RESEARCH PAPER

ML Techniques for Infant Health Forecasting

Sanjay Makwana[1*], Seema Mahajan[2]
[1]Department of Computer Engineering, Indus University, Ahmedabad, India.
Email: sanjaymakwana.it@gmail.com; hod.ce@indusuni.ac.in. ORCID: https://orcid.org/0009-0005-1584-6257 (Corresponding Author)
[2]Department of Computer Engineering, Indus University, Ahmedabad, India

ABSTRACT
Newborns, particularly in the first few weeks of life are vulnerable to life-threatening diseases that may progress rapidly when not identified in a timely manner. An early diagnosis and prediction of neonatal diseases are necessary to increase the survival percentages and make correct treatment. However, limited clinical records and class imbalance are major obstacles for disease prediction. In this paper, we propose machine learning-based multi-class classification framework for neonate Diseases prediction. Due to limited availability of real clinical data, a synthetic dataset with 100,000 records including 63 neonatal and maternal parameters was created to mimic real life situations. The dataset was pre-processed in a systematic manner with feature engineering, class balancing, and feature selection based on ANOVA F-test. Different types of machine learning and deep learning models were implemented and compared. Among all models, LightGBM achieved the best accuracy of 79.60% and F1-Score of 78.47% with consistent performance across all disease categories. The study demonstrates that, with careful feature selection, class-imbalance treatment and ensemble modeling, neonatal disease forecasting may significantly improve. This framework provides a strong foundation in the direction of well performing clinical decision support systems to improve neonatal care.
Keywords: Neonatal disease prediction, ensemble modeling, LightGBM, class imbalance, feature selection
How to cite this article: Makwana S, Mahajan S. ML Techniques for Infant Health Forecasting. Int J Drug Deliv Technol. 2026;16(20s): 399-413. DOI: 10.25258/ijddt.16.20s.48
Source of support: Nil.
Conflict of interest: None

1 Introduction

Newborn babies, or infants less than 28 days old are called neonates. This is a very crucial period of human development. The situation becomes even more critical if the baby is premature or has a low birth weight. Premature infants usually require special care in the ICU known as Neonatal Intensive Care Unit (NICU). The NICU is designated to continuously monitor the baby's vital signs and help doctors take action if any medical complications occur.

In most case, a nurse or a dedicated staff is responsible for recording neonatal health data. They track important parameters such as heart rate, oxygen, temperature, blood pressure, gas, humidity and more. In this process, accuracy and timely data collection became extremely important. In developing countries like India, Newborn health is a major concern, especially due to large population and limited medical facilities. According to the National Health Report of India (2021&2022), the neonatal death rate remains high. Common causes include preterm birth, birth asphyxia, pneumonia, sepsis, malformation, diarrhea, and other serious diseases. Studying these reports can help identify trends and challenges in neonatal care and suggest ways to improve Newborn health outcomes. All over the world, child health issues are very common and critical. The World Health Organization (WHO) reported that 2.3 million children died in the first 20 days of life in 2022 ([1]).

Recent advances in machine learning and sensor technology have made it possible to develop smart devices that can monitor health in real time and also predict diseases in advance. However, there are still major challenges in this field. Often, the available data is unorganized, not in digital form, or based on very small sample sizes. Sometimes, it is also difficult to collect specific types of data needed for disease prediction.

Many researchers have worked on this topic using different combinations of sensors to collect neonatal health data. Most existing studies mainly focus on building monitoring systems, while only a few studies have explored disease prediction. In neonatal care, continuous monitoring is very crucial, but early disease prediction could be life-saving. To build such an intelligent system, proper planning is required. This includes selecting the right combination of sensors, designing an effective system, carefully implementing it, and applying strong domain knowledge. Along with real-time sensor data, historical datasets also play an important role. These datasets help the machine learning models learn patterns and make accurate predictions.

To understand the work done by other researchers, we collected papers from well-known databases such as PubMed, Google Scholar, Scopus, and Web of Science. Initially, 265 papers were selected. After carefully removing duplicate papers, out-of-scope studies, and unrelated work, 59 papers were finally selected for detailed review. The process of paper selection is shown in Fig 1. The rest of the paper is structured as follows. Section 2 covers the related work. Section 3 explains the proposed methodology, including data preparation, feature selection, model building, and the steps followed in model evaluation. Section 4 presents the experimental results along with the discussion. Finally, Section 5 concludes the paper and suggests possible directions for future research.

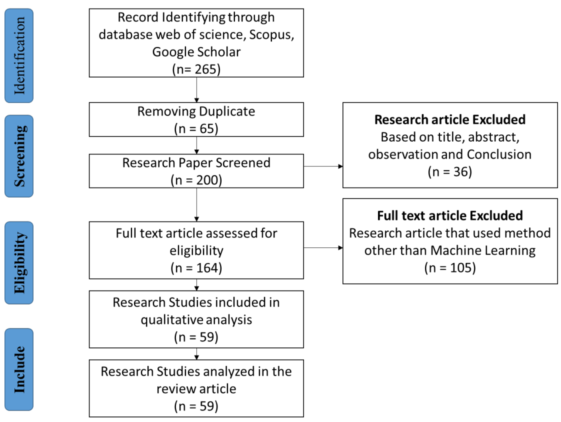

Fig. 1. PRISMA diagram for inclusive / exclusive criteria (adapted from [60])

2 Related Work

Many researchers have worked on neonatal health monitoring systems using different health parameters like temperature, humidity, heart rate, movement, sound, gas levels, and weight. To build a smart system, they have used sensors such as DHT11, LM35, pulse sensors, and many more. They have use microcontrollers like Arduino, NodeMCU, and ESP8266. These systems often use IoT platforms like ThingSpeak for real-time monitoring and sending alerts. Several studies have focused on creating low-cost, smart, and remote monitoring systems to improve Newborn care. Some researchers have also used machine learning and deep learning algorithms with datasets like MIMIC-III and NICU-collected data to predict diseases such as sepsis, jaundice, hypothermia, and asphyxia. These research efforts aim to provide early detection, continuous monitoring, and quick response to health problems. It can help reduce neonatal death rates and improve Newborn healthcare. The details of the measured parameters, sensors, and hardware used in these studies are provided in Table 1 and 2. An overview of the algorithms, targeted diseases, and datasets used by different researchers is presented in Table 3.

Table 1. Summary of Measured Parameters and Hardware Used

Year of Publication	Parameter Measured	Hardware Used	Reference
2017	Incubator Temperature	LM35 Sensor	[2]
2018	Light, Temperature, Pulse, Gas, Humidity	NodeMCU, ESP8266, DHT11, Light Sensor, Gas Sensor, Pulse Sensor, Arduino	[3]
2018	Weight, Humidity, Motion,	DHT11, Arduino, FSR-402,NFC Module	[4]

Year of Publication	Parameter Measured	Hardware Used	Reference
	Temperature		
2014	Temperature, Humidity	DHT11, Arduino	[5]
2021	Incubator Temperature, Humidity, Heart Rate, Movement	Arduino, DHT11, Oximeter MAX30100	[6]
2019	Temperature, Humidity, CO2, Water Reservoir level	Arduino, DHT11, CO2 Gas Sensor, Water Level/ Ultrasound sensor	[7]
2020	Temperature, Humidity	Arduino, DHT11, DHT22, ESP8266, NodeMCU	[8]
2021	Temperature, Humidity	LM35, Arduino, Humidity Sensor	[9]
2018	Temperature, Humidity, Heart rate	DHT11,KY039, NodeMCU, ESP8266	[10]
2019	Pulse Rate, Temperature, Vibration, Pressure	Heart Rate Sensor,LM35, MPX2050GSX, Arduino, Vibration, RFID	[11]
2021	Temperature, Humidity, Pulse Rate, Gas, Light	DHT11, Heart Rate Sensor, Gas Sensor, Light Sensor, ESP8266	[12]
2021	Temperature, Pulse Rate	Pulse Rate Sensor, LM35, ESP8266, Arduino	[13]
2021	Temperature, Humidity, Heart Rate	ATmega328 Microcontroller, Heart Rate Sensor, DHT22, Arduino	[14]
2020	Temperature, Humidity	DS18B20 sensor, DHT22, Noise Sensor, Air Flow Sensor	[15]
2021	Temperature, Weight, Heart Rate	DS18B20 sensor, Load Cell, Pulse Heart Rate Sensor,SIM800L	[16]
2019	Temperature,	DHT11	[17]

Year of Publication	Parameter Measured	Hardware Used	Reference
	Humidity,		
2021	Sound, Humidity	Arduino, NodeMCU, Sound Sensor, Humidity Sensor	[18]
2022	Temperature, Humidity, Sound	DHT22, KY037, Raspberry PI 3	[19]
2020	Temperature, Heart Rate	LM35, Pulse Sensor, Arduino, HC-05	[20]
2021	Temperature, Humidity, Heart Rate	ESP8266, DHT11, Pulse Sensor	[21]
2022	Temperature, Heartbeat, Sound, Weight	LM35,HX711, Sound Sensor, Pulse Sensor	[22]
2020	Temperature, Humidity, Pulse Rate, Gas,	LM35, DHT11, Q135,MAX30102,MQ3	[23]
2015	Temperature, Humidity,	Arduino, ESP8266, LM35, Heart Rate Sensor	[24]

Year of Publication	Parameter Measured	Hardware Used	Reference
	Heart Rate		
2020	Temperature, Humidity, Pulse Rate, Respiration, SPO2,	LM35,DHT11, MQ-3,MQ-135,	[25]
2021	Temperature, Humidity	Temperature Sensor, Humidity Sensor, Arduino	[26]
2018	Heart Rate, Body weight, body length	Load sensor, Pulse Sensor, Arduino	[27]
2021	Temperature, Humidity, Blood O2, Heart Rate,	DHT2,MAX30102,DS18B20, AD8232 ECG	[28]
2018	Temperature, Humidity, Heart Rate	Arduino, DS18B20,HT-BPM,ESP8826	[29]
2013	Temperature, Humidity	Microcontroller 18F4550, SYHS2XX, Temperature Sensor	[30]

Year of Publication	Parameter Measured	Hardware Used	Reference
2021	Temperature, Humidity	DHT22,DS1820	[31]
2015	Temperature, Humidity, Origination, Light	DS18B20, DHT11, LDR,MQ-7, Arduino,	[32]
2021	Temperature, Humidity, Heart Rate	Atmega328, MAX30100,MLX90614	[33]
2019	Temperature, Humidity	LM35, DHT22	[34]
2022	Temperature, Humidity	DHT11, Arduino	[35]
2017	Temperature, Humidity	Atmega328, LM356	[36]
2021	Temperature, Humidity	DHT-11, ATmega8535	[37]
2021	Temperature, Humidity, Sound, Gas, Heart Rate, Water Level, Vibration	DS18B20,DHT22, SEN-11574,LM393, SW-420,MQ-135, Wet Sensor	[38]
2020	Temperature, Humidity	DHT11, NodeMCU	[39]
2022	Temperature, Heart Beat, Sound, Gas,	Arduino, Temperature sensor, Heart Beat Sensor, Sound Sensor, Gas Sensor	[40]
2019	Temperature, Humidity, Heart Rate, Light, Gas	Pulse Sensor,MQ-6 Gas Sensor, DHT11 ,Light Sensor, PIC Microcontroller	[41]
2018	Temperature, Humidity	Arduino, DHT11,	[42]
2022	Temperature, Humidity, Sound,	DHT22,Max4466,ESP32	[43]
2019	Temperature, Humidity, Sound,	Arduino,LM35, DHT11,Sound sensor, wet sensor	[44]

Year of Publication	Parameter Measured	Hardware Used	Reference
	Water Level		
2019	Temperature, Movement, Heart Rate, Sound	Lm35,pir,pulse rate sensor, sound detection sensor, Arduino Uno	[45]
2019	Temperature, Pulse Rate	LM35, Pulse Sensor, Arduino, esp8266	[46]
2015	Temperature, Light, Weight, CO2	LM35, Gas sensor, Weight sensor, LDR	[47]
2019	temperature, humidity	DHT22,ESP8266 Node MCU	[48]
2020	temperature, humidity, heart rate, breathing	- DHT11,Heart beat sensor,-Respiration sensor, Raspberry Pi	[49]
2018	weight, temperature, head circumference of infant	NodeMCU ESP8266,Loadcell sensor, SHT Sensor),Ultrasonic HC-SR04 sensor	[50]
2017	Temperature, Weight	not specified	[51]

Table 1 shows the year-wise summary of the different parameters measured and the hardware used by researchers in neonatal monitoring studies. The table helps to see how both the health parameters and hardware choices have changed over time. Some studies used simple sensors and basic boards, while others moved to more advanced setups. By looking at the table, it's clear that researchers have used a variety of tools to track neonatal health in different ways.

Table 2. Overview of Parameters referred by researchers

Parameter Name	No of Time Referred in Paper	Parameter Name	No of Time Referred in Paper
Incubator Temperature	95	Wet	2
Humidity	80	Light	2
Heart Rate	32	Blood Pressure	2
Body Temperature	27	Air Flow	2
Sound	21	Vibration, Respiration	1
Weight	11	Air Quality	1
Gas	11	Water Level	1
Movement	8	Thermal Camera	1
O2	6	Ultra Sonic Sensor	1
Blood O2	5	Pressure Sensor	1

Co2	3	Conductivity	1
Light Sensor	3		

Table 2 gives an overview of the different health parameters that have been used by researchers for neonatal monitoring. Some of the common ones are temperature, heart rate, humidity, and oxygen levels. A few studies also looked at gas levels and sound, which are less commonly used. It can be seen that different studies focus on different sets of parameters, and there's quite a bit of variety in how neonatal health is tracked.

Table 3. Overview of Algorithms, Targeted Diseases, and Datasets

Year of Publication	Algorithm Used	Diseases Targeted	Dataset Name	Sample Size	Reference
2019	Not Specified	Not Given	Not Specified	4000	[52]
2016	Not Specified	Hypothermia	Not Specified	Not Specified	[53]
2017	Not Specified	Sepsis, jaundice, respiratory problems, prematurity, heart problems	MIMIC III	7863	[54]
2018	Not Specified	Premature birth, asphyxia	Not Specified	Not Specified	[55]
2019	Logistic regression with L2 regularization, Naïve Bayes, Support vector ma-chine (SVM) with a radial basis function kernel, K-nearest neighbours (KNN), Gaussian process, Random forest, AdaBoost, Gradient boosting	Sepsis	NICU at the Children's Hospital of Philadelphia (CHOP) between September 2014 and November 2017	2640	[56]

Year of Publication	Algorithm Used	Diseases Targeted	Dataset Name	Sample Size	Reference
2020	Logistic regression, Gaussian Naïve Bayes, decision tree, gradient boosting, adaptive boosting, bagging classifier, random forest, multilayer perceptron	Sepsis	MIMIC III	7870	[57]
2019	Not Specified	Sepsis	MIMIC III	1446	[58]
2019	Boosted decision trees	Not Specified	Not Specified	831	[59]

Table 3 gives an overview of the different algorithms that researchers have used for neonatal disease prediction. It also shows which diseases were targeted in each study and what kind of datasets were used. Some studies worked with real clinical data, while others used simulated or small-sized datasets. Looking at the table, it's clear that the choice of algorithm often depends on the type of disease and the available data.

2.1 Key Observations

- **Variation in Assessed Parameters:** Researchers have used a wide range of parameters to assess neonatal health. These include incubator temperature, humidity, heart rate, body temperature, sound level, weight, gas composition, movement, and oxygen saturation.
- **Diversity in Sensor Technologies:** Various sensor technologies have been used to collect neonatal health data. DHT11, DHT22, and LM35 for monitoring humidity and temperature, pulse sensors for measuring heart rate, CO_2 sensors for detecting gas levels and ultrasonic sensors for measuring weight are the commonly used sensors.
- **Variety in Hardware Platforms:** Various hardware platforms have been used by researchers to facilitate data processing and acquisition. These include Arduino boards (such as Uno and Mega), NodeMCU, ESP8266, Raspberry Pi, and application-specific microcontrollers like Atmega328 and Atmega8535.

- **Underutilization of Available Data:** Although neonatal incubators and hospitals generate large amounts of data, this data—along with historical health records—is not being used effectively for disease prediction and diagnosis.
- **Issues with Data Standardization and Availability:** Many studies rely on manually recorded, non-standardized data. This makes it difficult to compare results across studies and limits the ability to reproduce experiments.
- **Incomplete Use of Clinical Features**: Even though multiple clinical parameters are available, most studies only use a small number of them. Manual data collection further restricts the ability to capture complete data, reducing the potential for real-time disease prediction.
- **Limited Application of Machine Learning:** Despite significant advancements in machine learning, very few researchers have applied these techniques to neonatal disease prediction. Most existing studies focus mainly on monitoring and alarm systems. The use of sensor data for predictive purposes remains an underexplored area.

3 Methodology

The complete workflow of the neonatal disease prediction system is shown in Fig. 2. The flowchart provides a step-by-step summary of the entire process, starting from data collection and ending with the final model selection. Each step is explained in detail in this sections.

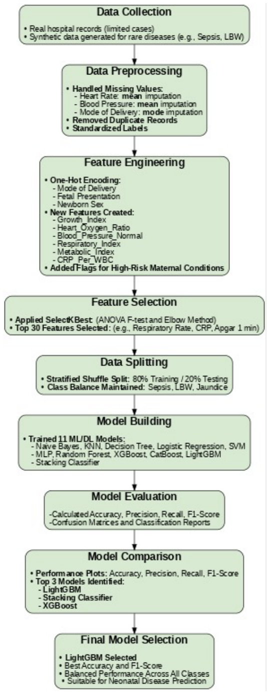

Fig. 2. Flow chart of methodology

The synthetic dataset[61] was loaded, and an initial review was done to get a basic understanding of the features, the structure, and the class labels. A few simple checks were made to see how the data was spread. To explore this further, histograms were plotted for the numerical features. These plots made it easier to see the spread and shape of the data. Some features, like heart rate and blood glucose, looked fairly even, while others, such as temperature and birth weight, showed more variation. Taking a look at these patterns gave a useful starting point for the later steps. The histograms are shown in Fig. 3.

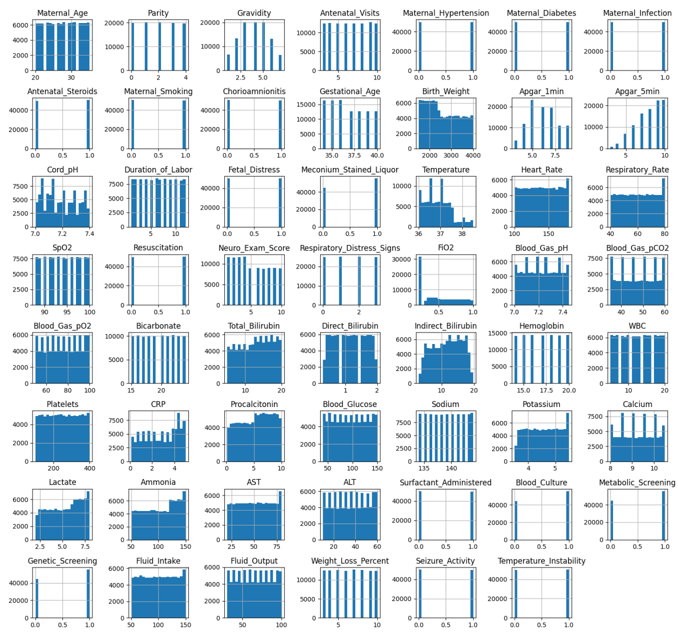

Fig. 3. Histogram of Numeric features

After this, the class distributions were checked to find any imbalance, and missing values were carefully reviewed. Pre-processing was done step by step. Missing values were filled using imputation where needed. Categorical features were changed into numerical form using One-Hot Encoding. Numerical features were scaled using the StandardScaler so that all features would be on a similar range. Once this was done, the dataset was split into training (80%) and testing (20%) sets for model evaluation.

Feature engineering was done to create new features that could help the model learn better. For example, a Growth Index was calculated by dividing birth weight by gestational age. This was added to capture growth patterns that may be useful for disease prediction. Since the dataset was imbalanced, random oversampling was applied to make sure each disease class had enough samples. This helped the models to learn fairly across all classes. Feature selection was carried out using the ANOVA F-test with the SelectKBest method. The Elbow method was used to choose the best number of features. It was found that keeping 30 features worked best. The feature selection curve is shown in Fig. 4.

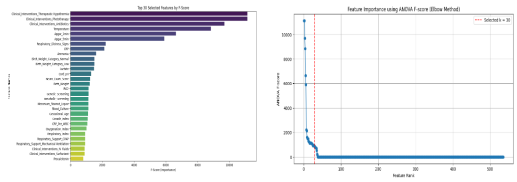

Fig. 4. Top 30 Selected Features and Elbow method for features selection

Different machine learning and deep learning models were trained using the final dataset. Traditional models like Logistic Regression, Decision Tree, K-Nearest Neighbours (KNN), Naive Bayes, and Support Vector Machine (SVM) were tested along with ensemble models such as LightGBM, XGBoost, and CatBoost. A Multi-Layer Perceptron (MLP) was also used to try a basic deep learning approach. Finally, a Stacking Classifier was built by combining the best models into one. The variables used in the datasets are listed in Table 4.

Table 4. Variables used in Datasets

Category	Variable Name
Maternal Information	Maternal_Age, Parity, Gravidity, Antenatal_Visits, Maternal_Hypertension, Maternal_Diabetes, Maternal_Infection, Antenatal_Steroids, Maternal_Smoking, Chorioamnionitis
Neonatal Information	Gestational_Age, Birth_Weight,Mode_of_Delivery, Fetal_Presentation, Apgar_1min,Apgar_5min, Cord_pH,Newborn_Sex
Initial Neonatal Vitals	Temperature, Heart_Rate, Respiratory_Rate, Blood_Pressure,SpO2, Resuscitation, Respiratory_Support
Clinical Indicators	Neuro_Exam_Score, Respiratory_Distress_Signs, FiO2, Blood_Gas_pH, Blood_Gas_pCO2, Blood_Gas_pO2,Bicarbonate, Chest_Xray
Laboratory Results	Total_Bilirubin, Direct_Bilirubin, Indirect_Bilirubin,Hemoglobin, WBC, Platelets, CRP, Procalcitonin, Blood_Glucose, Sodium, Potassium, Calcium, Lactate, Ammonia,AST, ALT
Advanced Diagnostic Tests	Cranial_Ultrasound, Echocardiography, Surfactant_Administered, Blood_Culture, Metabolic_Screening, Genetic_Screening
Fluid Management	Fluid_Intake, Fluid_Output, Weight_Loss_Percent,Dehydration_Signs
Critical Conditions & Interventions	Seizure_Activity, Temperature_Instability, Clinical_Interventions, Response_to_Treatment
Target Variable	Predicted_Disease

The models were checked using accuracy, F1-Score, confusion matrix, and ROC curve. This gave a good idea of how well the models performed, both overall and for each disease category. The results for all models are shown in Table 5. The Accuracy vs. F1-Score comparison is displayed in Fig. 5. All results, trained models, and graphs were saved for later use and for future reference.

4 Experimental Results

The proposed system was tested for predicting different neonatal diseases using several machine learning and deep learning models. Each model was checked using accuracy, F1-Score, and confusion matrix to see how well it could predict across all disease classes. A quick summary of the model performances can be found in Table 5.

Table 5. Model Performance Comparison

Model	Accuracy (%)	F1-Score (%)
LightGBM	79.60	78.47
CatBoost	79.51	78.44
XGBoost	79.41	78.40
Stacking Classifier	78.93	77.98
Multi-Layer Perceptron	78.64	77.72
Random Forest	78.85	77.70
Support Vector Machine	76.27	74.85
Logistic Regression	72.53	71.57
Decision Tree	70.99	71.17
K-Nearest Neighbors	61.91	59.48
Naive Bayes	59.17	51.78

When going through the results, it was quite clear that ensemble models like LightGBM, CatBoost, and XGBoost worked better than the traditional models. Out of all the models tested, LightGBM seemed to perform the best, reaching an accuracy of 79.60% and an F1-Score of 78.47%. CatBoost and XGBoost also gave good results, though their scores were a little lower when compared to LightGBM. On the other hand, the traditional models like Logistic Regression, Decision Tree, Naive Bayes, and K-Nearest Neighbors (KNN) did not perform as expected. These models showed more errors, especially when classifying the more difficult or overlapping disease categories.

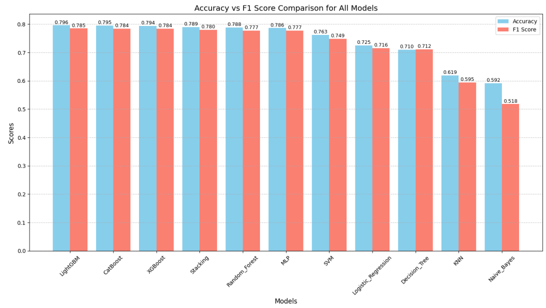

Fig. 5. Accuracy vs F1 score comparison

The Accuracy vs. F1-Score plot, which is shown in Fig. 5, helped in quickly comparing the models. It was easy to see that the ensemble models generally stayed ahead in both accuracy and class-wise prediction. A closer look at the confusion matrices also gave some useful information. These matrices helped in understanding how each model handled the different disease classes one by one. LightGBM and CatBoost gave more balanced predictions across all the classes. XGBoost and the Stacking Classifier worked reasonably well but showed a few drops in precision, especially for some of the smaller disease categories.

It was also noticeable that the traditional models made more mistakes when the diseases had overlapping features. In those cases, their predictions were not very reliable, which was expected based on how these simpler models learn.Overall, LightGBM stood out as the most reliable model in this study. It had the highest accuracy, good F1-Score, and was also faster to run compared to others. The steady performance across different disease types and its quick processing time make it a strong choice for this task.

Another thing that became clear during this study was how much class balancing and feature selection mattered. Random oversampling really helped the models to learn more fairly across all classes. Also, careful feature selection allowed the system to focus on the most important data, without losing prediction strength. Even though the study used synthetic data, the system worked well, and the overall process seems useful for developing future neonatal healthcare solutions.

The results showed that ensemble models like LightGBM and CatBoost gave the most balanced predictions. They performed consistently well across all disease categories. XGBoost and Stacking classifiers also gave good results, but they had slightly lower accuracy in predicting some of the smaller disease classes. Traditional models like Logistic Regression, Decision Tree, K-Nearest Neighbors, and Naive Bayes made more mistakes. These models struggled more with diseases that had overlapping patterns.From all the models tested, LightGBM came out as the best performer. It showed the highest accuracy and F1-Score while also being faster to compute. This study shows that ensemble models can handle complex neonatal data better than traditional ones. It also shows that taking care of class imbalance and choosing the right features can improve predictions. Even though synthetic data was used, the process worked well and can guide future research in neonatal disease prediction. This system could help build smarter healthcare solutions in the future.

5 Conclusion and Future Work

This study focused on building a machine learning system for predicting neonatal diseases. The dataset used was synthetic, designed carefully to follow clinical patterns as much as possible. However, some differences from real hospital data might still exist. Multiple machine learning and deep learning models were tested during this work. From the results, it appeared that ensemble models, especially LightGBM, performed better overall. LightGBM provided higher accuracy and gave more balanced predictions across disease categories. Traditional models did not perform as well when tested on the same dataset.

One thing that became obvious while working on this system was how much class balancing mattered. Random oversampling helped the models learn from all disease groups more fairly, which likely improved their prediction results. Another thing noticed was the effect of feature selection. It helped to reduce unnecessary data while keeping the important parts, which made the system more efficient. Even though synthetic data was used, the system worked well enough to suggest that this method could have practical value in neonatal healthcare.

There are still things to improve. Future work should test this system using real clinical datasets collected from multiple hospitals. This would help to confirm whether the system can really perform in real hospital settings. Also, using more advanced deep learning models might give better results, especially if the data grows larger and more complex. It would also help to add explain ability tools like SHAP or LIME. These can give doctors clearer reasons for why the system makes certain predictions. This could make it easier for the system to be trusted and used in real practice.

Overall, this study offers a good starting point. With more development, systems like this might help doctors make faster decisions and improve care for Newborns.

Statements and Declarations

Funding: No funding was received for conducting this study.

Competing Interests: The authors have no relevant financial or non-financial interests to disclose.

Ethical Approval: This article does not contain any studies with human participants or animals performed by any of the authors.

Data Availability Statement

The dataset utilized in this study is publicly available and can be accessed via Zenodo. The full dataset description and download link are provided in the following reference [61]: S. Makwana and S. Mahajan, “A Synthetic, Clinically Inspired Neonatal Dataset for Multi-Class Disease Prediction and Trust-Aware Machine Learning”, Zenodo, Jan. 13, 2026, doi: 10.5281/zenodo.18329706.

REFERENCES

[1] World Health Organization 2024 *Newborn* mortality: fact sheet. World Health Organization Geneva

[2] Babhulkar A 2017 *Temperature* monitored IoT based intelligent system. In: Proc. 2017 Int. Conf. I-SMAC (IoT in Social, Mobile, Analytics and Cloud) Palladam, India. New York: IEEE p. 1–6

[3] Sivamani Chinnaswamy Rajasekar R and Seshadri A 2018 *Intelligent* system using Internet of Things. Int J Mod Trends Sci Technol 4(9): 23–27

[4] Sendra S Ureña R Lloret J and Canarias M 2018 *Smart* infant incubator based on LoRa networks. In: 2018 IEEE/ACS 15th Int. Conf. Computer Systems and Applications (AICCSA) Agadir, Morocco. New York: IEEE p. 1–6

[5] Nisha ZA and Elahi A 2014 *Low* cost neonatal incubator with smart control system. In: 8th Int. Conf. Software, Knowledge, Information Management & Applications (SKIMA) Dhaka, Bangladesh. Dhaka: IEEE p. 1–5

[6] Ahmed F Sufyan A Hussain S Rehman A Amin S and Salam Y 2021 *Smart* baby incubator. In: 2021 Int. Conf. Computing, Electronic and Electrical Engineering (ICE Cube) Dubai, UAE. New York: IEEE p. 1–4

[7] Mohammed Sheet SS Mohammed ZG Khaleel KN and Abbas AA 2019 *Smart* infant incubator based on mega microcontroller. In: 2019 2nd Int. Conf. Engineering Technology and its Applications (IICETA) Colombo, Sri Lanka. New York: IEEE p. 1–5

[8] Shabeeb AG Al-Askery AJ and Nahi ZM 2020 *Remote* monitoring of a premature infant's incubator. In: 2018 IEEE/ACS 15th Int. Conf. Computer Systems and Applications (AICCSA) Agadir, Morocco. New York: IEEE p. 1232–1238

[9] Latif A Malik R and Khan S 2021 *Temperature* monitoring system for baby incubator based on Visual Basic. J Robotics Control 2(3): 47–50

[10] Ghosh A Mandal S Banerjee S Pal S and Bhattacharjee A 2018 A smart incubation system. *In:* 2018 IEEE Int. Conf. Recent Innovation in Electrical, Electronics & Communication Engineering (ICRIEECE) Bangalore, India. New York: IEEE p. 1–5

[11] Subramanian M Kumar A Singh P and Kumar S 2019 *Security* and health monitoring system of the baby in incubator. Int J Eng Adv Technol 8: 3582–3585

[12] Abd El-Aziz RM and Taloba AI 2021 *Real* time monitoring and control of neonatal incubator using IoT. arXiv preprint arXiv:2112.02071

[14] Anonymous 2021 *Smart* Asher infant incubator for accurate monitoring and control. Int J Emerg Technol Innov Res 8: c531–c536

[15] Ainiyah S Nurjasmi D Sari R Nurfadilah S and Rahmawati L 2020 *Development* of incubator analyzer based on computer with temperature and humidity parameters. J Electronics Electromedical Eng Med Inform 2(2): 48–57

[16] Irmansyah M Nugroho AH Wardhani NP Hidayat R and Rahmawati D 2021 *Monitoring* on portable baby incubator based on microcontroller and notification using short message service (SMS). Indones J Electronics Electromedical Eng Med Inform 3(4): 140–147

[17] Inam S Farooqui NM Khan HA Shaikh M and Qureshi R 2019 *Android* based internet accessible infant incubator. In: 2019 8th Int. Conf. Inf Commun Technol (ICICT) Bangkok, Thailand. New York: IEEE p. 1–5

[18] Bhasha P Sreelekshmi T Sharma S Anu K and George S 2021 *An* IoT-based Blynk server application for infant monitoring alert system to detect crying and wetness of a baby. In: ISCDA 2020: Int. Conf. Intelligent & Smart Comput. Data Anal. Kuala Lumpur, Malaysia. Cham: Springer p. 1–7

[19] Irawan B Hadi N Putra AC Pratama R and Fadillah S 2022 *Design* and development of a baby sleep monitoring system based on Internet of Things (IoT). Int J Sci Technol Manage 3(4): 835–844

[20] Sheril Amira O Aziz AMA Taloba AI Hassan M and Ibrahim R 2020 *Neonatal* health monitoring system with IoT application. J Phys Conf Ser 1529(5): 052071

[21] Sony MD Sharma R Patel K and Chouhan M 2021 *IoT* based infant healthcare monitoring system. J Phys Conf Ser 1937(1): 012045

[22] Suthagar MG and Tamilsevana KS 2021 *Baby* incubator monitoring system using global system for mobile technology. J Kejuruteraan

[23] Shalannanda W Noor A Hamzah T and Zulkifli N 2020 *Implementation* of the hardware module of IoT-based infant incubator monitoring system. In: 2020 14th Int. Conf. Telecommun. Syst., Serv. Appl. (TSSA) Penang, Malaysia. New York: IEEE p. 1–5

[24] Wahab MA and Nor DM 2021 *Safety* and health monitoring system for baby incubator using IoT. Evol Elec Electron Eng 2: 256–264

[25] Shalannanda W Zakia I Sutanto E and Fahmi F 2020 *Design* of hardware module of IoT-based infant incubator monitoring system. In: 2020 6th Int. Conf. Wireless Telematics (ICWT) Bandung, Indonesia. New York: IEEE p. 1–6

[26] AlQaheri H Hasan SF Ahmed NAH and Ramzi M 2021 *Toward* an autonomous incubation system for monitoring premature infants. Stud Inform Control 30(4): 121–131

[27] Purwiyanti S Sulistiyanti SR Arinto Setyawan FX Wibisono BM Atmaja KS and Fitriawan H 2018 *Multisensors* system for real time detection of length, weight, and heartbeat of premature baby in the incubator. In: ICECOS 2018: Int. Conf. Electr. Eng. Comput. Sci. Mataram, Indonesia. New York: IEEE p. 1–5

[28] Kalilani M Lai NS and Abdulla R 2021 *IoT* based neonatal incubator for the developing world and conflict zones. J Appl Technol Innov 5(4): 44–50

[29] Koli M Ladge P Prasad B Boria R and Balur NJ 2018 *Intelligent* baby incubator. In: ICECA 2018: 2nd Int. Conf. Electron. Commun. Aerosp. Technol. Coimbatore, India. New York: IEEE p. 1–4

[30] Joshi NS Kamat RK and Gaikwad PK 2013 *Development* of wireless monitoring system for neonatal intensive care unit. Int J Adv Comput Res 3(3): 106–111

[31] Luthfiyah S Kristya F Wisana IDGH and Thaseen M 2021 *Baby* incubator monitoring center for temperature and humidity using Wi-Fi network. J Electronics Electromedical Eng Med Inform 3(1): 8–13

[32] Mittal H Mathew L and Gupta A 2015 *Design* and development of an infant incubator for controlling multiple parameters. Int J Emerg Trends Electr Electron 11(5): 65–72

[33] Al-Asadi HA Hasan SF Ahmed NAH and Ramzi M 2021 *Internet* of hospital things (IoHT): Designing a monitor system of healthcare for neonates. Iraq Univ Coll J Eng Appl Sci 2(1): 45–52

[34] Widhiada W Utami ZA Prasetyo A and Handayani D 2019 *The* robust PID control system of temperature stability and humidity on infant incubator based on Arduino at Mega 2560. IOP Conf Ser Earth Environ Sci 248: 012034

[35] Palled A GA Bhoomika SL and Shobha G 2022 *Infant* incubator monitoring system. Int Res J Modernization Eng Technol Sci 4(2): 112–118

[36] Oyebola B 2017 *Development* of a microcontroller based temperature and humidity controller for infant incubator. J Ind Technol 2(4): 78–89

[37] Latif A Pasha U Khan M and Sadiq S 2021 *Temperature* and humidity controlling system for baby incubator. J Robotics Control 2(3): 190–193

[38] Kolhe S Mishra SS and Shah MS 2021 *Improved* design of IoT based infant incubator monitoring and control system. Int Res J Eng Technol 8(9): 2837–2842

[39] Nwaneri S Sojobi JW Oyelade AO and Ezenwa BN 2020 A low cost *IoT* based neonatal incubator for resource poor settings. West Indian J Eng 42(2): 45–52

[40] Labhade ST Mali DA Thakur SM and Patil V 2021 *IoT* based smart crib for baby surveillance. Int Res J Modernization Eng Technol Sci 3(4): 67–73

[41] Rakhmawati R Irianto I Murdianto FD Luthfi A and Rahman AY 2019 *Thermal* optimization on incubator using fuzzy inference system based IoT. Int J Adv Res Ideas Innov Technol 5(1): 69–74

[42] Ali M Abdelwahab M Awadekreim S and Abdalla S 2018 *Development* of a monitoring and control system of infant incubator. In: ICCCE 2018: Int. Conf. Comput. Control Electr. Electron. Eng. Cairo, Egypt. New York: IEEE p. 1–4

[43] Paramartha IKN Hamzah T Utomo B and Rachman F 2022 *Lost* data and transmission speed analysis on incubator analyzer based IoT technology. Int J Adv Health Sci Technol 2(1): 39–46

[44] Raharja Iswanto NMA Sulistyo ME Prasojo I and Anindiyahadi F 2019 *Monitoring* baby conditions in the baby box based on IoT technology. J Adv Res Dynam Control Syst 11(Special Issue-08): 1089–1095

[45] Karthick R Rinoj B Venish Kumar TV Prabaharan AM and Selvaprasanth P 2019 *Automated* health monitoring system for premature fetus. Asian J Appl Sci Technol 3: 17–23

[46] Najib SM Hassan NB Mohd Shair NTA Rahim N and Ab Wahab N 2019 *Intelligent* neonatal monitoring system based on android application using multi sensors. In: 2019 IEEE 9th Symp. Comput. Appl. Ind. Electron. (ISCAIE) Langkawi, Malaysia. New York: IEEE p. 131–135

[47] Turkane SM Rode V Rahane S and Shinde S 2015 *Intelligent* neonatal monitoring based on smartphone. Int J Inform Futur Res 2(3): 754–761

[48] Netto SPN Chacon PFS de Morais Junior AMC Costa CLBR and Brasil FL 2019 *Remote* monitoring of temperature and humidity—a reliable and inexpensive device development applied in neonatal

incubators. In: XXVI Brazilian Congr Biomed Eng (CBEB 2018) Búzios, Brazil. Cham: Springer p. 695–699

[49] Kshirsgar P More V Hendre V Chippalkatti K and Paliwal K 2020 *IoT* based baby incubator for clinic. In: ICCCE 2019: 2nd Int. Conf. Commun. Cyber Phys. Eng. Jaipur, India. Cham: Springer p. 349–355

[50] Firmansyah R Widodo A Romadhon AD Hudha MS Saputra PPS and Lestari NA 2019 *The* prototype of infant incubator monitoring system based on the Internet of Things using NodeMCU ESP8266. J Phys Conf Ser 1171: 012015

[51] de Oliveira Filho JI and da Mota Almeida O 2017 *Cloud-based* monitoring system and risk management for premature newborns. In: 2017 2nd Int. Symp. Instrum. Syst. Circuits Transducers (INSCIT) Fortaleza, Brazil. New York: IEEE p. 1–6

[52] Ibrahim DM Hammoudeh MA Ambreen S and Mohammadi S 2019 *Raspberry* Pi-based smart infant monitoring system. Int J Eng Res Technol 12(10): 1723–1729

[53] De D Mukherjee A Sau A and Bhakta I 2017 *Design* of smart neonatal health monitoring system using SMCC. Healthc Technol Lett 4(1): 13–19

[54] Singh H Yadav G Mallaiah R Joshi P Joshi V Kaur R Bansal S and Brahmachari SK 2017 *INICU–Integrated* neonatal care unit: Capturing neonatal journey in an intelligent data way. J Med Syst 41(8): 132

[55] Puneeth C Lepcha S Jain SS Kumar S Batra AS Saw AK Singhal N Shukla S and Unal A 2019 *Smart* neonatal intensive care for urban hospitals. J Adv Res Dynam Control Syst 11(Special Issue-08): 1372–1378

[56] Masino AJ Harris MC Forsyth D Ostapenko S Srinivasan L Bonafide CP Balamuth F Schmatz M and Grundmeier RW 2019 *Machine* learning models for early sepsis recognition in the neonatal intensive care unit using readily available electronic health record data. PLoS One 14(2): e0212665

[57] Song W Jung SY Baek H Jang JY and Lee J 2020 A predictive model based on machine learning for the early detection of late-onset neonatal sepsis: *Development* and observational study. JMIR Med Inform 8(7): e15965

[58] Thakur J Pahuja SK and Pahuja R 2019 *Non-invasive* prediction model for developing countries to predict sepsis in neonates. Biomed Eng Appl Basis Commun 31(01): 1950001–1950008

[59] Semenova O Kurbatova NV Lester M Fedorov A Maksyutov R and Minullina L 2019 *Prediction* of short-term health outcomes in preterm neonates from heart-rate variability and blood pressure using boosted decision trees. Comput Methods Programs Biomed 180: 104996

[60] Page MJ Moher D Bossuyt PM Boutron I Hoffmann TC Mulrow CD and Shamseer L 2021 PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. BMJ 372: n160

[61] S. MAKWANA and S. MAHAJAN, "A Synthetic, Clinically Inspired Neonatal Dataset for Multi-Class Disease Prediction and Trust-Aware Machine Learning". Zenodo, Jan. 13, 2026. doi: 10.5281/zenodo.18329706.