

생징후 데이터와 랩 코드 데이터를 이용한 심정지 예측

채민수*, 윤건일**, 조유진**, 김영민**, 송찬영**, 이화민**

*순천향대학교 소프트웨어융합학과

**순천향대학교 컴퓨터소프트웨어공학과

e-mail:leehm@sch.ac.kr

Prediction of Cardiac Arrest using Vital Signs and Lab code

Minsu Chae*, Geonil Yu**, Yujin Jo**, Yeongmin Kim**,
Chanyeong Song**, HwaMin Lee**

*Dept. of Software Convergence, Soonchunhyang University

**Dept. of Computer Software Engineering, Soonchunhyang University

요약

심정지는 갑작스럽게 심장이 멈추어 혈액이 흐르지 않는 질병이다. 그로인해 초기 대응이 느릴 경우 혈액이 흐르지 않아 신체 기능이 손상된다. 우리는 순천향대학교 천안병원 내 병원 환자를 대상으로 2016년부터 2019년까지의 병원 환자 대상으로 연구를 진행하였다. 학습 항목으로는 성별, 나이, SBP, DBP, 체온, 호흡, 맥박, 알부민, 크레아티닌, 헤모글로빈을 사용하였다. 우리는 LSTM 모델, 결정트리, 랜덤포레스트, 로지스틱회귀를 이용하여 심정지 예측을 하였다. 성능 비교결과 LSTM 모델이 가장 좋은 것을 확인하였다.

3. 실험 및 결과

1. 서론

심정지는 갑작스럽게 심장이 멈추어 혈액이 흐르지 않는 질병이다. 그로인해 초기 대응이 느릴 경우 혈액이 흐르지 않아 신체 기능이 손상된다. 또한 뇌손상도 일으킬 수 있는 질병이다. 병원에서는 이러한 문제를 최소화하기 위해 신속 대응팀을 운영중에 있으며 MEWS(Modified Early Warning Score)를 사용하여 운영한다[1]. 그러나 MEWS는 낮은 민감도 문제와 의료 인력 낭비라는 문제점이 존재한다[1-2]. 그에 따라 본 논문에서는 병원 환자대상으로 생징후 데이터와 랩 코드 데이터를 이용하여 심정지 예측을 하고자 한다.

2. 관련연구

현재, 심정지 예측을 위한 연구들이 진행중에 있다[2-3]. 가장 대표적인 연구는 DEWS(Deep learning-based Dearly Warning System)이 있다[2]. DEWS는 8시간의 생징후를 이용하여 점수를 심정지 점수를 예측한다[2]. 예측한 점수에 임계치를 사용하여 심정지 예측을 하였다[2].

3.1 데이터셋

본 연구팀은 순천향대학교 천안병원 내 병원 환자를 대상으로 연구를 진행하였다. 2016년부터 2019년까지 병원 환자를 대상으로 하였다. 표 1은 환자 데이터 수를 나타낸다. 한 시간당 데이터를 하나의 레코드로 구성하였다. 환자별로 36 시간씩을 사용하였다. 심정지 환자는 코드블루가 발생하기 36시간부터 코드블루가 발생한 시간까지 사용하였다. 그 외 환자는 입원일로부터 36시간까지의 데이터를 사용하였다.

[표 1] 항목 별 데이터 수

항목	데이터 수
심정지 발생한 데이터 수	1,154
그 외 병동 데이터 수	82,389

그 외 병동 데이터에는 심정지가 발생한 환자가 다른 사유로 입원하였으나 심정지가 발생하지 않은 것도 포함한다. 또한 각 환자가 입원을 하고 퇴원(혹은 사망 포함)까지를 데이터 하나로 보며 여러 번 입원 한 경우 여러 개의 데이터로 보았다. 표 2는 학습 항목이다. 나이는 입원 시 나이이다. 본 연구팀은 LSTM 모델은 3차원의 데이터로 처리하고, 결정트리,

로지스틱회귀는 2차원 이하의 데이터만 처리가능하므로 시퀀스 순서대로 입력피처를 추가하였다.

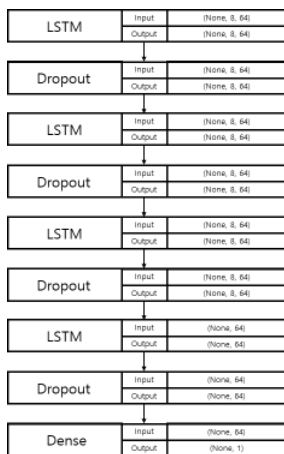
[표 2] 학습 항목

항목	데이터 수
성별	1은 남자, 2는 여자
나이	입원 시 나이
SBP	수축기 혈압
DBP	이완기 혈압
체온	체온
호흡	분당 호흡수
맥박	맥박
Albumin	알부민 수치
AlbuminCheck	알부민 값 존재 여부
Creatinine	크레아티닌 수치
CreatinineCheck	크레아틴 값 존재 여부
HB	헤모글로빈 수치
HBCheck	헤모글로빈 값 존재 여부

3.2 데이터 전처리

병원에서는 데이터를 24시간을 측정하지 않고, 의사가 환자의 상태를 고려하여 측정하도록 진료를 한다. 그에 따라 생장후 데이터 및 랩 코드의 수가 환자마다 상이한 문제가 있다. 그에 따라 누락된 데이터들은 마지막으로 측정된 항목을 그대로 사용하도록 하였다. 심정지 여부는 심정지가 발생하기 24시간부터 심정지 여부의 값에 1로 설정하였다. 그리고 병원 내 환자 중 심정지 환자의 수는 상대적으로 적기 때문에 학습이 어려운 문제점이 존재한다. 그에 따라 학습 데이터에 SMOTE(synthetic minority oversampling technique)을 이용하여 해결하였다. 테스트 데이터는 검증을 위해 SMOTE를 사용하지 않았다.

3.3 LSTM 모델



[그림 1] LSTM 모델을 적용한 심정지 예측 모델의 아키텍처

그림 1은 LSTM 모델 4개층을 이용한 심정지 예측 모델의 아키텍처이다. 표 3은 테스트 데이터셋을 이용한 혼동행렬 결과이다.

[표 3] LSTM 모델을 적용하여 얻은 혼동행렬 결과

		예측	
		심정지	심정지 X
실제	심정지	5,536	3
	심정지 X	7,758	471,253

3.3 결정트리

표 4는 결정트리를 적용하여 테스트 데이터셋을 이용한 혼동행렬 결과이다.

[표 4] 결정트리를 적용하여 얻은 혼동행렬 결과

		예측	
		심정지	심정지 X
실제	심정지	5,479	60
	심정지 X	1,404	477,607

3.4 랜덤포레스트

표 5는 랜덤포레스트를 적용하여 테스트 데이터셋을 이용한 혼동행렬 결과이다.

[표 5] 랜덤포레스트 적용하여 얻은 혼동행렬 결과

		예측	
		심정지	심정지 X
실제	심정지	5,527	12
	심정지 X	1,173	477,838

3.5 로지스틱회귀

표 6은 로지스틱회귀를 적용하여 테스트 데이터셋을 이용한 혼동행렬 결과이다.

[표 6] 로지스틱회귀를 적용하여 얻은 혼동행렬 결과

		예측	
		심정지	심정지 X
실제	심정지	4,452	1,087
	심정지 X	86,851	392,160

3.5 성능비교

[표 7] 머신러닝 및 딥러닝 모델별 성능 비교

모델	정확도	정밀도	재현도	F1 Score	AURO C
LSTM	98.398 31%	41.642 85%	99.945 84%	58.790 42%	1.00
결정트리	99.697 86%	79.601 9%	98.916 77%	88.214 46%	0.01
랜덤포레스트	99.755 44%	82.492 54%	99.783 35%	90.317 84%	0.00
로지스틱회귀	81.851 61%	4.8760 7%	80.375 52%	9.1943 6%	0.12

표 7은 LSTM 모델과 결정트리, 랜덤포레스트, 로지스틱회귀의 성능평가 결과이다. 정확도, 정밀도, 재현도, F1 Score는 소수점 여섯 번째 자리에서 반올림하였다. AUROC는 소수점 셋째 자리에서 반올림하였다. 단순하게 정밀도가 보면 랜덤포레스트가 가장 높고 그다음 결정트리가 높다. 그러나 AUROC의 값도 낮고, LSTM 모델에 비해 심정지 환자이나 심정지가 아니라고 예측한 수가 높다. 머신러닝이나, 딥러닝에서는 정확도, 정밀도가 중요한 성능평가 지표로 사용되나, 의료에서는 실제 질병이 걸린 것에 대해 정확히 예측하는 것이 더 중요한 성능평가이다. 그에 따라 본 연구팀이 수행한 모델에서는 LSTM 모델이 가장 좋은 성능을 나타내었다.

4. 결론

기존 시스템에서는 심정지 예측이 가능하나 낮은 민감도 문제가 있다. 본 논문에서 구현한 생징후 데이터와 랩 코드 데이터를 이용한 심정지 예측의 효과는 다음과 같다. 첫째, 생징후 데이터와 랩코드 데이터를 이용함으로써 기존의 머신러닝 알고리즘보다 더 높은 성능을 갖는다. 둘째, 심정지의 예측을 통해 환자의 상태를 고려하여 사전에 대비할 수 있다. 현재에는 DEWS의 임계치에 따라 정밀도가 높을 때도 있고, 낮을 때도 존재하는 문제점이 있다. 향후에는 이러한 문제를 해결하여 더 높은 정밀도를 가지고자 한다.

Acknowledgement

본 연구는 과학기술정보통신부 및 정보통신기획평가원의 대학ICT연구센터육성지원사업의 연구결과로 수행되었음 (IITP-2020-2015-0-00403). 이 논문은 2019년도 정부(과학기술정보통신부)의 재원으로 한국연구재단 -현장맞춤형 이공계 인재양성지원사업의 지원을 받아 수행된 연구임(No. 2019H1D8A1105622).

참고문헌

- [1] Subbe, C. P., et al. "Validation of a modified Early Warning Score in medical admissions." QJM: An International Journal of Medicine, Vol 94. No.10, pp. 521-526, 2001.
- [2] Kwon, Joon myoung, et al. "An algorithm based on deep learning for predicting in hospital cardiac arrest." Journal of the American Heart Association, Vol 7. No.13, e008678, 2018.
- [3] Park, Sung Bum, et al. "Developing a risk prediction model for survival to discharge in cardiac arrest patients who undergo extracorporeal membrane oxygenation." International journal of cardiology, Vol 177. No.3, pp. 1031-1035, 2014.