



OPEN

Development of artificial neural networks for early prediction of intestinal perforation in preterm infants

Joonhyuk Son^{1,5}, Daehyun Kim^{2,5}, Jae Yoon Na^{3,5}, Donggoo Jung², Ja-Hye Ahn³, Tae Hyun Kim⁴✉ & Hyun-Kyung Park³✉

Intestinal perforation (IP) in preterm infants is a life-threatening condition that may result in serious complications and increased mortality. Early Prediction of IP in infants is important, but challenging due to its multifactorial and complex nature of the disease. Thus, there are no reliable tools to predict IP in infants. In this study, we developed new machine learning (ML) models for predicting IP in very low birth weight (VLBW) infants and compared their performance to that of classic ML methods. We developed artificial neural networks (ANNs) using VLBW infant data from a nationwide cohort and prospective web-based registry. The new ANN models, which outperformed all other classic ML methods, showed an area under the receiver operating characteristic curve (AUROC) of 0.8832 for predicting IP associated with necrotizing enterocolitis (NEC-IP) and 0.8797 for spontaneous IP (SIP). We tested these algorithms using patient data from our institution, which were not included in the training dataset, and obtained an AUROC of 1.0000 for NEC-IP and 0.9364 for SIP. NEC-IP and SIP in VLBW infants can be predicted at an excellent performance level with these newly developed ML models. <https://github.com/kdhRick2222/Early-Prediction-of-Intestinal-Perforation-in-Preterm-Infants>.

Intestinal perforation (IP) in preterm infants is a life-threatening condition that requires urgent surgical intervention. Delayed diagnosis of IP can lead to serious complications and even death in certain patients. There are two major causes of IP in premature infants: necrotizing enterocolitis (NEC) and spontaneous intestinal perforation (SIP)^{1–5}. NEC and SIP are acquired diseases, which means that IP occurs due to the actions of various risk factors. IP associated with NEC (NEC-IP) is an advanced stage of NEC in which NEC has progressed to a severe state. On the other hand, SIP can occur unexpectedly without preceding symptoms or signs⁶. Many previous studies have investigated the risk factors and pathogenesis of these two types of IP. According to the literature, possible risk factors include prematurity, low birth weight, perinatal use of medications (ibuprofen, steroids, indomethacin, and inotropic agents), maternal chorioamnionitis, patent ductus arteriosus (PDA), and high-grade intraventricular hemorrhage (IVH grade III or IV)^{5,7–22}. Despite these efforts, due to the nature of retrospective studies with a small number of patients, there is thus far no consensus concerning risk factors. Furthermore, the pathogenesis of IP in premature infants is still not clearly understood; as a result, neonatologists and pediatric surgeons have difficulty predicting and preventing the occurrence of NEC-IP and SIP^{23–26}.

Recently, artificial intelligence (AI) technologies have been used to solve various medical problems, including diagnosing diseases, choosing treatments, and predicting the risks and progression of diseases^{27–30}. Machine learning (ML) algorithms, specifically artificial neural networks (ANNs), allow more accurate estimation of complex patterns through learning from training data and are able to predict which patients will develop a disease as an early and supportive diagnostic tool for clinicians. To date, only a few studies have used ML algorithms to predict IP in premature infants. Irls et al.³¹ reported an ML model predicting NEC-IP, and Lure et al.³² reported an ML model to differentiate NEC-IP and SIP prior to surgical interventions. Lin³³ et al. reported an

¹Department of Pediatric Surgery, Hanyang University College of Medicine, 222 Wangsimni-ro, Seongdong-gu, Seoul 04763, Korea. ²Department of Artificial Intelligence, Hanyang University, 222 Wangsimni-ro, Seongdong-gu, Seoul 04763, Korea. ³Department of Pediatrics, Hanyang University College of Medicine, 222 Wangsimni-ro, Seongdong-gu, Seoul 04763, Korea. ⁴Department of Computer Science, Hanyang University, 222 Wangsimni-ro, Seongdong-gu, Seoul 04763, Korea. ⁵These authors contributed equally: Joonhyuk Son, Daehyun Kim and Jae Yoon Na. ✉email: taehyunkim@hanyang.ac.kr; neopark@hanyang.ac.kr

	A	B	C	p-value	
	No perforation (n = 11,826)	NEC-IP (n = 521)	SIP (n = 208)	A vs. B	A vs. C
Gestational age (weeks, mean \pm SD)	28.58 \pm 3.02	25.84 \pm 2.41	26.25 \pm 2.23	<0.001	<0.001
Birth weight (g, mean \pm SD)	1092.38 \pm 284.05	820.03 \pm 249.41	858.38 \pm 248.15	<0.001	<0.001
Sex-male—n (%)	5919 (50.1)	285 (54.7)	130 (62.5)	0.038	<0.001
Maternal chorioamnionitis—n (%)	3426 (29.0)	162 (31.1)	64 (30.8)	0.296	0.571
PROM—n (%)	4073 (34.4)	196 (37.6)	80 (38.5)	0.135	0.227
Antenatal steroid use—n (%)	9341 (79.0)	414 (79.5)	169 (81.3)	0.794	0.427
Resuscitation at delivery—n (%)	10,472 (88.6)	489 (93.9)	204 (98.1)	<0.001	<0.001
RDS—n (%)	9080 (76.8)	487 (93.5)	198 (95.2)	<0.001	<0.001
Surfactant use—n (%)	9114 (77.1)	489 (93.9)	200 (96.2)	<0.001	<0.001
Steroid use—n (%)	2589 (21.9)	198 (38.0)	91 (43.8)	<0.001	<0.001
Indomethacin use—n (%)	54 (0.5)	8 (1.5)	2 (1.0)	0.001	0.252
Ibuprofen use—n (%)	3539 (29.9)	232 (44.5)	83 (39.9)	<0.001	0.002
Hypotension—n (%)	2838 (24.0)	325 (62.4)	130 (62.5)	<0.001	<0.001
Inotropic use—n (%)	548 (4.6)	59 (11.3)	26 (12.5)	<0.001	<0.001
IVH grade 3,4—n (%)	934 (7.9)	146 (28.0)	48 (23.1)	<0.001	<0.001
Sepsis—n (%)	2224 (18.8)	266 (51.1)	91 (43.8)	<0.001	<0.001
PDA on medication—n (%)	3615 (30.6)	241 (46.3)	87 (41.8)	<0.001	<0.001
PDA ligation—n (%)	1145 (9.7)	142 (27.3)	48 (23.1)	<0.001	<0.001

Table 1. Clinical characteristics of the patients. *NEC-IP* intestinal perforation associated with necrotizing enterocolitis, *SIP* spontaneous intestinal perforation, *SD* standard deviation, *PROM* premature rupture of membranes, *RDS* respiratory distress syndrome, *IVH* intraventricular hemorrhage, *PDA* patent ductus arteriosus.

ML technique for individualized NEC risk scores using intestinal microbiota data. However, despite these efforts, studies concerning ML technologies in neonatal IP are still lacking, and there are no definitive tools to predict NEC-IP and SIP in advance. Therefore, our study aimed to develop our own ML models for predicting IP in very low birth weight (VLBW) infants using a nationwide cohort dataset and to compare the performance of the new models with that of classic ML models. Furthermore, to ensure that our novel predictive ML models were valid in real neonatal intensive care (NICU) settings, we tested new ML models with individual cases from our institution and assessed whether these new models achieved high predictive accuracy.

Results

Baseline characteristics of the patients. The baseline characteristics of the infants are shown in Table 1. Of the 12,555 patients, 521 patients had NEC-IP (4.1%), and 208 patients had SIP (1.7%). Three groups of VLBW infants were compared for analysis: (A) control group without NEC-IP or SIP, (B) patients with NEC-IP, and (C) patients with SIP. Factors such as low gestational age, low birth weight, need for resuscitation at delivery, respiratory distress syndrome (RDS) and administration of surfactant, use of certain medications (steroids, ibuprofen, and inotropes), hypotension, IVH grade III/IV, sepsis, and PDA on treatment were significantly associated with IP, including both NEC-IP and SIP.

Proposed machine learning algorithms. In the work of Irls et al.³¹, ML-based approaches were proposed to predict NEC values, and our neural approaches produced promising NEC prediction results using a dataset that included 852 positive cases. However, it is more difficult to collect data for NEC-IP and SIP than for NEC; thus, the number of datasets, especially for positive cases, is very limited (521 and 208 positive cases for NEC-IP and SIP, respectively). To solve this problem, we utilized relevant information from a neural network trained for NEC prediction to elevate the diagnostic accuracy of SIP and NEC-IP.

In this study, we introduced several deep neural networks for predicting NEC, NEC-IP, and SIP in infants by taking a 54-dimensional input vector as the network input, and the output of each network was a single value for binary classification problems. Specifically, we first developed our baseline neural network (Model 1) based on the conventional multilayer perceptron (MLP) architecture and then trained Model 1 to predict NEC, NEC-IP, and SIP separately. For ANNs, it is well known that simply adding layers can lead to performance improvement by enforcing more nonlinearity; however, it causes the overfitting problem when the training dataset is insufficient. Therefore, we stacked the layers more deeply than the conventional models in diagnosis, as introduced in a previously published study³¹, but determined the hyperparameters (e.g., the number of channels) carefully and added more advanced techniques, such as batch normalization³⁴ and drop-out³⁵, to avoid the overfitting problem and facilitate stable training. Specifically, Model 1 is a binary classifier that is composed of 5 hidden layers; each layer is formed from a block (Fig. 1a) arranged with an activation function, batch normalization and dropout. It takes a 54-dimensional vector as an input and, after the application of all layers, renders feature vectors of dimension. Model 1 is a typical neural approach and can produce promising results when a large number of

datasets (e.g., NEC) are available. However, the outcomes from Model 1 can be vulnerable in real-world scenarios where the training dataset is insufficient.

Therefore, we attempted to further improve our baseline Model 1 to predict NEC-IP and SIP more accurately since the lack of data problem for NEC-IP and SIP is more serious than that for NEC. To alleviate this problem, we developed a new approach that transferred information from the network to predict NEC values to help estimate NEC-IP and SIP. Note that transfer learning in the field of deep learning is one of the most widely used approaches to solve lack of data problems^{36–40}. Therefore, we present additional models (Model 2 and Model 3) that can exploit information achieved from the NEC dataset to predict SIP and NEC-IP.

Specifically, Model 2 is composed of two different MLPs. One branch predicts NEC, and the other branch predicts either SIP or NEC-IP (Fig. 1b). Notably, at the 4th layer of the MLP branch for NEC-IP and SIP in Model 2, feature vectors from the third layer of the network for NEC are fed as an additional input. By concatenating the feature vectors from NEC, we can utilize the information for NEC in predicting NEC-IP/SIP.

Unlike that of Model 2, the network architecture of Model 3 is the same as that of Model 1. We employed conventional transfer learning to utilize information from NEC to estimate NEC-IP/SIP and fine-tuned the pretrained Model 1 to the specific NEC-IP/SIP datasets (Fig. 1c).

Comparison of performance between classic ML models and proposed ANN models. We provide prediction results in Table 2 to compare traditional ML models with our neural approaches. We observed that the proposed neural approach (Model 1) outperformed traditional learning-based methods in terms of area under the receiver operating characteristic curve (AUROC) scores for all cases (NEC, NEC-IP, and SIP).

Moreover, our extended networks (i.e., Model 2 and Model 3), which we designed to mitigate the lack of data problem for NEC-IP (521 positive cases) and SIP (208 positive cases), showed improved results in predicting NEC-IP and SIP over the baseline network (Model 1).

In particular, compared to Model 1, which was trained on the NEC dataset, our proposed methods for NEC-IP and SIP (Model 2 and Model 3) exhibited improvements. Notably, Model 2 directly utilizes features distilled from Model 1, Model 3 fine-tunes Model 1 for the prediction of either NEC-IP or SIP, and Model 2 outperforms Model 3 in terms of AUROC, as shown in Table 3 and Fig. 2. The performance metrics of these models using balanced validation dataset are described in Supplementary Tables 1 and 2. This performance improvement achieved in Model 2 and Model 3 indicates that information extracted to predict NEC can also be used to predict NEC-IP and SIP more accurately. Moreover, the proposed direct feature distillation of Model 2 rather than the conventional fine-tuning approach (i.e., Model 3) can be a recommendable option for addressing problems with limited data.

Application of the new ANN models in a real clinical environment. To assess the feasibility of the algorithm in a real clinical environment, we tested our newly developed algorithms using the patient data from our institution, which were not included in the training dataset. A total of 57 VLBW infants who were born at our hospital between 2019 and 2020 were included in this test analysis. Among the three ANN models, Model 2 achieved the highest AUROC scores: 1.0000 for the prediction of NEC-IP and 0.9364 for the prediction of SIP (Table 4 and Fig. 3).

Discussion

Our novel ML algorithms predicted NEC-IP and SIP in VLBW infants with favorable AUROC scores, outperforming all other classic ML algorithms. One of our algorithms exhibited an AUROC score of 1.0000 for predicting NEC-IP and 0.9364 for predicting SIP in real clinical settings. Our study shows the integration of a vast nationwide dataset with ML, and the resulting model can be used to predict the possibility of specific medical conditions in patients who may not perfectly represent the signs and symptoms of the disease.

Among preterm infants in the NICU, various medical problems exist at the same time, and multidisciplinary collaboration is required to make medical decisions for these patients. IP is one of the most devastating medical conditions that occurs in the NICU. Early diagnosis, swift judgment, and prompt surgical intervention are required to prevent severe complications and poor outcomes⁴¹. However, as we explained earlier, predicting the occurrence of IP is difficult. By running this algorithm, it is possible to analyze every preterm infant who is admitted to the NICU and predict each patient's likelihood of developing NEC-IP and SIP. Thus, early predictions of these serious medical conditions could provide clinicians with a much more stable management environment, enabling them to make better treatment decisions.

In recent years, AI and big data have been increasingly integrated into medicine because it is difficult for the unaided human clinician to acquire all the latest published knowledge, as is required by modern evidence-based medicine^{30,42}. AI is an important resource for medical research in that it can efficiently process large amounts of data. Additionally, AI can produce consistent and unbiased results without fatigue. Several studies in healthcare research have reported sufficient or even better risk prediction by AI methods compared to existing models^{43–47}. Especially these days, in solving difficult problems with big data which have complex distribution, neural approaches exhibit excellent performance. Since neural networks can have non-linearity from adding layers and fine-tune parameters by transfer learning^{36,40}, they can get the upper hand in complex tasks. To sum up, ANN is an appropriate model for IP prediction, as it can handle imbalanced big data efficiently. However, even if the overall cohort is large, diseases with low prevalence always suffer from a lack of data. Due to the character of ML, it can produce excellent results only with a large amount of training data. Thus, applying ML to diseases with low prevalence remains a challenge^{48,49}. To overcome the data imbalance problem, several studies have applied data processing techniques such as oversampling⁵⁰ and undersampling⁵¹ as we did in our ANN models.

To further improve the performance of our models, we modified them by adding another branch or pre-training it based on an algorithm that predicts other relevant disease with higher prevalence in order to help

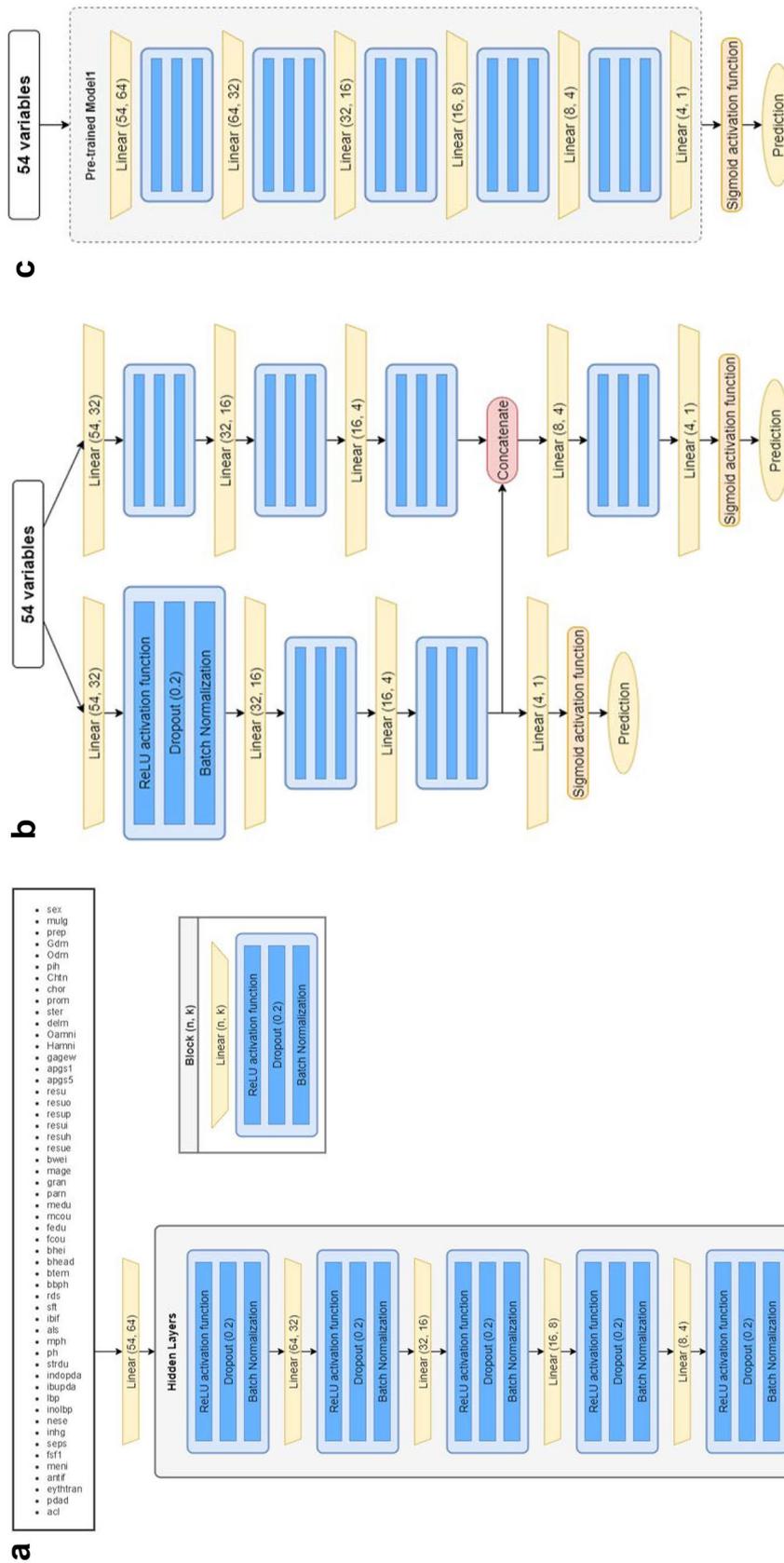


Figure 1. Architecture of machine learning models. (a) Model 1 is a baseline neural network based on the conventional multilayer perceptron (MLP) architecture. (b) Model 2 is composed of two different MLPs. One branch predicts NEC, and the other branch predicts either NEC-IP or SIP. The feature vectors from the 3rd layer of the network for NEC are concatenated with the 4th layer of the MLP branch for NEC-IP and SIP. (c) Model 3 has the same network architecture as that of Model 1. Pretrained Model 1 for NEC was further fine-tuned to estimate NEC-IP/SIP.

AUROC	NEC	NEC-IP	SIP
Linear SVM	0.7632	0.8618	0.8162
Radial SVM	0.7567	0.8195	0.7481
Logistic regression	0.7641	0.8644	0.8044
K-NN	0.6229	0.6337	0.5759
Decision tree	0.5145	0.5377	0.5004
XGBoost	0.6758	0.7748	0.7452
LightGBM	0.7087	0.7758	0.7477
Random forest	0.7495	0.8051	0.7687
MLP (model 1)	0.8128	0.8665	0.8498

Table 2. Model performance of classic ML models for predicting NEC, NEC-IP, and SIP. *ML* machine learning, *NEC* necrotizing enterocolitis, *NEC-IP* intestinal perforation associated with necrotizing enterocolitis, *SIP* spontaneous intestinal perforation, *SVM* support vector machine, *K-NN* k-nearest neighbor, *XGBoost* extreme gradient boosting, *GBM* gradient boosting machine learning, *LightGBM* light gradient boosting machine learning, *MLP* multilayer perceptron.

the model predict the target disease more accurately. As a result of these adjustments, the modified algorithms (Model 2 and 3) achieved better performance than the original model. According to previous studies, NEC-IP and SIP are regarded as separate disease entities, and the pathogenesis of SIP does not appear to correlate with that of NEC^{5,6,52–55}. Notably, however, training of NEC prediction improved not only the accuracy of NEC-IP prediction but also the accuracy of SIP prediction. These results show that pre-ML training with more prevalent medical conditions can help AI predict the occurrence of the target disease more accurately. Our study also highlights that it is necessary to customize the algorithm for each disease to apply an ML model in real clinical settings, especially if the disease is rare.

Although our study showed a favorable outcome, it had its share of limitations. First, a limited number of factors are included because only data collected from the Korean Neonatal Network (KNN) were used. It is expected that the collection of further IP-related data, such as clinical symptoms, vital signs, and radiologic findings, will enable the model to produce better results. In addition, the limitations of AI studies, such as representation, homogeneity, and accuracy, were observed in this study. Another limitation is that it is difficult to determine how AI methods generate results due to the nature of self-extracted data from large datasets^{49,56,57}.

In conclusion, we developed our own ANN models to predict IP early in VLBW infants, and these new models achieved higher accuracy than classic ML algorithms. To our knowledge, this is the first study to develop an ML model to predict both NEC-IP and SIP using nationwide VLBW infant data. In addition, the newly proposed ANN models showed excellent performance within real NICU clinical settings. When more clinical data, such as vital signs, radiologic findings, biomarkers, and laboratory results, are gathered, we believe that a more accurate ML model will be developed, thereby achieving early prediction of these serious medical conditions and better clinical outcomes for VLBW infants.

Methods

Data collection. We derived data from infants registered in the KNN, a nationwide prospective cohort registry of VLBW infants⁵⁸. Their clinical data were collected from 74 participating NICUs across the country and analyzed retrospectively for this study. Prior to participation in the KNN registry, informed consent was obtained from the parents of each infant, and all methods were carried out following relevant guidelines. This study was approved by the Hanyang University Institutional Review Board (IRB No. 2013-06-025-043).

The cohort comprised 12,555 VLBW infants born between January 5, 2013, and December 31, 2018, weighing less than 1500×g.

Disease definitions. NEC was defined according to Bell's modified staging grade \geq II. NEC-IP was diagnosed when patients with NEC underwent any kind of abdominal surgical intervention (peritoneal drainage or laparotomy). SIP was defined when the patients underwent surgical intervention due to IP and the surgeon found no predisposing causes, such as NEC, intestinal atresia, or meconium plug. The full list of 54 variables used in ML analysis is shown in Supplementary Table 3.

Comparisons of baseline characteristics. A total of 54 variables, including various maternal and perinatal factors, were collected for ML. Among them, 18 clinical factors that were proposed as possible risk factors for either NEC-IP or SIP in previous studies were analyzed using conventional statistical methods. Student's t-test was performed to analyze the continuous variables, and the chi-squared test was used to analyze categorical variables. Statistical significance was set at $P < 0.05$. The Statistical Package for the Social Sciences version 22.0 for Windows software program (IBM Corp., Armonk, NY, USA) was used in all statistical analyses.

Data preprocessing. Our dataset was composed of 12,555 infants in total; we divided them into training and evaluation datasets (Table 5). Moreover, to facilitate the network training procedure, a data preprocessing

	Model 1	Model 2	Model 3
AUROC			
NEC	0.8128	–	–
NEC-IP	0.8665	0.8832	0.8692
SIP	0.8498	0.8797	0.8633
F1-score^a			
NEC	0.7701	–	–
NEC-IP	0.7181	0.8093	0.7273
SIP	0.8059	0.8204	0.8041

Table 3. Performance of the proposed ANN models in predicting NEC, NEC-IP, and SIP. ANN artificial neural network, NEC necrotizing enterocolitis, NEC-IP intestinal perforation associated with necrotizing enterocolitis, SIP spontaneous intestinal perforation, AUROC area under the receiver operating characteristic curve. ^aScores were found with balanced validation dataset. Positive cases were oversampled.

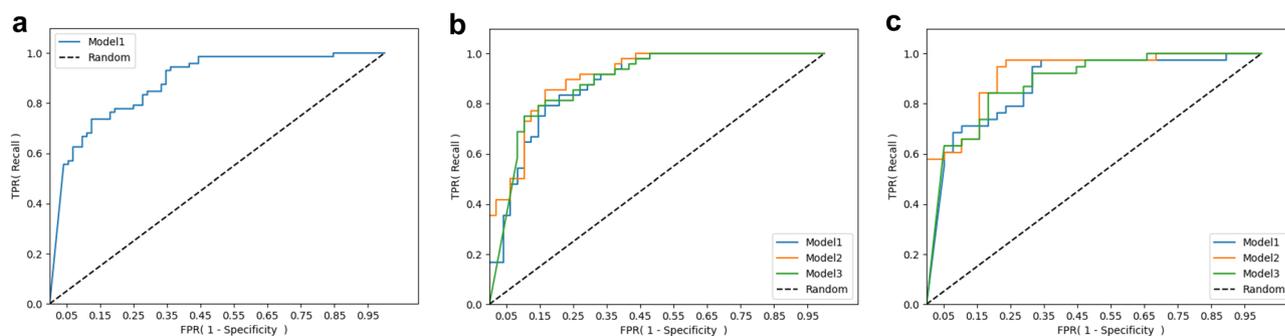


Figure 2. Receiver operating characteristic curves of proposed ML models for (a) NEC prediction, (b) NEC-IP prediction, and (c) SIP prediction.

	Model 1	Model 2	Model 3
AUROC			
NEC	0.6745	–	–
NEC-IP	1.0000	1.0000	0.8704
SIP	0.9000	0.9364	0.8818
F1-score^a			
NEC	0.6903	–	–
NEC-IP	0.8571	0.9076	0.8710
SIP	0.7241	0.7179	0.7925

Table 4. Test results for 57 cases within a real NICU environment. NICU neonatal intensive care unit, AUROC area under the receiver operating characteristic curve, NEC necrotizing enterocolitis, NEC-IP intestinal perforation associated with necrotizing enterocolitis, SIP spontaneous intestinal perforation. ^aScores were found with balanced validation dataset. Positive cases were oversampled.

step was applied (Fig. 4). First, to solve the missing data problem in the given dataset, we imputed the missing (null) values with plausible values. Specifically, input values were categorized into ordinal, continuous and categorical types. In the case of ordinal inputs, we imputed the null values with the mode (most frequently occurring) values, and in the case of continuous inputs, we imputed the null values with the mean values. Finally, missing values of categorical inputs were replaced with the median values. After recovering the data, we mapped the dataset between 0 and 1 by using min–max normalization. Finally, to mitigate the data imbalance problem (e.g., 11,703 negative and 852 positive cases for NEC in Table 5), which is a common intrinsic feature of disease datasets, we oversampled the smaller category (positive cases) and undersampled the bigger category (negative cases), as suggested in previous studies^{59,60}.

Training. We used the binary cross entropy (BCE) loss to train Model 1, and Model 1 was separately trained to predict NEC, NEC-IP, and SIP. For Model 3, pretrained Model 1 for NEC was further fine-tuned, and the separately updated parameters using the BCE loss were used to predict NEC-IP and SIP.

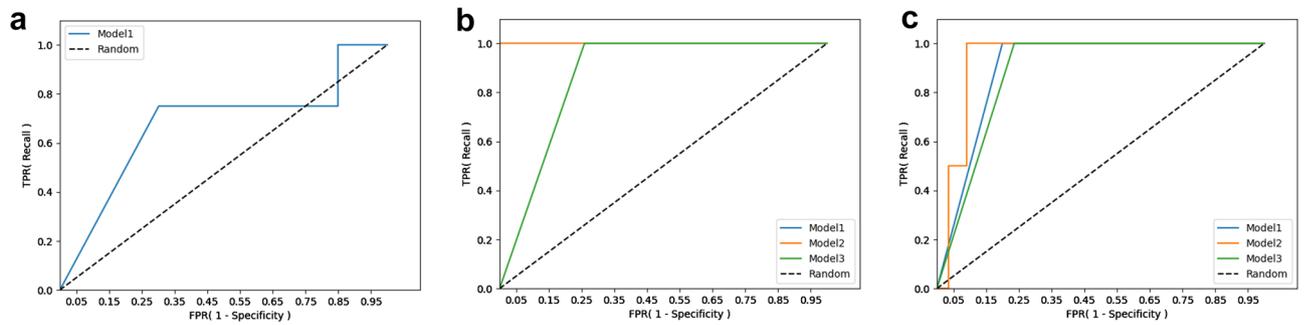


Figure 3. Receiver operating characteristic curves of proposed ML models from 57 test cases. **(a)** NEC prediction, **(b)** NEC-IP prediction, and **(c)** SIP prediction.

	Training dataset		Evaluation dataset		Total
	Negative	Positive	Negative	Positive	
NEC	10,000	770	1703	82	12,555
NEC-IP		430	2035	91	
SIP		170	2348	38	

Table 5. Size of training and evaluation datasets. *NEC* necrotizing enterocolitis, *NEC-IP* intestinal perforation associated with necrotizing enterocolitis, *SIP* spontaneous intestinal perforation.

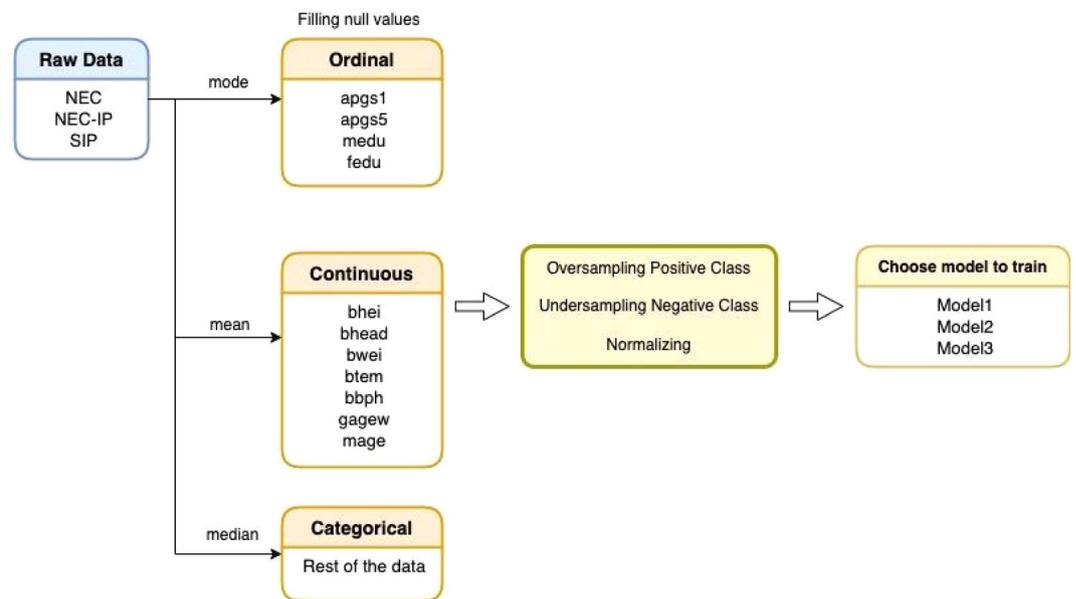


Figure 4. Flowchart of data processing. Input values were categorized into ordinal, continuous and categorical types. To solve the data imbalance problem, oversampling and undersampling technique were applied. Then, the data were normalized.

Unlike Model 1 and Model 3, Model 2 was trained in a two-stage manner. In the first stage, the left branch of the network (Fig. 1b) was trained for NEC with the BCE loss. Then, the BCE loss as well as the BCE for NEC in the left branch were used to jointly optimize the network for either NEC-IP or SIP. In both training steps, oversampling and undersampling technique were used to alleviate the data imbalance problem. We employed the two-stage training scheme rather than using the one-shot joint training approach, as we could retain more stability during the training from the two-stage approach.

To train the three proposed models (i.e., Model 1, Model 2, Model 3), the Adam optimizer⁶¹ with a learning rate of 0.0001 was used for loss minimization. We used a dropout rate of 0.2³⁵ and a batch size of 128³⁴. Instead of fixing the number of iterations, we stopped the training using the early stopping technique to avoid the overfitting

problem. When the loss increased seven times in a row, we stopped the training (i.e., early stop) and took the parameters just prior to the loss increase. The models were implemented in PyTorch with Python⁶², and the Scikit-learn library⁶³ was used to evaluate the results (e.g., AUROC, F1-Score, and ROC curve).

Data availability

According to the Korean Neonatal Network (KNN) Publication Ethics Policy, all information about patients is confidential. The information contained in the data must be protected as confidential, and only available to individuals who have access for the permitted research activity.

Received: 16 December 2021; Accepted: 7 July 2022

Published online: 15 July 2022

References

- Hunter, C. J., Chokshi, N. & Ford, H. R. Evidence vs experience in the surgical management of necrotizing enterocolitis and focal intestinal perforation. *J. Perinatol.* **28**(Suppl 1), S14–17 (2008).
- Karila, K., Anttila, A., Iber, T., Pakarinen, M. & Koivusalo, A. Outcomes of surgery for necrotizing enterocolitis and spontaneous intestinal perforation in Finland during 1986–2014. *J. Pediatr. Surg.* **53**(10), 1928–1932 (2018).
- Vongbhavit, K. & Underwood, M. A. Intestinal perforation in the premature infant. *J. Neonatal Perinatal Med.* **10**(3), 281–289 (2017).
- Zozaya, C. *et al.* Neurodevelopmental and growth outcomes of extremely preterm infants with necrotizing enterocolitis or spontaneous intestinal perforation. *J. Pediatr. Surg.* **56**(2), 309–316 (2021).
- Shah, J. *et al.* Intestinal perforation in very preterm neonates: Risk factors and outcomes. *J. Perinatol.* **35**(8), 595–600 (2015).
- Pumberger, W., Mayr, M., Kohlhauser, C. & Weninger, M. Spontaneous localized intestinal perforation in very-low-birth-weight infants: A distinct clinical entity different from necrotizing enterocolitis. *J. Am. Coll. Surg.* **195**(6), 796–803 (2002).
- Ragouilliaux, C. J., Keeney, S. E., Hawkins, H. K. & Rowen, J. L. Maternal factors in extremely low birth weight infants who develop spontaneous intestinal perforation. *Pediatrics* **120**(6), e1458–e1464 (2007).
- Rayyan, M. *et al.* Risk factors for spontaneous localized intestinal perforation in the preterm infant. *J. Matern. Fetal Neonatal Med.* **31**(19), 2617–2623 (2018).
- Rose, A. T. & Patel, R. M. A critical analysis of risk factors for necrotizing enterocolitis. *Semin. Fetal Neonat. M.* **23**(6), 374–379 (2018).
- Samuels, N., van de Graaf, R. A., de Jonge, R. C. J., Reiss, I. K. M. & Vermeulen, M. J. Risk factors for necrotizing enterocolitis in neonates: A systematic review of prognostic studies. *BMC Pediatr.* **17**(1), 105 (2017).
- Youn, Y. A., Kim, E. K. & Kim, S. Y. Necrotizing enterocolitis among very-low-birth-weight infants in Korea. *J. Korean Med. Sci.* **30**(Suppl 1), S75–80 (2015).
- Arnaudovic, T. I. *et al.* Antenatal risk factors associated with spontaneous intestinal perforation in preterm infants receiving post-natal indomethacin. *J. Pediatr.* **232**(59–64), e51 (2021).
- Kandragu, H. *et al.* Association of co-exposure of antenatal steroid and prophylactic indomethacin with spontaneous intestinal perforation. *J. Pediatr.* **235**(34–41), e31 (2021).
- Stavel, M. *et al.* Effect of prophylactic indomethacin administration and early feeding on spontaneous intestinal perforation in extremely low-birth-weight infants. *J. Perinatol.* **37**(2), 188–193 (2017).
- Yilmaz, Y. *et al.* Preeclampsia is an independent risk factor for spontaneous intestinal perforation in very preterm infants. *J. Matern. Fetal Neonatal Med.* **27**(12), 1248–1251 (2014).
- Houben, C. H. *et al.* Spontaneous intestinal perforation: The long-term outcome. *Eur. J. Pediatr. Surg.* **27**(4), 346–351 (2017).
- Prasad, U., Mohnani, A. & Hussain, N. Spontaneous intestinal perforation associated with premature twin infants. *J. Neonatal Perinatal Med.* **14**(3), 403–409 (2021).
- Maheshwari, A., Patel, R. M. & Christensen, R. D. Anemia, red blood cell transfusions, and necrotizing enterocolitis. *Semin. Pediatr. Surg.* **27**(1), 47–51 (2018).
- Wang, Z. L. *et al.* Risk factors of necrotizing enterocolitis in neonates with sepsis: A retrospective case-control study. *Int. J. Immunopathol. Pharmacol.* **34**, 2058738420963818 (2020).
- Tatli, M. M. *et al.* Spontaneous intestinal perforation after oral ibuprofen treatment of patent ductus arteriosus in two very-low-birthweight infants. *Acta Paediatr.* **93**(7), 999–1001 (2004).
- Ndour, D. *et al.* Adverse events related to ibuprofen treatment for patent ductus arteriosus in premature neonates. *Arch. Pediatr.* **27**(8), 452–455 (2020).
- El Manouni El Hassani, S. *et al.* Predictive factors for surgical treatment in preterm neonates with necrotizing enterocolitis: A multicenter case-control study. *Eur. J. Pediatr.* **180**(2), 617–625 (2021).
- Eaton, S., Rees, C. M. & Hall, N. J. Current research on the epidemiology, pathogenesis, and management of necrotizing enterocolitis. *Neonatology* **111**(4), 423–430 (2017).
- Neu, J. Necrotizing enterocolitis: The mystery goes on. *Neonatology* **106**(4), 289–295 (2014).
- Frost, B. L., Modi, B. P., Jaksic, T. & Caplan, M. S. New medical and surgical insights into neonatal necrotizing enterocolitis: A review. *JAMA Pediatr.* **171**(1), 83–88 (2017).
- Cho, S. X. *et al.* Characterization of the pathoimmunology of necrotizing enterocolitis reveals novel therapeutic opportunities. *Nat. Commun.* **11**(1), 5794 (2020).
- Krittawong, C. *et al.* Future direction for using artificial intelligence to predict and manage hypertension. *Curr. Hypertens. Rep.* **20**(9), 75 (2018).
- Obermeyer, Z. & Emanuel, E. J. Predicting the future: Big data, machine learning, and clinical medicine. *N. Engl. J. Med.* **375**(13), 1216–1219 (2016).
- Zhang, Z. *et al.* Derivation and validation of an ensemble model for the prediction of agitation in mechanically ventilated patients maintained under light sedation. *Crit. Care Med.* **49**(3), e279–e290 (2021).
- Na, J. Y. *et al.* Artificial intelligence model comparison for risk factor analysis of patent ductus arteriosus in nationwide very low birth weight infants cohort. *Sci. Rep.* **11**(1), 1–10 (2021).
- Irls, C. *et al.* Estimation of neonatal intestinal perforation associated with necrotizing enterocolitis by machine learning reveals new key factors. *Int. J. Environ. Res. Public Health.* **15**(11), 2509 (2018).
- Lure, A. C. *et al.* Using machine learning analysis to assist in differentiating between necrotizing enterocolitis and spontaneous intestinal perforation: A novel predictive analytic tool. *J. Pediatr. Surg.* **56**(10), 1703–1710 (2020).
- Lin, Y. C., Sallab-Aouissi, A. & Hooven, T. A. Interpretable prediction of necrotizing enterocolitis from machine learning analysis of premature infant stool microbiota. *BMC Bioinform.* **23**(1), 104 (2022).
- Ioffe, S. & Szegedy, C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. in *Proceedings of the 32nd International Conference on Machine Learning*, vol 37, 448–456 (PMLR, 2015).

35. Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I. & Salakhutdinov, R. Dropout: A simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.* **15**, 1929–1958 (2014).
36. Pan, S. J. & Yang, Q. A. A survey on transfer learning. *IEEE Trans. Knowl. Data Eng.* **22**(10), 1345–1359 (2010).
37. Rawat, W. & Wang, Z. H. Deep convolutional neural networks for image classification: a comprehensive review. *Neural Comput.* **29**(9), 2352–2449 (2017).
38. Bengio, Y. Deep learning of representations for unsupervised and transfer learning. in *Proceedings of ICML workshop on unsupervised and transfer learning*. (JMLR Workshop and Conference Proceedings, 2012).
39. Tan, C. Q. *et al.* A survey on deep transfer learning. *Artif. Neural Netw. Mach. Learn.* **11141**, 270–279 (2018).
40. Zhuang, F. Z. *et al.* A comprehensive survey on transfer learning. *Proc. IEEE* **109**(1), 43–76 (2021).
41. Bethell, G. S., Knight, M. & Hall, N. J. Surgical necrotizing enterocolitis: Association between surgical indication, timing, and outcomes. *J. Pediatr. Surg.* **56**(10), 1785–1790 (2021).
42. Barry-Jester, A. M., Casselman, B. & Goldstein, D. The new science of sentencing. *Marshall Project.* **4**, 2015 (2015).
43. Alaa, A. M., Bolton, T., Di Angelantonio, E., Rudd, J. H. F. & van der Schaar, M. Cardiovascular disease risk prediction using automated machine learning: A prospective study of 423,604 UK Biobank participants. *PLoS ONE* **14**(5), e0213653 (2019).
44. Jhee, J. H. *et al.* Prediction model development of late-onset preeclampsia using machine learning-based methods. *PLoS ONE* **14**(8), e0221202 (2019).
45. Lee, Y. *et al.* Machine learning-based prediction of acute kidney injury after nephrectomy in patients with renal cell carcinoma. *Sci. Rep.* **11**(1), 15704 (2021).
46. Safavi, K. C. *et al.* Development and validation of a machine learning model to aid discharge processes for inpatient surgical care. *JAMA Netw. Open.* **2**(12), e1917221 (2019).
47. Weng, S. F., Reys, J., Kai, J., Garibaldi, J. M. & Qureshi, N. Can machine-learning improve cardiovascular risk prediction using routine clinical data?. *PLoS ONE* **12**(4), e0174944 (2017).
48. Murdoch, W. J., Singh, C., Kumbier, K., Abbasi-Asl, R. & Yu, B. J. *Interpretable Machine Learning: Definitions, Methods, and Applications* (Springer, 2019).
49. Rajkomar, A., Dean, J. & Kohane, I. Machine learning in medicine. *N. Engl. J. Med.* **380**(14), 1347–1358 (2019).
50. Goh, K. H. *et al.* Artificial intelligence in sepsis early prediction and diagnosis using unstructured data in healthcare. *Nat. Commun.* **12**(1), 711 (2021).
51. Liu, X. Y., Wu, J. X. & Zhou, Z. H. Exploratory undersampling for class-imbalance learning. *IEEE Trans. Syst. Man Cybern. B* **39**(2), 539–550 (2009).
52. Okuyama, H. *et al.* A comparison of the clinical presentation and outcome of focal intestinal perforation and necrotizing enterocolitis in very-low-birth-weight neonates. *Pediatr. Surg. Int.* **18**(8), 704–706 (2002).
53. Buchheit, J. Q. & Stewart, D. L. Clinical comparison of localized intestinal perforation and necrotizing enterocolitis in neonates. *Pediatrics* **93**(1), 32–36 (1994).
54. Coates, E. W., Karłowicz, M. G., Croitoru, D. P. & Buescher, E. S. Distinctive distribution of pathogens associated with peritonitis in neonates with focal intestinal perforation compared with necrotizing enterocolitis. *Pediatrics* **116**(2), e241–246 (2005).
55. Mintz, A. C. & Applebaum, H. Focal gastrointestinal perforations not associated with necrotizing enterocolitis in very-low-birth-weight neonates. *J. Pediatr. Surg.* **28**(6), 857–860 (1993).
56. Murdoch, W. J., Singh, C., Kumbier, K., Abbasi-Asl, R. & Yu, B. Definitions, methods, and applications in interpretable machine learning. *Proc. Natl. Acad. Sci. USA.* **116**(44), 22071–22080 (2019).
57. Rai, A. Explainable AI: From black box to glass box. *J. Acad. Mark. Sci.* **48**(1), 137–141 (2020).
58. Chang, Y. S., Park, H.-Y. & Park, W. S. The Korean neonatal network: an overview. *J. Korean Med. Sci.* **30**(Suppl 1), S3–S11 (2015).
59. Chawla, N. V., Bowyer, K. W., Hall, L. O. & Kegelmeyer, W. P. SMOTE: Synthetic minority over-sampling technique. *J. Artif. Intell. Res.* **16**, 321–357 (2002).
60. Johnson, J. M. & Khoshgoftaar, T. M. Survey on deep learning with class imbalance. *J. Big Data Ger.* **6**(1), 1–54 (2019).
61. Kingma, D. P. & Ba, J. *Adam: A Method for Stochastic Optimization*. CoRR. <http://arxiv.org/abs/1412.6980> (2015).
62. Vasilev, I., Slater, D., Spacagna, G., Roelants, P. & Zocca, V. *Python Deep Learning: Exploring Deep Learning Techniques and neural Network Architectures with Pytorch, Keras, and TensorFlow* (Packt Publishing Ltd, 2019).
63. Pedregosa, F. *et al.* Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).

Acknowledgements

This research was supported by funds (2019-ER7103-02#) from Research of the Korea National Institute of Health and research funds from Hanyang University MEB (Global Center for Developmental Disorders, HY-20210000002865), an Institute of Information & Communications Technology Planning & Evaluation (IITP) grant funded by the Korean government (MSIT) [No. 2020-0-01373, Artificial Intelligence Graduate School Program (Hanyang University)], and the Bio & Medical Technology Development Program of the National Research Foundation of Korea (NRF) funded by the Korea government (MSIT) [Grant Number NRF-2019M3E5D1A01069363].

Author contributions

H.K.P. and T.H.K. conceptualized, designed, and supervised the study. J.S. and J.H.A. were involved in data collection and initial statistical analysis. T.H.K., D.K., and D.J. performed the computational analyses, development of the algorithms, and interpretation of the data. J.S., D.K. and J.Y.N. were involved in the writing, and editing of the paper. All authors reviewed and approved the revised manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-16273-5>.

Correspondence and requests for materials should be addressed to T.H.K. or H.-K.P.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022