

DNN을 이용한 중환자 상태 징후 조기 예측

윤현석*, 박길식^o, 주해종*

*강남대학교 인공지능융합공학부,

^o경희대학교 AI기술경영학과

e-mail: {201804221, hjjoo}@kangnam.ac.kr*, gspark@khu.ac.kr^o

An Efficient Dynamic Workload Balancing Strategy

Hyun-Suk Yoon*, Gil-Sik Park^o, Hae-Jong Joo*

*Dept. of Artificial Intelligence Convergence Engineering, Kangnam University,

^oDept. of AI Technology Management, Kyung-Hee University

● 요약 ●

국내외에서 AI기반 의료 솔루션 시장은 빠른 속도로 확장 중이며 이에 따른 다양한 의학 분야에서 많은 기법을 통한 의료 AI 시스템이 등장하고 있다. 그러나 기존 다양한 AI 연구가 이뤄짐에도 아직 중환자의 징후 예측에는 많은 어려움이 있다. 또한, 중환자의 경우 현재 의료진만으로 모든 환자를 필요한 시기에 진료하기엔 어려움이 있고 환자 상태 조기 예측이 필수적임을 관련 다양한 의학 기사를 통해 쉽게 인지할 수 있다.

본 연구에서는 위와 같은 문제점을 해결하고자 중환자의 진료 결과 데이터를 활용하여 환자의 진료 후 상태를 예측하는 모델을 생성하였다. ‘용인시산업진흥원’에서 제공하는 60만여 건에 달하는 환자 데이터를 수집하여, 중환자 상태 징후를 조기에 예측할 수 있는 머신러닝/딥러닝 기반 알고리즘으로 구현한 여러 모델에 대해 비교했을 때 딥러닝(DNN) 기반 모델이 약 92%의 분류 정확도를 측정할 수 있었다.

키워드: 딥러닝(DNN), 머신러닝(ML), 중환자 예측

I. Introduction

중환자의 상태 조기 예측은 미리 조치를 취하여 생명을 구하는데 효율적이고 결정적인 역할을 할 수 있으며 의료진은 치료 계획을 조정하고 적절한 의료 리소스를 할당하여 치료 효율성 및 결과를 개선할 수 있다.

위와 같은 이유로 중환자를 조기 예측하기 위해 의료계 기업은 인공지능 바이탈케어와 같은 소프트웨어를 개발하고 많은 병원이 도입하고 있고 해외에서는 이전부터 병동에서 상태가 악화되는 환자를 조기에 파악하고 치료를 시행하여 심정지와 비계획적 중환자실 전동을 감소시킬 목적으로 1990년대에 호주, 영국, 미국에서 조기대응체계(Rapid response system[RRS]) 및 조기인지시스템(physiological track and triggerwarning system)을 도입하였다[1].

중증 질병의 경우 패혈증과 같이 사망률이 25-56%를 기록하며 갑작스러운 쇼크로 사망할 수 있어 조기예측은 필수적임이 대두된다[2].

중환자는 치료에 비용이 많이 들며 상태 악화에 따른 심리적 불안함을 유발할 수 있고 비용과 환자의 징후가 어떻게 될지 모르는 것이 중환자에게 가장 큰 불안함을 유발하는 요소이다[3]. 조기예측이 가능할 경우 비용절감과 심리적 부담을 덜어줄 수 있다.

현재 모든 중환자들의 상태를 주기적으로 확인하며 조치를 취하기엔 의료진이 부족한 상황이다. AI를 통한 중환자 상태 조기 예측은 부족한 의료진의 현상황을 대체하여 개선할 수 있으며 의료진의 높은 피로도 또한 개선할 수 있다. 이는 업무 효율과 피로에 의한 실수를 예방할 수 있어 제안하는 연구는 가치 있는 연구임이 분명하다.

본 연구는 수집한 데이터를 바탕으로 중환자의 징후를 예측하고 더 나아가 의료 효율성과 환자의 불안감을 덜어준다.

본 연구는 중환자의 위험을 조기 예측하고 웨어러블 기기를 통해 실시간 모니터링 시스템을 도입하기 위한 것이며, 구체적인 내용은 다음과 같다.

- 1) 중환자의 위험을 예측할 변수의 관계성을 분석한다.
- 2) 중환자의 위험을 예측할 모델을 비교하며 적합한 모델을 선정한다.
- 3) 모델 성능을 높이기 위해 파라미터를 조절하며 최적의 파라미터를 선정한다.
- 4) 중환자의 위험을 미리 예측하여 의료 효율성 및 환자의 불안감을 덜어준다.

II. Preliminaries

의료 헬스케어 AI는 지금까지 많은 발전을 이루고 활발한 연구가 이뤄지고 있으며 AI 헬스케어 시장도 지속적으로 커지고 있는 추세이다. AI 헬스케어 중에서도 앞서 설명했듯 중환자의 조기 예측은 필수적이며 많은 연구가 이뤄지고 있다. 조기 예측 방법론으로 데이터 수집 측면에서 가치 있는 데이터를 수집하고 관리하기 위해 중환자 데이터 베이스 구축 후 관리할 수 있는 플랫폼을 제작하여 실시간 환자의 생체 데이터를 수집 후 가공하는 ICU 플랫폼 설계[4] 연구 결과가 있다. AI 모델링 측면에서는 특정 질병들을 집중적으로 다루며 다양한 방법론의 연구가 이뤄졌다. 폐혈증의 경우 2008년부터 2019년까지 의료 정보 마트(MIMIC-IV, v.1.0) 데이터베이스에 저장된 데이터를 사용하였고 XGBoost 알고리즘을 통해 89.5% 정확도로 예측한 연구 결과가 있다[5]. C형 간염 환자 분류 에는 Kaggle에서 제공하는 C형 간염환자 데이터를 사용하여 MLP(다층 퍼셉트론)와 QUEST 의사결정 트리를 포함한 앙상블 모델을 개발하였고 최종 개선된 분류 정확도는 94.67%를 기록하였다[6].

III. The Proposed Scheme

3-1 데이터 수집

본 연구는 수집된 수많은 데이터를 바탕으로 유의미한 결과를 도출하기 위해 중복되거나 데이터의 불균형을 초래할 수 있는 변수들을 제거하여 차원축소를 진행하고 다양한 분석기법을 비교 분석하여 최적의 모델을 사용하여 중환자 위험군을 분류한다. 중환자의 실시간 모니터링을 위해 웨어러블 기기를 통한 앱을 설계한다. 연구 대상자는 약 10년간 국내 종합병원에서 항생제를 처방받은 입원/외래 환자를 대상으로 하며 암호화된 환자 내원번호와 입력일자로 구분되어 있어 환자의 비밀을 유지한다. 환자의 정보로 연령대와 성별 진료부서는 식별할 수 있으며 환자의 위험을 예측 및 분석하기 위해 환자의 생체신호, 진단명, ICD-10코드, 항생제 처방 종류, 항생제 처방 용량, 검사결과, 비당질환, ATC 내용을 독립변수로 사용하고 환자의 치료결과를 종속변수로 위험한 중환자를 분류한다.

자료 수집은 ‘용인시산업진흥원’에서 제공하며 현재 ‘감염병 빅데이터 거래스’에서 운영되는 검사실 소견, 환자 비당질환, 환자 생장후, 항생제 처방 종류, 배양검사 결과, 치료결과 자료를 수집하였다.

3-2 데이터 분석 절차

수집된 자료를 분석하기 위해 중복되는 변수와 데이터 불균형을 초래할 수 있는 불균형 데이터 제거가 선행적으로 이뤄졌고 중환자를 분석하기 위해 일반환자와 응급환자는 분류대상에서 제거하였으며 분류할 치료결과 데이터는 48시간 이내 사망, 48시간 이후 사망, 호전 인원은 위험으로 분류 경계, 원인은 호전됨으로 분류를 진행하였다. 변수들 사이의 범위 차이를 조정하기 위해 표준화 작업을 진행한 후 분석하였다. 자세한 자료 분석 방법은 다음과 같다.

- 1) 환자의 변수의 연관관계와 상관관계를 분석한다.
- 2) 중환자의 위험을 예측 분류하기 위해 앙상블, 로지스틱 회귀, 딥러닝 기법을 사용한 후 성능을 비교하여 최적의 모델을 채택한다.
- 3) 채택된 모델을 검증하기 위해 교차 검증을 진행한다.
- 4) 채택한 모델의 성능을 개선하기위해 파라미터를 조절한다.

IV. Experiment result

중환자 상태를 예측 분류하기 위한 알고리즘으로는 랜덤포레스트, 로지스틱 회귀, Xgboost, DNN, 랜덤포레스트와 로지스틱 DNN을 조합하여 투표를 진행한 커스텀 앙상블 5가지를 비교하였다.



Fig. 1. 중환자 예측 모델 정확도 비교

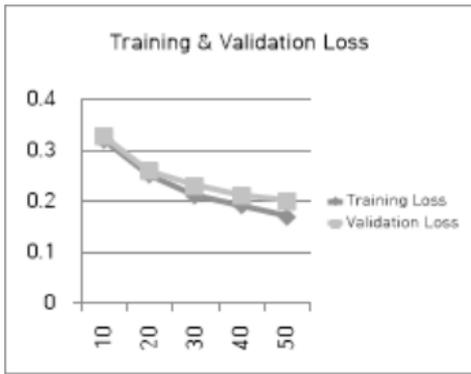
분류 예측을 진행하였을 때, 전체적으로 정확률이 높게 나오고 커스텀 앙상블이 분류 정확도 0.97을 기록하며 가장 우수한 성능을 보인 듯 하지만 DNN과 로지스틱 회귀분석을 제외하고 다른 모델은 과적합 문제를 극복하지 못하였다. 따라서 본 연구는 분류 정확도 92%에 해당되는 DNN을 최종 모델로 채택 하였다.

DNN은 총 3개의 은닉층으로 설정하며 퍼셉트론 개수는 각각 24, 8, 1개로 구성하고 은닉층 함수는 각각 ReLu, ReLu, Sigmoid로 구성하였다.



epoch	10	20	30	40	50
Training acc	0.858	0.897	0.915	0.921	0.927
Validation acc	0.856	0.895	0.905	0.913	0.920

Fig. 2. DNN 분류 정확도



epoch	10	20	30	40	50
Training Loss	0.330	0.250	0.216	0.194	0.179
Validation Loss	0.355	0.257	0.230	0.211	0.20

Fig. 3. Training & Validation Data Loss

학습한 모델을 Validation data에 적용하였을 때 위 그래프와 같이 호전될 환자와 위험 환자를 0.92의 분류 정확도로 분류하였다. 손실함수는 Binary-Cross-Entropy를 사용하였고 최종 유효 손실은 0.20을 기록하였다.

이진 분류 모델의 성능을 평가하기 위해 Roc 커브를 확인하였다. AUC = 0.98을 기록하며 우수한 성능을 보였다.

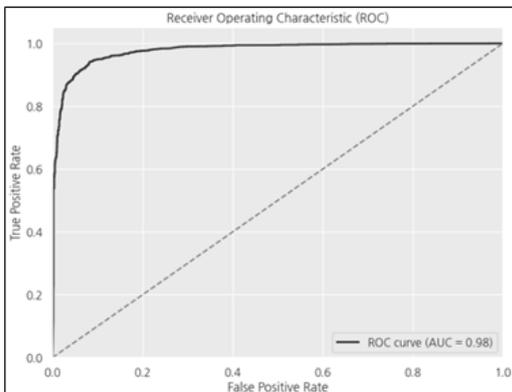


Fig. 4. ROC Curve

V. Conclusions

DNN을 통해 중환자의 상태 징후가 호전될지 위험할지 예측해 보였다. DNN은 환자의 조기 위험을 예측 분류에 있어 분류 정확도 92%로 우수한 성능을 나타냈다. 데이터베이스 관리 및 수집 기술이 지속적으로 발전함에 따라 다양한 환자 정보를 수집하게 된다면 다양한 데이터를 활용하여 더욱 우수한 연구 결과를 전망한다. 본 연구를 통해 중환자 모니터링에 딥러닝 기술이 접목된다면 의료진의 높은 업무 효율과 환자 관리, 환자의 안정감 등 많은 부분에 이점을 기대한다.

REFERENCES

- [1] Lee A, Bishop G, Hillman KM, Daffurn K. The medical emergency team. *Anaesthesia and Intensive Care*. 1995;23(2):183-186.
- [2] Lee Ju Ry, Choi Hye Ran (2014). Medical Alert Team, Asan Medical Center. Validation of a Modified Early Warning Score to Predict ICU Transfer for Patients with Severe Sepsis or Septic Shock on General Wards. *J Korean Acad Nurs Vol.44 No.2*, 219-227
- [3] Han, Kyoung Shin, & Park, Young Im (2002). The Level of Anxiety and Relating Factors of ICU Patients. *The Journal of Korean Academic Society of Nursing Education*, 8(1), 155-166.
- [4] Yong-Hwan Kim, Jang-Yong Kim, & Ki-Tae Bae (2019). A Study on Artificial Intelligence-based Intensive Medical Patient's Shock Pre-Detection System. *정보기술융합공학논문지*, 9(2), 49-56.
- [5] Hu, C., Li, L., Huang, W. et al. Interpretable Machine Learning for Early Prediction of Prognosis in Sepsis: A Discovery and Validation Study. *Infect Dis Ther* 11, 1117-1132 (2022).
- [6] Michael Onyema Edeh et al. Artificial Intelligence-Based Ensemble Learning Model for Prediction of Hepatitis C Disease. *Front. Public Health*, 27 April 2022 Sec. Digital Public Health. Volume 10 - 2022