

심전도 신호 이진 분류에 대한 언어 모델 활용

하석현⁰¹, 김수현¹, 이해인¹, 김예준¹, 강경래¹, 손연경², 권선영^{1,2,3*}

¹부산대학교 의생명융합공학부, ²부산대학교 정보융합공학과, ³부산대학교 인공지능융합연구센터
{tjtigus, ksh114612, haein1012, kyj361, kangkr1002, sy.kwon}@pusan.ac.kr

Leveraging Language Models for Binary Classification of ECG Signal

Seokhyeon Ha⁰¹, Suhyun Kim¹, Haein Lee¹, Yejun Kim¹, Kyungrae Kang¹, Yeongyeong Son², Sunyoung Kwon^{1,2,3*}

¹School of Biomedical Convergence Engineering,

²Dept. of Information Convergence Engineering,

³Center for Artificial Intelligence Research, Pusan National University

요약

본 연구는 심전도 신호의 정상과 비정상 분류를 위해 대형 언어 모델을 활용한다. 이를 위해, 샘플링 간격 변경, 데이터 압축, 심전도 신호의 임상적 특징을 추출하여 텍스트로 변환하는 데이터 전처리를 제안하였다. 실험 결과, 언어 모델의 입력 길이 제한과 메타 데이터의 활용이 성능 향상에 중요한 요소로 작용함을 확인하였다. 이러한 결과는 대형 언어 모델이 심전도 신호와 같은 복잡한 시계열 데이터를 효과적으로 분석할 수 있는 가능성을 보여주며, 적절한 프롬프트와 메타 데이터의 통합이 모델의 분류 성능 향상에 기여함을 확인하였다.

1. 서론

심전도는 심장의 전기적 활동을 측정하고 분석하여 심장 건강 상태를 평가하는 중요한 도구이다. 최근 인공지능을 활용한 심전도 연구 시도가 증가하고 있다 [1]. 다양한 도메인에서 탁월한 성과를 보이는 대형 언어 모델을 심전도 분야에서도 적용하려는 관심 또한 높아지고 있다 [2]. 본 논문은 심전도 분석에서 대형 언어 모델의 적용 방향성을 탐구하고, 대형 언어 모델이 심전도 분석에 활용될 수 있는 가능성을 제시한다.

2. 배경 및 관련 연구

2.1 심전도

심전도는 6개의 흉부 전극(V1-V6)과 6개의 사지 전극(I, II, III, aVR, aVL, aVF)으로 총 12리드로 측정되며, 심장 내 특정 부위에서 발생하는 전기적 변화를 반영하는 PQRST파로 구성된다. 임상에서는 이러한 심전도 신호를 P파, QRS복합체, T파 단위 혹은 PR간격, QRS 지속 시간, QT간격, ST분절로 분석하여 진단한다.

2.2 대형 언어 모델

T5 [3]는 텍스트-투-텍스트 접근 방식을 통해 일관된 방식으로

로 여러 작업을 처리하는 모델이다. BART [4]는 텍스트 복원을 통해 학습하는 모델로 텍스트 생성에 특화되어 있다. LED [5]는 긴 시퀀스를 처리하기 위한 효율적인 어텐션 메커니즘을 제공하여 기존 모델의 한계를 극복한 모델이다.

2.3 PromptCast

PromptCast [6]는 시계열 데이터에 적합한 프롬프트를 생성해 대형 언어 모델을 활용한 예측 작업을 수행하며, 프롬프트 기반 학습 패러다임을 제시하였다. 날씨, 에너지 소비 등 다양한 도메인에 따라 다른 프롬프트를 적용하여 대형 언어 모델이 시계열 데이터의 복잡한 패턴을 효과적으로 학습할 수 있음을 보였다.

본 연구에서는 이러한 패러다임을 심전도 데이터에 적용하였다. 이때, 예측 작업 대신 분류 작업으로 변경하여 심전도 신호의 정상과 비정상을 분류하는 실험을 수행하였다.

3. 제안 방법

각 대형 언어 모델의 입력 길이는 LED는 16,384 토큰, BART는 1,024 토큰, T5는 512 토큰으로 제한된다. 이를 극복하기 위해, 임상적으로 가장 널리 쓰이는 리드 II 신호만을 활용하였다. 또한 원본 심전도의 신호 샘플링 간격 변경, 데이터 압축, 임상적 특징 활용 방안을 제안한다. [표 1]에서 사용한 입력 처리 방식을 확인할 수 있다.

¹ 본 연구는 과학기술정보통신부 및 정보통신기획평가원의 인공지능융합혁신인재양성사업 연구 결과(IITP-2024-RS-2023-00254177)와 생성AI선도인재양성사업 연구 결과(IITP-2024-RS-2024-00360227)로 수행되었음

표 1. 심전도 분석을 위한 프롬프트

방식	입력 차원 (원본 : 5000)	템플릿	예시
샘플링 간격 변경	60 간격: 83 30 간격: 166 15 간격: 333 2 간격: 2,500	ECG Data: {ECG_Data} this data is a ECG signal sampled at(500/sampling_rate)Hz, length of 10 seconds	ECG Data : -0.075,-0.065,...-0.058 this data is a ECG signal sampled at 33Hz, length of 10 seconds
데이터 압축	25 간격: 200	ECG Data[value, number of same value (omitted when 1)] : {Compressed_ECG_Data} this data is a ECG signal sampled at(500/sampling_rate)Hz, length of 10 seconds	ECG Data[value, number of same value (omitted when 1)] : [-0.075, 2], -0.065,...,-0.058 this data is a ECG signal sampled at 33Hz, length of 10 seconds
임상적 특징 활용	6 개의 정보 (심박수, PR 간격, QRS 간격, QT 간격, ST 분절, T 파 반전)	Patient's ST segment analysis ({ST_segment}), Heart rate and rhythm analysis ({Heart_Rate} bpm, {Rhythm}), PR interval analysis ({PR_interval}), QRS duration analysis ({QRS_duration}), QT interval analysis ({QT_interval}), T-wave inversion analysis ({T_wave_inversion})	Patient's ST segment analysis (ST segment elevation: -0.020), Heart rate and rhythm analysis (Heart Rate: 73.7 bpm, Rhythm: regular), PR interval analysis (PR interval: 0.172 s), QRS duration analysis (QRS duration: 0.181 s), QT interval analysis (QT interval: 0.363 s), T-wave inversion analysis (T-wave inversion: False)
메타 데이터	3 개의 정보 (환자의 나이, 성별, 몸무게)	Demographic information {Age:{age}, Sex:{sex}, Weight:{weight}}.	Demographic information (Age: 48.0 years, Sex: Male, Weight: weight: 95.0). Based on this analysis, classify the condition as 'Normal' or 'Abnormal'.
공통 입력	Based on this analysis, classify the condition as 'Normal' or 'Abnormal'.		
공통 출력	Based on this analysis, the most likely classification is:{Normal/ Abnormal}		

3.1 샘플링 간격 변경

[그림1]과 같이 입력 데이터 포인트 양을 줄이기 위해, 10초 길이의 500Hz 로 총 5000개를 가진 원본 신호를 60, 30, 15, 2 간격으로 분할한다. 예를 들어, 총 5000개의 신호를 60 간격으로 분할되어 83개의 신호만 입력한다

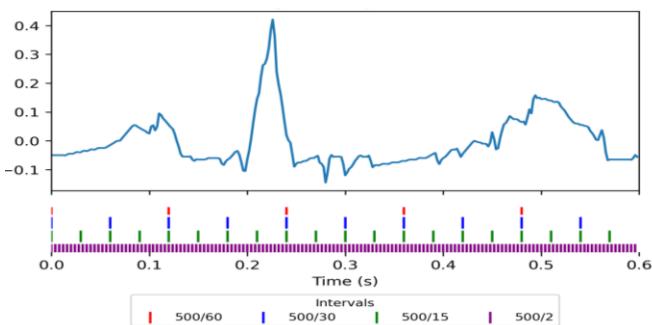


그림 1. 샘플링 간격 예시

3.2 데이터 압축

본 방법은 동일한 값이 빈번하게 나타나는 심전도 데이터를 효과적으로 압축한다. 총 5000개를 가진 원본 신호를 25 간격으로 분할한 후, 중복된 값은 [값, 중복된 횟수] 형식으로 처리

하여 압축한다. 예를 들어, [-0.017, -0.03, -0.03, 0.08, ...]의 신호를 가지는 경우 [-0.017, [-0.03, 2], 0.08, ...]의 형태로 처리한다.

3.3 심전도 신호의 임상적 특징 활용

NeuroKit2 [7] 라이브러리를 활용하여 심전도 데이터의 주요 임상적 지표인 심박수, PR 간격, QRS 간격, QT 간격, ST 분절, T파 반전을 추출한다. 따라서 심전도 신호가 아닌 6개의 임상 지표가 입력으로 사용된다. 이는 임상에서 의료진이 해당 지표로 진단하는 메커니즘을 모방한 것이다.

3.4 메타 데이터 추가

환자의 나이, 성별, 몸무게 정보가 담긴 메타 데이터를 앞서 언급된 입력과 함께 사용한다. 이는 심전도 신호를 압축하는 과정에서 발생하는 정보 손실을 보완하여 분류 정확도를 높이기 위함이다.

4. 실험

4.1 데이터

본 연구에서는 PhysioNet의 12리드로 구성된 PTB-XL 데이터를 활용하였다. NORM, HYP, MI, CD, STTC로 총 5개의 멀티레이블을 가지는 데이터를 이진 분류 작업으로 구성하기 위하여 NORM은 정상, 그 외는 모두 비정상으로 라벨링 하였다. 이때, 정상과 비정상이 혼재된 샘플은 제외하였다. 학습, 검증, 테스트 데이터는 7:1:2 비율로 분할하였다. 총 샘플 수는 17,053개이며, 이 중 정상 샘플은 7,229개(42.4%), 비정상 샘플은 9,824개(57.6%)로 두 클래스는 비교적 균등한 분포를 가진다.

4.2 실험 방법

본 연구에서는 PromptCast의 프롬프트 기반 학습 패러다임을 적용하여 대형 언어 모델을 심전도 데이터 분석을 위해 미세 조정하였다. 생성된 문자열 형태의 출력에서 분류 결과를 추출하기 위해 문자열 파싱 기법을 사용하였다. 평가 지표는 정확도를 사용하였다.

5. 실험 결과

5.1. 샘플링 간격 변경 결과

[표 2]에 샘플링 간격 별 성능을 나타내었으며, 모델 별 가장 높은 성능을 굵게 표시하였다. 모델의 입력 길이 제한으로 일부 작은 간격의 결과는 확인하지 못했다. 반면 LED는 긴 시퀀스를 처리할 수 있어 2 간격에서도 결과를 확인할 수 있었다. 동일 조건에서 대형 언어 모델과 CNN은 유사한 성능이 확인되어 활용 가능성을 시사한다. 전체적으로는 30, 15 간격에서 전반적으로 우수한 성능을 보였으며, 이는 해당 간격으로 샘플링한 결과가 심전도 신호 분석에 적합한 것으로 해석된다. 60 간격에서는 신호의 세밀한 변화를 놓쳐 성능이 저하되었고, 2 간격에서는 입력 길이 증가로 인한 과도한 정보로 인해 일부 모델 사용이 제한되거나 성능이 저하되었다.

표 2. 간격에 따른 성능

	60 간격	30 간격	15 간격	2 간격
CNN(baseline)	0.755	0.785	0.812	0.804
T5	0.759	0.783	0.798	-
BART	0.737	0.766	-	-
LED	0.656	0.763	0.755	0.743

5.2. 입력 처리 방식과 메타 데이터 활용 결과

제안하는 입력 처리 방식과 메타 데이터 사용 여부에 따른 성능 결과를 [표 3]에서 확인할 수 있다. 이때, 앞선 실험에서 우수한 성능을 보인 30, 15 간격의 성능과 비교하였다.

메타 데이터를 활용하지 않았을 때, 전반적으로 원본 신호의 샘플링 간격을 다르게 압축한 경우에 우수한 성능을 보였으며, 데이터 압축에서도 유사하게 성능이 나타났다. 이러한 결과는 중복되는 값을 처리하여도 정보 손실에 큰 영향을 미치지 않음을 시사한다. 반면 임상적 특징을 활용한 방법은 심전도 신

호 분석에 충분한 정보를 제공하지 않는 것으로 판단된다. 메타 데이터를 활용한 경우, BART를 제외하고 전반적으로 향상된 성능 결과를 보였다. 특히, 원본 데이터 처리 과정에서 손실된 정보가 큰 임상적 특징 활용에서 개선 효과가 두드러지게 나타난다. 이는 메타 데이터가 분류 정확도를 높이는 데 긍정적인 영향을 주는 것으로 확인된다.

표 3. 입력 처리 방식 및 메타 데이터 사용에 따른 성능

		메타 데이터 X	메타 데이터 O
30 간격	T5	0.783	0.798
	BART	0.766	0.740
	LED	0.763	0.773
15 간격	T5	0.798	0.804
	BART	-	-
	LED	0.755	0.773
데이터 압축	T5	0.767	0.795
	BART	0.755	0.775
	LED	0.775	0.778
임상적 특징 활용	T5	0.656	0.739
	BART	0.708	0.672
	LED	0.705	0.744

6. 결 론

본 논문은 대형 언어 모델이 심전도 데이터 분석에 적용할 수 있는 가능성을 실험적으로 확인하였으며, 심전도에 적합한 프롬프트 생성 방법을 제안하였다. 실험 결과는 원본 데이터를 단순 샘플링 기법을 적용했을 때 가장 우수하였으며, 대형 언어 모델이 텍스트 형태로 변환된 심전도를 학습할 수 있음을 보여주었다. 임상적 특징을 활용한 프롬프트의 경우 정보 손실에도 불구하고, 메타 데이터를 활용한 보완 능력을 증명하였다. 본 연구는 심전도에서의 대형 언어 모델의 적용 가능성을 확인하였으며, 향후 언어 모델을 활용한 심전도 분석 실험을 지속적으로 진행할 예정이다.

참 고 문 헌

[1] Nils Strodthoff, Patrick Wagner, et al. "Deep Learning for ECG Analysis: Benchmarks and Insights from PTB-XL" arXiv preprint arXiv:2004.13701v1, 2020

[2] Shervin Minaee, Tomas Mikolov, et al. "Large Language Model: A Survey", arXiv preprint arXiv:2402.06196v2, 2024

[3] Raffel, C., et al. "Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer", 2020

[4] Lewis, M., et al. "BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension", 2020

[5] Beltagy, I., et al. "Longformer: The Long-Document Transformer", 2020

[6] Hao Xue, F. D. et al. "PromptCast: A New Prompt-based Learning Paradigm for Time Series Forecasting." IEEE Transactions on Knowledge and Data Engineering, 6851-6864, 2024

[7] Makowski, D., et al. "NeuroKit2: A Python toolbox for neurophysiological signal processing. Behavior Research Methods",53(4),1689-1696, 2021