

머신러닝 알고리즘을 이용한 패혈증 조기 예측 분석

김수민¹, 권영인¹, 강봉근¹, 황동환^{1*}

대구경북첨단의료산업진흥재단¹

Analysis of Sepsis early prediction using machine learning

Su Min Kim¹, Young-in Kwon¹, Bongkeun Kang¹, Donghwan Hwang^{1*}

Daegu-Gyeongbuk Medical Innovation Foundation, Korea

*dh0746@kmedihub.re.kr

Abstract

Sepsis is a systemic inflammatory response syndrome caused by infection, which can lead to death. Early diagnosis can help to reduce the mortality in sepsis. Recently studies on sepsis prediction using machine learning have been conducted. In this study, we suggest the sepsis onset can be predictable from patient vital signs collected on MIMIC-IV. Seven vital sign measures selected from the MIMIC-IV and their changes over time were calculated and used for input data. We evaluated AUROC and the accuracy of vital signs by learning the LightGBM algorithm to predict the onset of sepsis. Our proposed model could be predicted sepsis with AUROC 0.76 and accuracy 0.87 using vital signs 4-6 hours before onset. Our model can predict sepsis early through vital signs, and these estimates can serve as important clinical indicators of a patient's condition, especially with regard to sepsis. Further, we are going to develop a high-performance sepsis prediction model and patients monitoring system.

1. 연구 배경

패혈증은 감염으로 인해 발생하는 전신염증반응증후군으로 다양한 징후와 증상을 보인다[1]. 2017년 기준 전 세계적으로 약 48,900만 건의 패혈증 발병 사례가 보고되었고, 약 1,100만 명이 패혈증으로 인해 사망하였다. 이는 전 세계 연간 사망자의 19.7% 비중을 차지하고 있다[2]. 패혈증 사망률은 치료 수준 및 의료 인프라 부족, 열악한 감염 예방 조치, 늦은 진단과 부적절한 임상 관리와 관련이 있다[3]. Liu VX 등이 실시한 연구에서 패혈증 치료를 위한 항생제의 적절한 조기 투여는 사망률을 감소시켰다고 보고하였다[4]. 패혈증 환자의 절반 이상은 감염 이후 적절한 관리로 예방할 수 있으며 조기진단으로 사망률을 낮출 수 있다[3]

최근 개정된 Sepsis-3 기준에 따르면 패혈증은 감염으로 인한 장기부전으로 정의한다. 장기부전은 SOFA 점수가 2점 이상일 때로 추정할 수 있고 패혈증으로 진단할 수 있다[5]. SOFA 점수는 6가지 지표(0~4점)로 평가되며, 지표의 총합을 SOFA 점수로 사용한다[6]. (표 1.)

표 1. 패혈증 진단에 활용되는 SOFA 점수

System or organ and measure	SOFA score			
	1	2	3	4
Respiration PaO ₂ /FiO ₂ , mmHg	<400	<300	<200	<100
Coagulation Platelets ×10 ³ /mm ³	<150	<100	<50	<20
Liver Bilirubin, mg/dl	1.2-1.9	2.0-5.9	6.0-11.9	>12.0
Cardiovascular Hypotension	MAP<70mm Hg	Dopamine≤5 or dobutamine (any dose)	epinephrine≤0.1 or norepinephrine ≤ 0.1	Dopamine>15 or epinephrine>0.1 or norepinephrine> 0.1
Central nervous system Glasgow Coma Score	13-14	10-12	6-9	<6
Renal Creatinine, mg/dl or urine output	1.2-1.9	2.0-3.4	3.5-4.9 or <500 ml/day	> 5.0 or <200 ml/day

최근 연구에 따르면 활력 징후를 이용한 패혈증 조기진단 연구가 보고 되고 있다[7,8]. 이는 활력징후와 패혈증 증상과의 관계가 있음을 시사한다.

본 연구에서는 Gradient Boosting Decision Tree 기반의 LightGBM 알고리즘[9]을 활용해 활력 징후와 패혈증 진단기준 중 하나인 Sepsis-3의 학습으로 패혈증의 조기진단이 가능한지 분석했다.

2. 연구 방법

본 연구에서는 미국 Beth Israel Deaconess Medical Center (BIDMC)에서 2008년부터 2019년 사이에 수집된 의료데이터베이스 MIMIC-IV(Medial Information Mart for Intensive Care-IV)에서 중환자실에 입원한 환자(n= 76,540) 데이터를 활용해 패혈증 환자와 비 패혈증 환자의 활력 징후를 분류했다[10]. Johnson 등의 연구에서 사용된 Sepsis-3 분류 코드를 활용해 패혈증 진단 기준인 Sepsis-3 정의에 따라 패혈증 환자를 구분했다[11]. Sepsis-3 기준, 패혈증으로 식별된 환자는 35,010명이다.

MIMIC-IV 데이터에서 심박수, 호흡수, 수축기혈압, 이완기혈압, 평균 동맥압, 산소포화도, 체온의 7가지 활력 징후 데이터를 학습 데이터셋을 계산하기 위해 선택하였다. 패혈증 환자 데이터는 패혈증 진단 시간을 기준으로 n시간(n=1,2,3) 전과 n+1시간 전, n+2시간 전 그리고 n+1과 n+2시간 전의 활력 징후 데이터의 변화량을 각각 계산하여 총 35개의 데이터를 학습데이터로 입력하였다. 패혈증 조기예측 학습모델은 A 모델은 n이 1시간일 때, B 모델은 n이 2시간일 때 C 모델은 n이 3시간일 때, D 모델은 n이 4시간일 때 각각 학습을 진행하였다. 비 패혈증 환자는 중환자실 입원 시간부터 입원 시간 2시간 뒤 데이터를 사용해 패혈증 환자와 같은 데이터 형태로 구성했다.

결측 데이터는 발생 시점 직전 데이터를 할당했고 동일 시간 내 다수의 측정 데이터가 존재할 시, 그 평균을 사용했다. 학습 수행 시, 학습 데이터에만 최적화되어 다른 데이터로 예측을 수행하면 성능이 저하되는 과적합을 방지하기 위해 10 fold 교차검증을 시행했고, 학습과 테스트 데이터의 비는 8대 2로 구성했다.

3. 연구 결과

본 연구는 MIMIC-IV 데이터에서 7개의 활력 징후를 선택하여 학습데이터로 활용하기 위해 데이터 전처리를 진행하였다. 전처리된 데이터는 총 35개 데이터로 계산되었으며 패혈증 조기 진단을 위한 머신러닝 모델 학습의 입력데이터로 활용되었다. 머신러닝 알고리즘을 이용해 패혈증 진단 시간으로부터 3시간 전에 패혈증을 예측했고, 0.75의 AUROC와 0.81의 정확도를 확인하였다. 아래 표 2는 시간에 따른 LightGBM을 사용한 패혈증 조기 예측 결과이다. 변수 중요도 결과 A 모델과 D 모델은 3시간 전 심박수 측정값에서 2시간 전 측정값의 변화량이 가장 높았고, B 모델과 C 모델은 3시간 전 체온 측정값에서 2시간 전 측정값의 변화량이 가장 높게 나타났다. 활력징후 변화는 학습 모델이 패혈증을 조기 예측할 수 있는 중요한 인자로 평가될 수 있다.

본 연구에서는 시간에 따른 활력 징후 변화가 머신러닝 모델의 성능에 미치는 영향을 분석하였다. 패혈증 진단 4시간 전, 5시간 전, 그리고 6시간 전 활력 징후 데이터로 학습한 모델은 0.76의 AUROC와 0.87의 정확도로 가장 높은 값을 나타냈다. 이 성능은 패혈증 진단 시점 이전에 조기 예측이 가능함을 시사할 수 있으며, 의료진이 환자의 상태를 빠르게 판단하는 데 도움이 될 것으로 기대한다. 본 연구를 바탕으로 향후 높은 성능을 나타내는 패혈증 조기 예측 머신러닝 모델 개발 및 실시간 환자 모니터링 시스템의 연구개발을 추진할 계획이다.

표 2. 시간에 따른 머신러닝 패혈증 조기 예측 성능

예측 시간(t*-n)	A	B	C	D
AUROC	0.75	0.75	0.76	0.76
ACCURACY	0.81	0.84	0.86	0.87

*t=패혈증 진단시간

4. Acknowledgements

본 연구는 2019년도 산업통상자원부 및 한국산업기술진흥원(KIAT)의 국제공동기술개발사업(인공지능과 WRS 바이오마커를 이용한 고성능 패혈증 환자 모니터링 시스템) 과제의 지원을 받아 수행하였음.

5. 참고 문헌

[1] O'Brien, James M Jr et al. "Sepsis." The American journal of medicine vol. 120,12 (2007): 1012-22.

[2] Rudd, Kristina E et al. "Global, regional, and national sepsis incidence and mortality, 1990-2017: analysis for the Global Burden of Disease Study." Lancet (London, England) vol. 395,10219 (2020): 200-211.

[3] World Health Organization. (2020). Global report on the epidemiology and burden of sepsis: current evidence, identifying gaps and future directions. World Health Organization. <https://apps.who.int/iris/handle/10665/334216>. License: CC BY-NC-SA 3.0 IGO

[4] Liu, Vincent X et al. "The Timing of Early Antibiotics and Hospital Mortality in Sepsis." American journal of respiratory and critical care medicine vol. 196,7 (2017): 856-863.

[5] Singer, Mervyn et al. "The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3)." JAMA vol. 315,8 (2016): 801-10.

[6] Vincent, J L et al. "The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure. On behalf of the Working Group on Sepsis-Related Problems of the European Society of Intensive Care Medicine." Intensive care medicine vol. 22,7 (1996): 707-10.

[7] Mao, Qingqing et al. "Multicentre validation of a sepsis prediction algorithm using only vital sign data in the emergency department, general ward and ICU." BMJ open vol. 8,1 e017833. 26 Jan. 2018

[8] Barton, Christopher et al. "Evaluation of a machine learning algorithm for up to 48-hour advance prediction of sepsis using six vital signs." Computers in biology and medicine vol. 109 (2019): 79-84

[9] Guolin Ke et al. "LightGBM: a highly efficient gradient boosting decision tree". Advances in Neural Information Processing Systems 30 (NIPS 2017). Pages 3149-3157

[10] Johnson, A., Bulgarelli, L., Pollard, T., Horng, S., Celi, L. A., & Mark, R. (2020). MIMIC-IV (version 0.4). PhysioNet.

[11] Johnson, Alistair Ew et al. "The MIMIC Code Repository: enabling reproducibility in critical care research." Journal of the American Medical Informatics Association : JAMIA vol. 25,1 (2018): 32-39.