

ORIGINAL ARTICLE

EARLY PREDICTION OF NEONATAL SEPSIS USING MACHINE LEARNING ALGORITHMS

Lakshman Narayanaswamy¹, Arvind Kasaragod², Kishore Kumar³, Jayalakshmi Srinivasan².

¹Ketaki solutions, Bangalore, Karnataka, India,

²Department of Pediatrics, Cloudnine Hospitals, Bangalore, Karnataka, India,

³Department of Paediatrics and Neonatology, Cloudnine Hospital, Jayanagar, Bangalore, Karnataka, India.

ABSTRACT

Objective: We aim to develop a machine learning model based on data from Indian hospitals to predict neonatal sepsis several hours ahead of the current diagnosis used by hospitals to determine sepsis.

Methods: Vitals data of neonates from nine Bangalore hospitals ICU was collected. The data was collated and normalized and then used as features for training a LSTM machine learning model built to predict sepsis 12 hours ahead. To train the model, patients with culture positive were taken as sepsis true and the rest were taken as sepsis false.

Results: Two models, one with hourly data input of 12 hours and the other with hourly data input of 24 hours were built. Both models predicted sepsis 12 hours ahead of the last given hourly input. Receiver Operating characteristics, an important model performance metric for the 12 hours model was 0.9324 and the 24 hour model was 0.8703.

Conclusion: The performance metrics of the machine learning model built using India neonatal vitals data shows lots of promise in aiding physicians in early detection on neonatal sepsis. The model requires field testing and collaboration with hospitals to ensure reliable data for broader adoption in neonatal sepsis detection.

Introduction

Neonatal Sepsis is a major cause of neonatal mortality rate in LMIC especially India, accounting for more than 50% neonatal deaths as per neonatal data currently available. The gold standard of neonatal sepsis detection is culture positive and for many babies the time taken to become culture positive may be too late. The goal of this study is to use Artificial Intelligence (machine learning models) to evaluate babies vital data from neonatal intensive care units and predict which of the babies are going to develop sepsis even before they become septic. This machine learning model will serve as a tool to predict sepsis early so that interventions can be put in to save the babies before it is too late.

Methods

The neonatal data for this study was collected from nine CloudNine tertiary neonatal intensive care units of perinatal hospitals located in various parts of Bangalore. These hospitals use an Electronic Medical Records (EMR) system out of which patient data was extracted for further processing. Bulk access to patient data is not supported by the EMR vendor resulting in a time consuming process of extracting individual patient data. Data available from the EMR is post 2019. The hospital also has a critical care monitoring system

Address for Correspondance: Lakshman Narayanaswamy, Ketaki solutions, MS Electrical Engineering, Bangalore, 560011, INDIA.

Email: nswamy.lakshman@gmail.com

©2026 Pediatric Oncall

ARTICLE HISTORY

Received 29 April, 2026

Accepted 26 May, 2026

KEYWORDS

Artificial intelligence, LSTM, Machine learning, Neonatal sepsis, Neonatal mortality.

that records vitals data periodically. Some patient data post 2021 was available from the critical care software. Patient data was made available in various formats. Demographic details and vitals reading were available as a PDF file, lab data when available was given in an Excel format. Patient data had to be corroborated between these various files to build and analyze a consistent data view for each patient. PDF file processing is done using PDF library, the primary key correlates the various patient files is the unique patient id. Vital data is extracted from PDF files and correlated with lab and sepsis results from csv files. All of the data processing pipeline is written in Python. For training the model, only patients whose culture results tested positive for sepsis were considered as Sepsis positive and the rest of the patients considered as negative Sepsis cases. All patient PII information from was removed and data anonymized before use in model training. This study and use of data was reviewed and cleared by the Institutional ethics committee.

After data collation and cleansing we had data from 415 Non sepsis patients and from 120 Sepsis positive patients. The admission and discharge timestamp serves as the start and end of the hourly sampling interval for each patient. Patient vitals data as available for each hour was collated from various patient files. While we started with a longer list of features, the features used to train the model were pruned and selected based on availability across all patient data. We are also motivated to build a model with features that are available across most health care institutions



so that our findings and methods are applicable and can be adopted across a wide spectrum of hospitals with varying sophistication and capabilities. Table 1 details the features and their availability. As shown, the hourly-sampled feature data varies across patients.

Table 1. List of features and their availability

Hourly data availability	Features labels
100%	gender, date of birth, admission date, discharge date, Sepsis
82% to 92%	Pulse, Respiration rate, Temperature
20%	SBP, DBP
6%	Blood glucose level

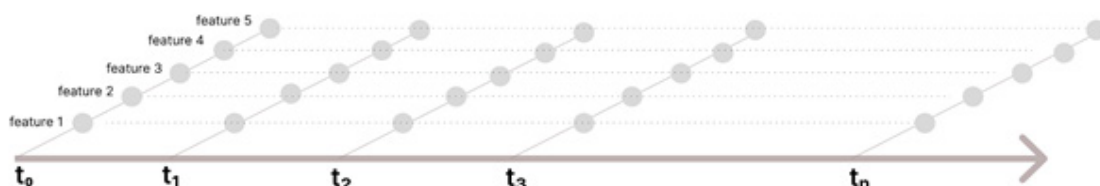
The number of data points available per patient is an important metric to ensure we can train a robust model. The number of records vary between patient for several reasons, some of them:

- Dependent on patient length of stay in the ICU
- Severity of case and frequency of vital recordings during patient stay
- Faithful recording practices of data after each measurement episode

The goal of our project is to develop a ML model that will predict neonatal sepsis several hours earlier than what is done using current hospital protocols. We start with the observation that the data is temporal and feature values collected for every hourly time period need to be interpreted with respect to the previous values. After analysis with several other machine learning model approaches we choose the LSTM (Long Short Term Memory) model. The LSTM approach 2 is well suited for time sequence data and has the ability to "remember" past knowledge. As shown in Figure 1, features samples at progressive timestamps are fed to the LSTM model as shown in timesteps t_0 , t_1 , t_2 , $t(n-1)$ up to t_n . During the learning stage the target label Sepsis (True/False) at each time period is provided as part of model training. After training, the model is used to predict Sepsis at a future timestamp. Inputs to the LSTM model can best be described as having a 3-D form. The three axes of the model inputs are time steps, features and batch size.

Our implementation uses Python and TensorFlow 3 library for data manipulation and model implementation

Figure 1: Overview of the LSTM model training process using feature samples at progressive timestamps



and ML Flow for documenting the details of each run. The data was split into training (70%), validation (20%) and test (10%). The model is composed of a LSTM layer with 12 units followed by a dense layer that converts the LSTM layer's outputs to model predictions. To predict several time steps ahead, we use a RNN multi-step model to predict individual time steps. Each model's output is fed back into itself at each step so that predictions can be made based on the previous output 4. When training the model, we vary the number of time steps as inputs while having the model predict 12 hours ahead. The optimizer used is "Adam" and the loss function is Mean Absolute Error.

Results

We have the model predict results for two scenarios. We train the model with 12 and 24 time steps and have the model predict Sepsis 12 hours into the future. Figure 2 graphs Receiver Operating Characteristic (ROC) curve, which shows the performance of the model. As shown in the graph the model does better with the 24 hour time interval as inputs. The model metrics are tabulated in Table 2.

Figure 2: Receiver Operating Characteristic (ROC) curve comparing the performance of the two models

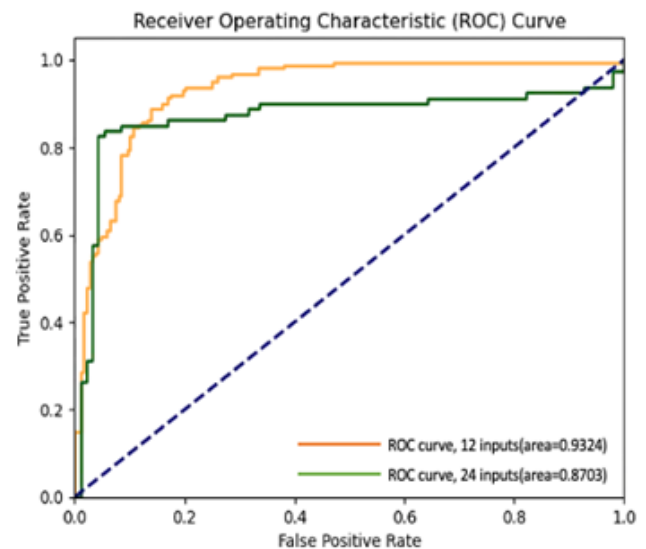


Table 2. Model performance for Sepsis prediction 12 hours ahead

	Input time steps = 12 hrs	Input time step = 24 hrs
accuracy	0.8285	0.8914
precision	0.7972	0.8584
recall	0.9153	0.9578
AUC	0.9324	0.8703

Discussion

In the case of sepsis prediction, where data is imbalanced, precision and recall help evaluate the model rather than accuracy 5,6. Quoting from the references, "Recall, also known as the sensitivity or True Positive Rate (TPR), denotes the rate of positive samples correctly classified, and is calculated as the ratio between correctly classified positive samples and all samples assigned to the positive class". Given that untreated sepsis can lead to death, we argue for choosing an operating threshold that maximizes declaring cases as Sepsis positives at the expense of False Positives cases. Our AUC results compare favorably with the studies in reference 7 .

We envisage a physician using our sepsis machine learning model as follows. After a neonate has been admitted in the hospital, their vital and features required for Sepsis valuation by the model are monitored hourly. At the end of 12 hours (or 24 hours, based on the model chosen) the recorded hourly feature values are input to the model. The model gives a prediction of Sepsis being True or False 12 hours into the future from the time of the last hourly input value. This helps the treating neonatologist to differentiate true sepsis vs sepsis mimicking conditions helping him/her to focus on other causes that could make a difference to the babies instead of assuming the diagnosis of sepsis till it is too late.

Our motivation in building this model is to maximize the usefulness and usability of the model across health care providers with varying sophistication. While we started with a longer list of possible features, based on ease of availability and cost of various tests and measurements, we pruned the feature selection to be ones that are easily available in an ICU setting. Non availability of data at various time intervals is a reality, hence a robust fill in strategy is must. The training of the model was done offline, we have built the inputs with the aim of using it in a hospital setting where data

can be fed in through integration with a EMR system or values of features for the time samples can be input manually. Our next step would be to test the solution on the field and confirm the accuracy and efficacy of the model. Ongoing model training and evaluation needs to be part of a production system.

The model results are promising in identifying babies with sepsis early on. The clinical challenge is to identify true sepsis from metabolic causes or other causes, where symptoms are similar to sepsis but the culture turns out negative. Using this model's prediction, physicians can take measures to control and treat sepsis early on, better manage their babies and potentially save lives.

Conclusions

This LSTM-based model shows strong potential in predicting neonatal sepsis up to 12 hours in advance using routinely available ICU data. Its design prioritizes real-world applicability across varied healthcare settings, enabling timely intervention and improved outcomes. Further clinical validation will support its integration into neonatal care.

Compliance with Ethical Standards

Funding: None

Conflict of Interest: None

References:

1. Narayanaswamy L, Garg D, Narra B, et al. Machine Learning Algorithmic and System Level Considerations for Early Prediction of Sepsis. Computing in Cardiology conference proceedings. 2019; 46. <https://www.cinc.org/archives/2019/>
2. Long short term memory model [internet]. Available from: https://en.wikipedia.org/wiki/Long_short-term_memory
3. Tensor flow documentation [internet]. Available from: <https://www.tensorflow.org/>
4. A Graves. Generating Sequences With Recurrent Neural Networks. arXiv. 2013: 1308.0850. [arxiv.org](https://arxiv.org/abs/1308.0850)
5. Hicks S, Strümke I, Thambawita V, et al. On evaluation metrics for medical applications of artificial intelligence. Scientific Reports. 2022. doi: 10.1038/s41598-022-09954-8
6. https://en.wikipedia.org/wiki/Precision_and_recall
7. Deng H-F, Sun M-W, Wang Y, et al. Evaluating machine learning models for sepsis prediction: A systematic review of methodologies. iScience. 2022;25:103651. doi: 10.1016/j.isci.2021.103651