, Apr. 2023, doi: 10.1016/j.micpro.2023.104794.

[12] M. J. Kaur, V. P. Mishra, and P. Maheshwari, "The convergence of digital twin, IoT, and machine learning: transforming data into action," *Digital twin technologies and smart cities*, pp. 3–17, 2020.

[13] J. Kumari, R. Karim, K. Karim, and M. Arenbro, "MetaAnalyser - A Concept and Toolkit for Enablement of Digital Twin," *IFAC-PapersOnLine*, vol. 55, no. 2, pp. 199–204, Jan. 2022, doi: 10.1016/j.ifacol.2022.04.193.

[14] I. Kononenko and M. Kukar, *Machine learning and data mining*. Horwood Publishing, 2007.

[15] Attaran, M., & Celik, B. G. (2023). Digital twin: Benefits, use cases, challenges, and opportunities. *Decision Analytics Journal, 6*, 100165. https://doi.org/10.1016/j.dajour.2023.100165

[16] Dassault Systèmes. (n.d.). *The Living Heart Project*. https://www.3ds.com/products-services/simulia/solutions/life-sciences-healthcare/the-living-heart-project/

[17] A. Manocha, Y. Afaq, and M. Bhatia, "Digital Twin-assisted Blockchain-inspired irregular event analysis for eldercare," *Knowledge-Based Systems*, vol. 260, p. 110138, Jan. 2023, doi: 10.1016/j.knosys.2022.110138.

[18] Wahab, S. M. A. A., & Saad, M. (2022). Digital transformation acceleration in health sector during COVID-19: Drivers and consequences. *Journal of Business and Management Sciences*, *10*(4), 164-179.

[19] C. Quilodrán-Casas, V. L. S. Silva, R. Arcucci, C. E. Heaney, Y. Guo, and C. C. Pain, "Digital twins based on bidirectional LSTM and GAN for modelling the COVID-19 pandemic," *Neurocomputing*, vol. 470, pp. 11–28, Jan. 2022, doi: 10.1016/j.neucom.2021.10.043.

[20] D. Chen, N. A. AlNajem, and M. Shorfuzzaman, "Digital twins to fight against COVID-19 pandemic," *Internet of Things and Cyber-Physical Systems*, vol. 2, pp. 70–81, Jan. 2022, doi: 10.1016/j.iotcps.2022.05.003.

[21] Q. Lv, R. Zhang, X. Sun, Y. Lu, and J. Bao, "A digital twin-driven human-robot collaborative assembly approach in the wake of COVID-19," *Journal of Manufacturing Systems*, vol. 60, pp. 837–851, Jul. 2021, doi: 10.1016/j.jmsy.2021.02.011.

[22] H. Neog, P. E. Dutta, and N. Medhi, "Health condition prediction and covid risk detection using healthcare 4.0 techniques," *Smart Health*, vol. 26, p. 100322, Dec. 2022, doi: 10.1016/j.smhl.2022.100322.

[23] X. Shi, Z. Chen, H. Wang, D.-Y. Yeung, W.-K. Wong, and W. Woo, "Convolutional LSTM network: A machine learning approach for precipitation nowcasting," *Advances in neural information processing systems*, vol. 28, 2015.

[24] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997, doi: 10.1162/neco.1997.9.8.1735.

[25] C. Quilodrán-Casas, R. Arcucci, C. Pain, and Y. Guo, "Adversarially trained LSTMs on reduced order models of urban air pollution simulations," *arXiv preprint arXiv:2101.01568*, 2021.

[26] A. Elsheikh, S. Yacout, and M.-S. Ouali, "Bidirectional handshaking LSTM for remaining useful life prediction," *Neurocomputing*, vol. 323, pp. 148–156, Jan. 2019, doi: 10.1016/j.neucom.2018.09.076.

[27] G. Liu and J. Guo, "Bidirectional LSTM with attention mechanism and convolutional layer for text classification," *Neurocomputing*, vol. 337, pp. 325–338, Apr. 2019, doi: 10.1016/j.neucom.2019.01.078.

[28] O. Yeler and M. F. Koseoglu, "Performance prediction modeling of a premature baby incubator having modular thermoelectric heat pump system," *Applied Thermal Engineering*, vol. 182, p. 116036, 2021.

[29] R. Cuervo, M. A. Rodríguez-Lázaro, R. Farré, D. Gozal, G. Solana, and J. Otero, "Low-cost and open-source neonatal incubator operated by an Arduino microcontroller," *HardwareX*, vol. 15, p. e00457, Sep. 2023, doi: 10.1016/j.ohx.2023.e00457.

[30] P. T. Kapen, Y. Mohamadou, F. Momo, D. K. Jauspin, N. Kanmagne, and D. D. Jordan, "Development of a neonatal incubator with phototherapy, biometric fingerprint reader, remote monitoring, and heart rate control adapted for developing countries hospitals," *Journal of Neonatal Nursing*, vol. 25, no. 6, pp. 298–303, Dec. 2019, doi: 10.1016/j.jnn.2019.07.011.

[31] A. Hannouch, T. Lemenand, K. Khoury, and C. Habchi, "Heat and Mass Transfer of Preterm Neonates Nursed inside Incubators-A Review," *Thermal Science and Engineering Progress*, p. 100553, 2020.

[32] V. Puyana-Romero *et al.*, "Reverberation time measurements of a neonatal incubator," *Applied Acoustics*, vol. 167, p. 107374, 2020.

[33] A. Chandrasekaran *et al.*, "Disposable low-cost cardboard incubator for thermoregulation of stable preterm infant – a randomized controlled non-inferiority trial," *EClinicalMedicine*, vol. 31, p. 100664, Jan. 2021, doi: 10.1016/j.eclinm.2020.100664.

[34] A. Fraguela, F. D. Matlalcuatzi, and Á. M. Ramos, "Mathematical modelling of thermoregulation processes for premature infants in closed convectively heated incubators," *Computers in biology and medicine*, vol. 57, pp. 159–172, 2015.

Author(s) Contributions

This article was collaboratively written by three authors who contributed to the research and manuscript preparation as follows:

Hatice Kabaoğlu, Conceptualization, Writing – Original Draft, Software Development, Data Analysis. Fecir Duran, Supervision, Methodology, Conceptual Guidance, Technical Evaluation. Emine Uçar, Software Verification, Data Support, Writing – Review & Editing. All authors have read and approved the final version of the manuscript.

Conflict of Interest Notice

There is no conflict of interest regarding the publication of this paper.

Support/Supporting Organizations

The dataset used in this study was provided by OKUMAN Health Company (https://okuman.com.tr/). No additional institutional or financial support was received for this research.

Ethical Approval and Informed Consent

In this section, the author(s) declaration regarding research and publication ethics will be included. This title must be included for all articles. The name, date and number of the Ethics Committee can be given in this section.

It is declared that during the preparation process of this study, scientific and ethical principles were followed, and all the studies benefited from are stated in the bibliography.

Artificial Intelligence Statement

No AI tools were used to generate original research content, data analysis, or interpretation.

Plagiarism Statement

This article has been scanned by iThenticate ™