

설명 가능한 딥 러닝을 이용한 심전도 데이터 분석 연구

요약

최근 AI 기술이 4차 산업혁명의 핵심기술로 자리잡으면서, 관련 연구가 활발하게 진행되고 있다. 특히 기계학습분야에 많은 발전이 있었고, 다양한 모델들이 다양한 분야에서 실험적으로, 또 실용적으로 사용되고 있다. 의료분야에서도 쓰이고 있는데 심전도 데이터(ECG) 분석의 경우가 대표적이다. 기존에는 숙달된 인력이 여러 파형을 보고 분석해 결과를 추론해야 했지만, 딥 러닝 모델을 통해 이 과정을 단순화, 가속화할 수 있음과 동시에 높은 정확성을 갖게 한다. 하지만 동시에 한계점을 가지는데, 바로 기계학습 모델의 블랙박스 특성으로 인해 어떤 과정을 통해 결론을 도출하였는지 알기 어렵다는 것이다. 때문에 이에 관한 연구를 통해 결과의 신뢰성을 높이고, 판별을 위해 기존에 사용되지 않았던 중요한 특징을 역으로 발견할 수도 있다.

1. 서론

1.1. 연구배경

딥 러닝 모델의 높은 정확도와 가능성으로 인해 쓰이지 않는 곳을 찾아보기가 더 어려워졌다. 이는 기존에 인간이 많은 시간과 정성을 들여서 해야 했던 일을 쉽게 대신 해내기 시작했고, 그를 넘어서 인간이 해내지 못했던 분야까지 해내고 있다. 하지만 모델이 점점 더 복잡해짐에 따라 '블랙박스'라는 특성을 갖게 되는데, 결과의 정확도는 높지만, 어떻게 그러한 결과를 도출했는지는 들여다보기 어렵게 된 것이다.

이 맹점은 실패했을 경우의 리스크가 큰 분야일수록 크게 다가온다. 특히나 의료분야에서 활발히 연구되고 있는 ECG 분석의 경우가 그렇다. 심전도 데이터로 크게는 부정맥이나 빈혈의 가능성을 진단할 수 있는데, 실패할 경우 인명에 영향을 미친다. 또한, 환자의 경우에는 증상에 대해 설명 받을 권리가, 의사의 경우 설명해주어야 할 의무가 존재하는데, 블랙박스 형태의 모델의 경우 여기서 벗어나 있다.

따라서 단순히 모델의 높은 정확도만을 목표로 할 것이 아니라, 동시에 특이점을 포착해서 실질적인 사용자(모델의 개발자가 아닌)가 이해하기 쉬운 정보로 표현할 수 있어야 한다. 이 과정이 잘 정착되면, 사용의 용이성뿐만 아니라, 딥 러닝 모델의 신뢰성을 확보할 수 있는 장치로 자리매김 할 것이다.

1.2. 연구목표

다양한 상태의 ECG 데이터를 기본으로 딥 러닝 모델을 이용한 데이터의 가공 및 해석을 진행한다. ECG 데이터는 천편일률적이지 않고 다양한 상태를 갖을 수 있기 때문에, Self-supervised learning 을 이용한 ECG 데이터의 추가적인 생성이나, 다른 모델을 통한 접근으로 건강의 이상 유무를 넘어서, 어떤 문제가 있는지 classification 하는 것까지 생각해 볼 수 있다.

실질적인 판단모델의 경우에는 특이점이 존재하는 Rhythm 을 포착해내고, 앞선 데이터 가공 모델이 효과적으로 기능할 경우에 해당 특이점이 어떤 질환을 암시하는지 classification 할 수 있도록 한다.

특이점의 포착은 실제 수치와 정상일 경우의 추정치를 동시에 표현해 사용자(의사 또는 환자)가 딥 러닝 모델의 판단 근거를 확인할 수 있도록 한다. 개발자가 아닌 실질적인 사용자가 쉽게 이해할 수 있도록 만드는 것 역시 주안점이다.

2. 관련연구

2.1. 딥 러닝 모델 종류

2.1.1. GAN

Generative Adversarial Network(생산적 적대 신경망)은 ECG 관련 연구에 활발히 쓰이고 있는 딥러닝 모델이다. 크게 비지도 학습을 통해 데이터를 만들어내는 Generative Model 과 지도 학습을 통해 진위를 가려내는 Discriminative Model 의 두 부분으로 나뉜다. 이 두 모델이 서로 생성하고 판별하는 과정을 거쳐가면서 Generator 가 점점 더 진짜 같은 데이터를 만들어내게 되는 모델이다.

2.1.2. Autoencoder

Autoencoder 는 비지도 학습 모델로, 입력 받은 데이터를 학습해 입력된 데이터에 최대한 가까운 데이터를 생성하는 것이 목표이다. 쉽게 말하면 입력을 예측하는 모델로, 주로 이상 감지에 사용된다. 예측한 입력 값과 실제 입력 값의 비교를 통해 이상을 감지하는 방식으로 사용된다.

2.1.3. DNN

Deep Neural Network 은 Hidden Layer 을 여러 층을 쌓아서 학습의 결과를 향상시키는 방법이다. Data Classification 에 사용되는 모델이며, 충분한 양의 데이터를 이용해 반복학습, 오류역전파 등의 기법을 통해서 높은 정확도의 분류가 가능하다. DNN 을 기반으로 여러 응용 알고리즘이 탄생했다. 여기서 Hidden Layer 의 층이 늘고, 구조가 복잡해짐에 따라, Black-box 특성이 탄생되었다.

2.2. Deep Neural Network Explanation Methods

앞서 서론과 연구배경에서도 언급했지만, 최근 기계학습 분야에 대한 관심도는 최고조 상태이고, 이에 따라 다양한 모델들의 탄생과 정확도 향상이 이루어졌다. 인력으로는 상당한 시간을 요하는 부분을 기계가 빠른 시간에 대신해내게 되었지만, 모델은 더 복잡 해졌고 인간이 이 과정을 들여다보는 것은 더 어려워졌다. 이에 따라 최근에는 모델의 시각적인 설명 프레임워크에 대한 연구 역시 대두되고 있다.

3. 프로젝트 내용

3.1. 모델 학습에 필요한 데이터

어떤 형태의 심전도 데이터를 이용하느냐에 따라 만들어지는 모델이 다르다. 기본적으로는 12-lead 심전도 데이터를 사용하는데, 정상적인 데이터만을 이용해 이상을 감지할 것인지, 비정상적인 데이터까지 이용해 특정 증상을 진단해낼 것인지, 또한 정적인 상태에서의 측정데이터만을 이용할 것인지 기타 활동 상태의 데이터까지 이용할 것인지를 결정해, 보다 활용성 높은 모델을 만들어 낼 수 있다.

3.2. 딥러닝 모델 종류

3.2.1. 새로운 ECG 데이터 생성 model

GAN 과 ECG Dataset 을 이용해 진짜와 굉장히 비슷한 데이터를 생성한다. 모델의 Generator 는 난잡한 ECG 데이터를 생성해내고, Discriminator 는 생성된 데이터의 진위를 판단한다. 이 판별결과를 통해 점점 Generator 는 진짜와 다를 것 없는 데이터를 만들어내게 되고, Discriminator 는 결국 진위 판별을 할 수 없는 불능상태가 된다. 이 과정을 통해, ECG Dataset 의 부족과 불균형을 해소하고, 보다 정확한 anomaly detection model 과 symptom classification model 을 학습시킬 수 있다.

3.2.2. anomaly detection model

Autoencoder 는 입력 받은 데이터를 통해 입력 받은 데이터를 출력한다. 이 특성을 통해, 입력 받은 심전도를 예측하고, 실제 입력 받은 값과의 차이를 통해, anomaly detection 이 가능해진다. 정확한 진단에 앞서 이상 유무를 판별하고 보여주는 역할을 할 수 있을 것으로 보인다.

3.2.3. symptom classification model

Dataset 이 충분하다면, 앞서 포착해낸 특이점이 어떤 증상을 암시하는지 진단해내는 모델이다. 앞서 GAN 을 통해 충분한 데이터를 생성해내는데 성공했다면, 이를 기반으로 DNN 에 증상의 진단이 labeling 된 ECG Data 를 학습시켜, classification model 을 만들어 낼 수 있다.

3.3. 모델 설명 방법

anomaly detection model 을 사용하기 때문에 특이점이 발생할 경우, Autoencoder 는 정상적인 상태일 경우 입력 받은 ECG Data 와 특이점이 발생한 실제 환자의 입력 Data 를 동시 출력해낼 수 있다. 이를 통해 두 ECG 의 파형을 그래프 형태로 출력하거나, P-wave, PR interval, QRS complex 등의 수치 차이를 제시하는 방법으로, 의사와 환자의 특이점 파악을 용이하게 한다.

4. 결론 및 기대효과

ECG 데이터의 가공을 통해 판단 근거가 부족해 질환의 판별이 불가능한 것을 대처하고, 전문인력 없이도 높은 정확도의 건강 이상 유무, 더 나아가서는 질환의 위험을 파악할 수 있게 한다.

또한 모델을 투명하게 제작하여 사용자가 특이점과 정상 수치를 직접 확인할 수 있게 한다. 이 특성으로 의사는 모델의 결과에 대한 신뢰를 갖게 되고 해당 부분을 환자와 공유함으로써 환자 역시 본인의 문제에 대해 직관적으로 이해할 수 있게 된다.

더 나아가서는 해당 모델이 정확성과 신뢰성 모두를 보장해, 의료시설이 미흡한 곳에서도 심전도의 측정과 분석이 용이해지고, 보편적 인간의 건강 증진으로 이어질 수 있다.

5. 참고문헌

- [1] Zhou, B., Liu, S., Hooi, B., Cheng, X. and Ye, J., 2019, August. BeatGAN: Anomalous Rhythm Detection using Adversarially Generated Time Series. In *IJCAI* (pp. 4433-4439)
- [2] Jeyakumar, J.V., Noor, J., Cheng, Y.H., Garcia, L. and Srivastava, M., 2020. How can i explain this to you? an empirical study of deep neural network explanation methods.
- [3] Kuznetsov, V.V., Moskalenko, V.A., Griбанov, D.V. and Zolotykh, N.Y., 2021. Interpretable Feature Generation in ECG Using a Variational Autoencoder. *Frontiers in Genetics*, 12.
- [4] Golany, T., Radinsky, K. and Freedman, D., 2020, November. Simgans: Simulator-based generative adversarial networks for ecg synthesis to improve deep ECG classification. In *International Conference on Machine Learning* (pp. 3597-3606). PMLR.
- [5] Ribeiro, A.H., Ribeiro, M.H., Paixão, G.M.M. et al. Automatic diagnosis of the 12-lead ECG using a deep neural network. *Nat Commun* 11, 1760 (2020).