March - April 2020
ISSN: 0193-4120 Page No. 4135 - 4143

An Early Predictive Model of Sepsis with White Blood Cell Calibration

Ju-Hyuck Han[1], Gi-Hun Park[2], Yong-Suk Kim[*3]
[1]Master student, Department of Medical IT Engineering, Konyang University, Daejeon City, Rep. of Korea.
[2]Ph. D. candidate, Department of Medical IT Engineering, Konyang University, Daejeon City, Rep. of Korea.
[*3]Professor, Department of Medical IT Engineering, Konyang University, Daejeon City, Rep. of Korea.
dnfwlq203@gmail.com[1], rlgns11190@naver.com[2], yongsuk@konyang.ac.kr[*3]

Article Info
Volume 83
Page Number: 4135 - 4143
Publication Issue:
March - April 2020

Abstract
In this paper, the possibility of an early prediction for sepsis using deep learning was investigated. The sepsis is a disease that infects microorganisms, resulting in systemic inflammatory reactions such as fever, rapid pulse rate, respiratory increase rate, increase or decrease of white blood cell count. Sepsis also currently incur the highest medical costs of all diseases, affecting about 30 million people a year worldwide. Sepsis, an infectious disease, is essential for early detection because it can lower mortality rates with treatment during the initial three hours of infection and faster antibiotic administration. However, it takes a lot of tests to detect sepsis. White Blood Cells (WBCs) are particularly important in diagnosing sepsis but require blood tests. This acts as an obstacle to early detection by having additional medical expenses and time spent on the examination. Thus, this paper studied a deep learning model that can initially predict sepsis by calibrating the white blood cell count values acting as an important factor in sepsis detection. The data set in this paper utilizes PhysioNet's 'Early Prediction of Sepsis from Clinical Data the PhysioNet Computing in Cardiology Challenge 2019. Data are inpatient data from the intensive care unit released by two hospitals, including biometric signals (1 to 8), body component test results (9 to 34), other information (35 to 40) and annotations (41). This study used nine of the above 41 items and annotations, especially WBC is an important factor in diagnosing sepsis. However, because WBC require blood tests, they are harder to measure than biometric signals. This omitted many values (91 percent). Thus, this paper studied algorithms to calibrate the WBC and, furthermore, predict sepsis early, using the GAN (Generative Adversarial Network) technique. In this paper, the indicators of objectification focus on the ROC (Receiver Operating Characteristics) Curve and AUC (Area Under the Curve). The study for sepsis prediction showed excellent performance by the LSTM (Long Short-Term Memory) method, which is advantageous for sequential pattern learning, because it predicts the patient's condition by time. The LSTM-style study was constructed using approximately 6,000 data and showed a performance of 0.929 AUC. Traditional studies have also used WBC as a factor to predict sepsis, but many of the missing data have not been reflected in the study. However, in this study, missing data values are generated and corrected through a deep learning model, which can be a strength for data-based deep learning algorithms. In this paper, mixed data and pre-calibration data were divided into separate cases and tested. Mixed data is composed of calibrated data and original data mixed at half-and-half ratio, and the pre-calibration data is original data. The test results showed that the two cases were 0.705 (original) and 0.98 (mixed) AUC, respectively. Thus, this paper has improved the performance of sepsis predictive model by calibrating the WBC among blood test factor. The future works will verify this model in a real hospital and first it will complement the WBC calibration model according to the patient's Vital Signal. Since then, cross-verification is also planned with data from various agencies.

Article History
Article Received*: 24 July 2019*
Revised*: 12 September 2019*
Accepted*: 15 February 2020*
Publication*: 26 March 2020*

Keywords: *Deep Learning, Sepsis, Early Detecting, GAN, White Blood Cell, Calibration Model*

1. Introduction

Sepsis is a disease that infects microbes and results in systemic inflammatory reactions such as fever, rapid pulse, increased breathing, increased or decreased white blood cell count [1]. Sepsis also currently leads to the highest medical costs of all diseases [2], affecting about 30 million people a year worldwide [3]. Sepsis remains the main cause of death despite advances in modern medicine such as vaccines, antibiotics and acute treatment, leading to the death of the most humans to date. Sepsis is especially fatal for patients in the Intensive Care Unit (ICU) [4]. However, Sepsis can lower mortality rates with treatment during the initial three hours of infection and faster antibiotic administration. This is a characteristic of infectious diseases and early detection is essential for fast response [5,6]. A deep learning study that detects sepsis is a study on the evaluation of infant sepsis according to the type of pathogen [7]. This predicted sepsis using a deep learning model that studied clinical sepsis positive/negative responses based on pathogens. However, these studies are not very interested in early prediction of sepsis. As mentioned earlier, sepsis, where treatment and antibiotic administration during the initial three hours are key to treatment, should be studied with the goal of early detection for a quick response. In addition, existing studies related to early detection of sepsis have constructed a LSTM (Long Short-Term Memory) method to facilitate sequential pattern learning to predict sepsis by entering patients' Vital Signs by time-series [8]. This used the process of storing patient data on an hourly basis in intensive care units and based on MIMIC-2 data, 9 items (pulse, oxygen saturation, body temperature, systolic blood pressure, average blood pressure, respiratory rate, white blood cell count, PH, age) and 6,000 cases of patient data showed a high AUC (Area Under the Curve) of 92.9%. Similar studies also include a machine-learning model that can predict sepsis by dividing four, six, eight, and 12 hours based on ICU's high-resolution biometric signals and EMR (Electro Medical Record) data [9]. This was learned from about 52,000 data and 65 elements in total. In addition, a combination of multiple complex models was implemented in the form of ensembles, showing 85% AUC regard to the prediction of sepsis four hours ago. However, each of these studies may be missing because they classified the data using SIRS (System Inflammatory Response Syndrome), a scoring system that can diagnose sepsis in the process of classifying the data. Data without verification of data for each of these items has many difficulties in applying deep-learning technology. If deep learning is carried out with missing values, it is difficult for learning to have an effect because the weights of the nodes differ according to the values.

2. Methods

In this paper, Materials and Methods are described in three configurations. The first data analysis analyzes the data, the materials of the study, and describes the filtering process. The second process, the pre-processing model configuration, discusses pre-processing models and corrections based on the GAN (Generative Adversarial Network). The last third process, the sepsis predictive model configuration, describes sepsis predictive models.

• Data Analysis.

The data set in this paper utilizes PhysioNet's 'Early Prediction of Sepsis from Clinical Data the PhysioNet Computing in Cardiology Challenge 2019'[10, 11]. Data are data from inpatients in intensive care units released by two hospitals, including biometric signals (1 to 8), body component test results (9 to 34), other information (35 to 40) and annotations (41). As shown in Table 1. this study used nine of the above 41 items (Heart Rate, Oxygen Saturation, Body Temperature, Systolic Blood Pressure

March - April 2020
ISSN: 0193-4120 Page No. 4135 - 4143

(SBP), Mean Blood Pressure (MBP), Diastolic Blood Pressure (DBP), Respiratory Rate, White Blood Cell, Age) and Annotations.

Table 1. is a table of data used in the study. The data are approximately 800,000 cases, classified as normal and sepsis and specified the number and rate according to 10 items. Also, as shown in Table 1. The number of white blood cells is high in missing values. On the other hand, age, annotations can see that there is no missing data and both numbers and ratios are zero. This shows that the more complex the inspection process, the more missing it is. In the case of body temperature, the missing value is high because of the non-automated method input by the nurse every hour. In the case of WBC, there are many omissions because they must be accompanied by a blood test.

As mentioned earlier, deep learning cannot have enough effect on learning due to its nature if there is a lot of missing data. Therefore, this study uses a GAN (Generative Adversarial Network) model to correct the number of missing white blood cells and uses seven biological signals and age as inputs

Table 1. Missing Items and Percentage of Data.

	Normal (779,079)		Sepsis (17,136)	
	Count	Rate (%)	Count	Rate (%)
Heart Rate	60,084	7.7	1,105	6.4
Oxygen Saturation	93,447	12.0	1,632	9.5
Temperature	551,616	65.7	11,698	68.3
SBP	117,194	15.0	3,007	17.5
MBP	79,477	10.2	1,381	8.1
DBP	373,195	47.9	7,102	41.4
Respiration Rate	75,922	9.7	1,336	7.8
WBC	715,201	91.8	15,666	91.4
Age	0	0	0	0
Annotation	0	0	0	0

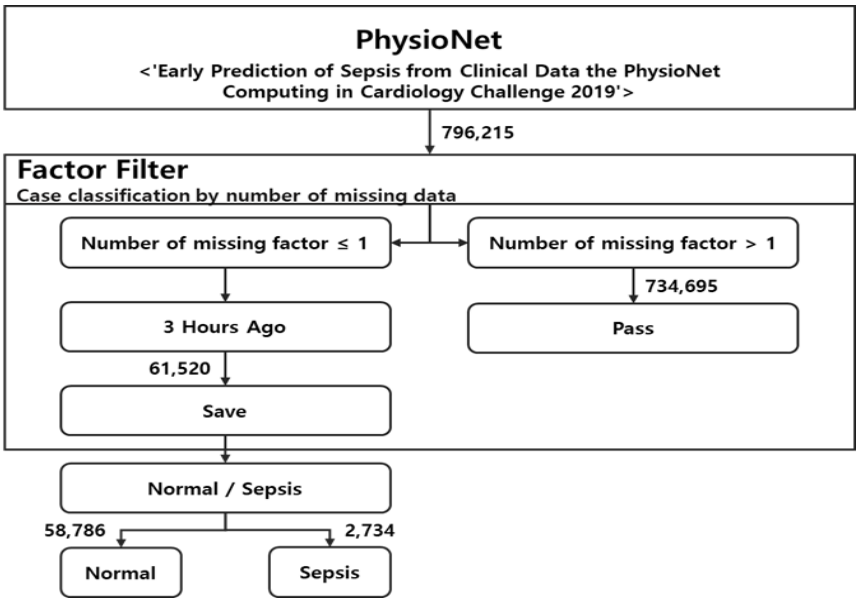

Figure 1. Data Classification and Configuration.

March - April 2020
ISSN: 0193-4120 Page No. 4135 - 4143

As you can see in Figure 1, this study has a preprocessing process before data is entered the model. Preprocessing is a factor filtering process in which data is classified by case to find missing items in one of the items in Table 1. The factor filters are shown in Figure 2. And the details are explained in the following chapters.

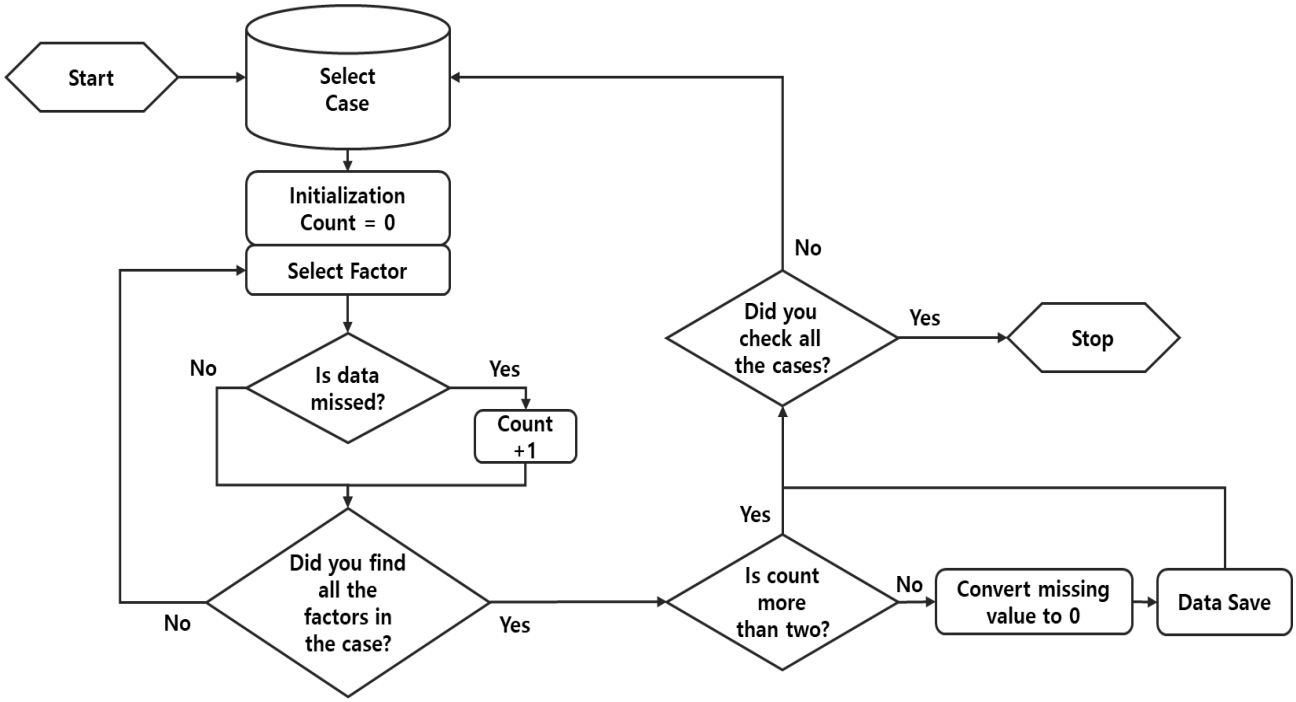

Figure 2. Factor Filter.

The factor filter examines all items on one case to examine missing values for the selected data on a case-by-case basis. This can be confirmed by Count and data is not selected when more than two missing value is generated for one case. The above is then performed on all cases and the missing values are replaced with zero for the selected data and stored.

The selection criteria for missing values considered the percentage of missing values an item has, as shown in Table 1. The purpose of this filter is to find cases that are missing and cases that are not. A few of the extracted data has WBC, but other items are missing, or all items exist. In this paper, data calibration is performed through a calibration model using GAN, so data from cases where WBC is not omitted is required.

Table 2. Examples of Data Classification.

Sortation	Save#1	Save#2	Pass#1	Pass#2
Heart Rate	61	64	66	58
Oxygen Saturation	99	98	99	100
Temperature	36.44	0	37	0
SBP	124	126	121	130
MBP	65	64	64	68
DBP	43	40	0	40
Respiration Rate	17.5	15.5	14	16
WBC	0	87	0	0
Age	75	81	76	76

March - April 2020
ISSN: 0193-4120 Page No. 4135 - 4143

Table 2 describes how case-by-case data are classified through the preprocessing. The case separated by Save#1 is stored because the number of missing values in nine items is one (WBC). Save#2 is stored because all but body temperature values exist. However, in the case of Pass#1, the list of relaxed blood pressure, white blood cell count was omitted, and in the case of Pass#2, the body temperature and white blood cell count were omitted. Thus, as shown in Figure 1. 61,520 data were extracted through the preprocessing process of 796,215 initial data. The extraction process was based on the number of items missing from the single case data. Since then, we have classified normal and sepsis and entered 58,786 normal and 2,734 sepsis into the model.

• Configuration of Preprocessing Model.

Preprocessing model 2.1 uses final classification data of data analysis as input. In this paper, a model that can compensate for missing data is composed of the GAN (Genetic Adversarial Network) of deep learning. The preprocessing model also consists of nine factors, excluding annotations, which are intended to calibrate the WBC.

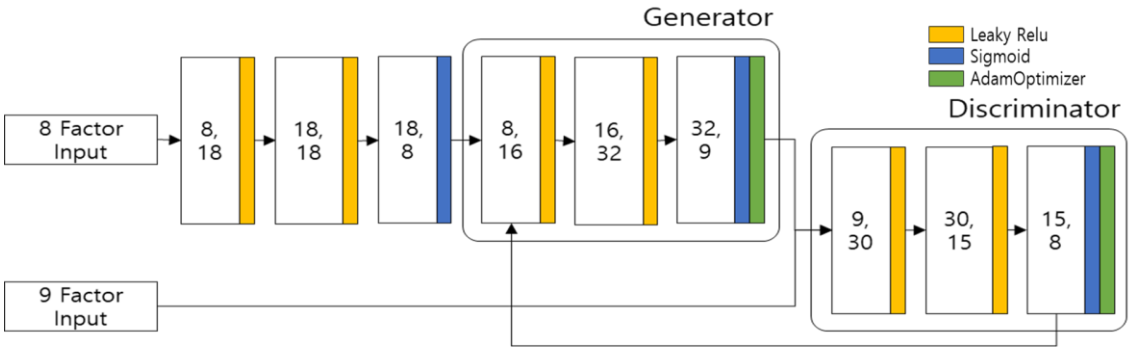

Figure 3. Preprocessing Model for White Blood Cell Correction.

As shown in Figure 3 the input of the model consists of two types. The first is the eight Factor data that is entered the generator of the GAN. This is missing WBC data and the WBC is populated by subsequent calibration. The second is the nine factor data that is entered the Discriminator. 8 factors except WBC and 9 factors including WBC are applied to the generator of the GAN and the discriminator respectively. The discriminator then uses the nine factors created by the creator as the second input to compare the data initially entered with the data generated. This adjusts the creator's weight so that the generated data is like the original data. Thus, as the learning progresses, the creator receives eight factors, which form nine like the original, and the discriminator further strengthens them.

All preprocessing models consist of FC (Fully Connected). The preprocessing model is also designed to target low complexity models. Because the data in this study does not have a larger amount of data than images or waveforms, Leaky Relu and Sigmoid were adopted as active functions to maintain the variables of deep learning.

The layer consists of the following. 8 Factor input is input by the constructor through 18 nodes on the first eight nodes and by the discriminator with 9 nodes, the output of the constructor. The

discriminator is then compared to the original case and is then passed back to the constructor with eight nodes.

The data from the 9 Factor, which was then produced by a fully-learned constructor, adds the WBC to the eight factor data that the WBC is missing. This makes it possible for the WBC to calibrate missing data to have all nine factors. In this paper, a case with the minimum difference value was selected and proceeded in the course of adding. The minimum difference value is a method of selecting and replacing data with the least difference value for all factors, between the original case and the calibrated case. This is the most intuitive way to compare similarities between cases, and the smaller the difference values, the more similar the case is.

Step 1: Calculation of the difference values is performed to compare similarity between the two cases, using the method of standard deviation. D means the sum of the squares of the difference values, and the WBC generated for the two cases with the least difference is replaced.

$$D = \sum_{i=1}^{9}(c_i - o_i)^2 \quad (1)$$

Where c is the nine factors that were finally generated by the model's creator, o is the original case. In addition, i corresponds to the number of the factor, so the above expression is the standard deviation of all factor differences between the two cases. Thus, the original data that the WBC is missing can calibrate the WBC of the data that is most like itself.

• Configuration of Sepsis Early Predictive Model.

In this paper, the preprocessing model and sepsis prediction model discussed earlier are composed of one. The sepsis predictive model combines the original data in which the data completed the calculation of the standard deviation by case above are intact. Data merges utilized the Concat function of Tensorflow, after which they were mixed at random to conduct learning and testing.

The sepsis early prediction model is constructed as shown at the bottom of Figure 4. In addition, data integrated into Concat is applied as an input for model learning. The optimization function of the model is AdamOptimizer, acting on the final layer of the creator, discriminator, and sepsis prediction model, and performed independently only in the last layer for binary classification. The criteria for binary classification are based on the annotation of the original data, with 0 being normal and 1 being sepsis.

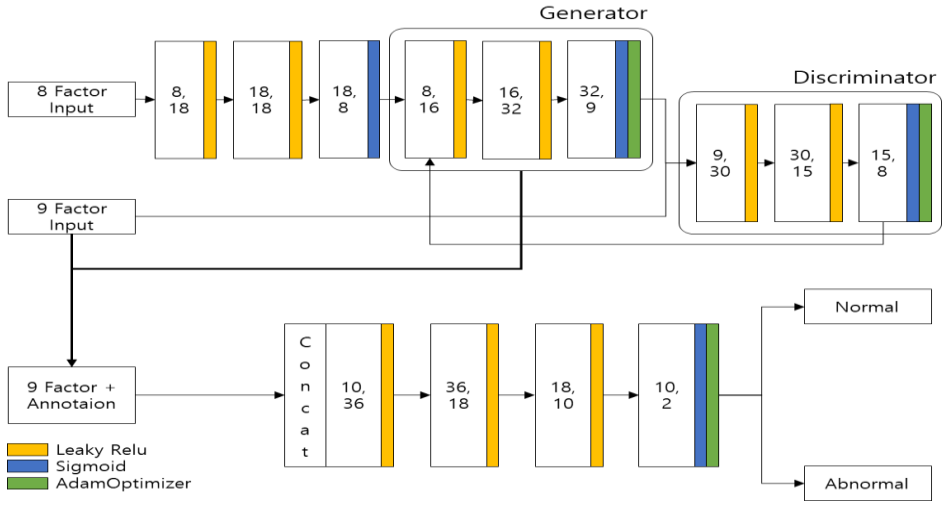

Figure 4. Sepsis Early Predictive Model.

March - April 2020
ISSN: 0193-4120 Page No. 4135 - 4143

3. Results

Here, we present the tremendous information intends to deal with the undertaking of work methodology delineated in Section 3. The fundamental objective of these architectures is to perceive how the information picture is managed since implementation. In any case, we base on Hadoop structure, and Spark system, and propose two designs for depiction experience as appeared in Fig. 1. No ifs, and or buts, the depiction organize addresses one of the standard bits of the proposed work process. Truly, the game-plan step packs each gathering of biomedical pictures (lunch hurt, pelvis, skin picture, etc.) with every sales. At long last, expressive and appraisal time will be confined both for expert or CNN calculations. From this time forward, the strategy step must be well-organized.

In this study, approximately 60,000 data were used through filtering based on approximately 800,000 data. The data were mixed at random and composed of 70%, Validation Set: 20%, Test Set: 10%. The tests were also conducted in two cases to verify the performance of the preprocessing model and sepsis detection model. The first case is a case that uses only the original data and the second case is a case that has corrected the 8 Factor data of the original data. The second case consisted of half the original data without missing values and half the corrected data. In addition, as shown in Figure 1 data from three hours ago was utilized based on sepsis annotations.

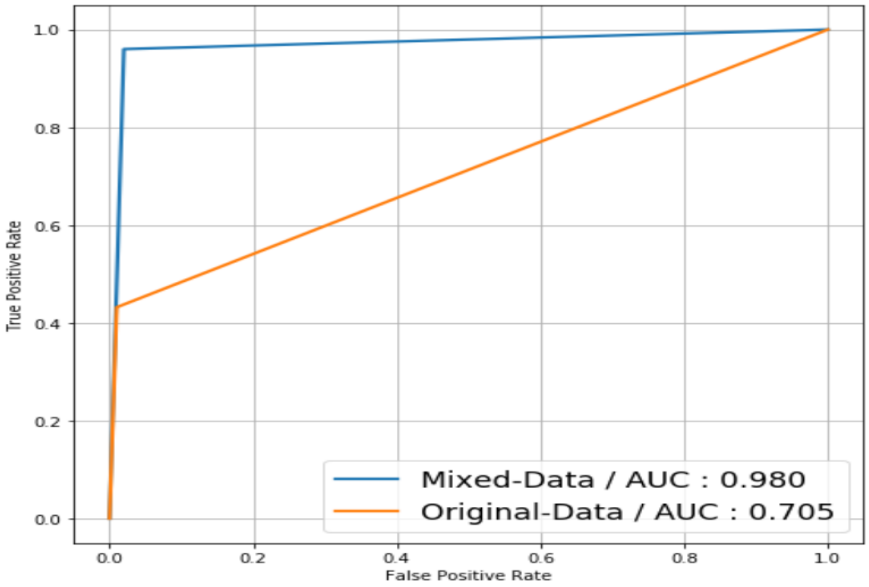

Figure 5. ROC Curves by Case.

This study shows one critical point, as shown in Figure 5. as a binary classification of normal and sepsis. Figure 5.is a ROC Curve (Receiver Operating Characteristic Curve), which uses the ratio of sepsis among the correct answers as horizontal axis and the proportion of sepsis not judged to be correct as vertical axis. In addition, the portion of the graph where one line is deflected be an indicator point in the classification. The area from the graph to the horizontal axis is referred to as AUC (Area Under the Curve), and the closer the 1 is, the more accurate it is determined. ROC Curve and AUC are used as indicators of machine learning and deep learning.

March - April 2020
ISSN: 0193-4120 Page No. 4135 - 4143

In Figure 3, the blue line means the original + corrected data, and the orange line is the original data.

As shown in Table 3, the experimental results of the original case showed that in normal cases, the Precision was 0.64, the Recall was 0.96, and the F1-Score was 0.77. In addition, for sepsis, Precision was 0.92, Recall was 0.45 and F1-Score was 0.60. The data entered in the original case were processed into a total of 2,000 cases, with 1,000 each for normal/sepsis. Thus, overall, the original cases showed Precision: 0.78, Recall: 0.70, F1-Score: 0.68, AUC: 0.705. Tests using original data showed better performance on top than sepsis. In addition, the accuracy of the model was 70.5%, like the results of several related studies.

Table 3: Test Results by Case

	AUC:	0.98			
Mixed-Case		Precision	Recall	F1-score	Number of Data
	Normal	1.00	0.96	0.98	500
	Sepsis	0.96	1.00	0.98	500
	Avg/Total	0.98	0.98	0.98	1,000
	AUC:	0.705			
Original-Case		Precision	Recall	F1-score	Number of Data
	Normal	0.64	0.96	0.77	1,000
	Sepsis	0.92	0.45	0.60	1,000
	Avg/Total	0.78	0.70	0.68	2,000

.

The top of the mixed case was Precision: 1.00, Recall: 0.96 and F1-Score: 0.98. Sepsis also showed Precision: 0.96, Recall: 1.00, F1-Score: 0.98. As the data in the mixed case is shown by the data configuration in Figure 1 the number of data was not large, so 500 of normal and sepsis were randomly mixed to use a total of 1,000. Thus, the overall mixing case shows Precision: 0.98, Recall: 0.98, F1-Score: 0.98, AUC: 0.98. Unlike the original case, mixed cases can be seen to be relatively better at judging sepsis than normal. Tests using mixed data show 98% accuracy, about 5% higher than 92.9% of studies using traditional LSTM methods and approximately 13% higher performance than 85% of studies using 65 items.

4. Conclusion

This paper researched and verified sepsis early prediction model with performance of about 98% through data correction. Early detection of sepsis is advantageous in sequential pattern learning as it receives and utilizes patient biometric information as a time series. However, if clinically important factors in diagnosing sepsis, such as the number of white blood cells, can be utilized in the predictive model, they can perform better than sequential pattern learning. This paper used a calibration model based on GAN to calibrate missing data and formed a Factor Filter to extract data applied by input to the calibration model. The corrected values were then replaced in the order in which

the difference values were less than the original values. The experimental results of this paper showed that the sepsis predictive model through data correction has improved performance over the existing sepsis predictive model. It is also thought that differences from conventional studies arise from the number of data and omission of input data. The quality of data has a significant effect on the study of predictive models using deep running. In future research, the model will be verified at a real hospital and will first complement the WBC calibration model according to patients' Vital Signals. Since then, cross-verification is also planned with data from various agencies

Acknowledgment

The research was conducted with the support of the Korea Research Foundation-Korea Research Foundation, a project to support field-specific science and engineering talent development (No.2019030267) with the funding of the government (Science, Information and Communication Ministry) in 2019.

References

[1] Stehr SN, Reinhart K. Sepsis as a global health problem-why we need a global sepsis alliance. Shock. 2013 May; 39(1):3-4.

[2] Moerer O, Schmid A, Hofmann M, Herklotz A, Reinhart K, Werdan K, Schneider H, Burchardi H. Direct costs of severe sepsis in three German intensive care units based on retrospective electronic patient record analysis of resource use. Intensive Care Medicine. 2002 Nov; 28(10):1440-6.

[3] Hajj J, Blaine N, Salavaci J, Jacoby D. The "Centrality of Sepsis": A Review on Incidence Mortality and Cost of Care. Healthcare. 2018 July; 6(3): 90-11.

[4] Topiwala R, Patel K, Twigg J, Rhule J, Barry M. Retrospective Observational Study of the Clinical Performance Characteristics of a Machine Learning Approach to Early Sepsis Identification. Critical Care. 2019 Sep; 1(9):46-6.

[5] Seymour CW, Gesten F, Prescott HC, Friedrich ME, Iwashyna TJ, Phillips GS, et al. Time to Treatment and Mortality during Mandated Emergency Care for Sepsis. New England Journal of Medicine. 2017 Jun; 376(23):2235-2244.

[6] Levy MM, Evans LE, Rhodes A. The Surviving Sepsis Campaign Bundle: 2018 update. Intensive care medicine. 2018 Jun; 44(6):925-928.

[7] Masino AJ, Harris MC, Forsyth D, Ostapenko S, Srinivasan L, Bonafide CP, et al. Machine learning models for early sepsis recognition in the neonatal intensive care unit using readily available electronic health record data. PloS one. 2019 Feb 22; 14(2):e0212665-23.

[8] Kam HJ, Kim HY. Learning representations for the early detection of sepsis with deep neural networks. Comput Biol Med. 2017 Oct; 89(1):248-7.

[9] Nemati S, Holder A, Razmi F, Stanley MD, Clifford GD, Buchman TG. An Interpretable Machine Learning Model for Accurate Prediction of Sepsis in the ICU. Critical Care. 2018 Apr; 46(4):547-6.

[10] Reyna MA, Josef CS, Jeter R, Shashikumar S, Westover MB, Nemati S, et al. Early prediction of sepsis from clinical data: the PhysioNet/Computing in Cardiology Challenge 2019. Critical Care. 2019 Aug (version 1.0.0). Available from: PhysioNet. https://doi.org/10.13026/v64v-d857

[11] Reyna MA, Josef CS, Jeter R, Shashikumar S, Westover MB, Nemati S, et al. Early prediction of sepsis from clinical data: the PhysioNet/Computing in Cardiology Challenge 2019. Society of Critical Care Medicine. 2020 Feb; 48(2):210-7.