**8조 최종 보고서**

**[1] 예비보고서**

**1. 프로젝트 제목**

* DNN 모델에 근거한 심장병 유무 예측

**2. 프로젝트 시작 계기**

* 세계 보건 기구 (WHO)에 따르면 주요 사망 순위 (2019년 연말기준) 1위를 달성한 심장병은 현재에도 사망률이 매우 높은 질환으로 다양한 원인에 의해 발생한다. 따라서 환자 스스로가 자신이 해당 병에 걸릴 요인과 확률을 알고 대처 하기에는 매우 어려운 질병이다. 또한 심장병에 대한 정확한 예측을 위해선, 훨씬 더 복잡한 데이터와 상호관계 분석이 필요하다. 따라서 magic box인 DNN을 이용한다면 이와 같은 문제를 해결할 수 있을 것같아 이 주제를 선정하였다.

**3. 프로젝트 개요**

* kaggle에서 303명 환자기록으로 부터 얻은 14가지 속성 데이터를 사용한다.

           예) ​​나이, 성별, 가슴 통증 유형, 혈청 콜레스테롤,​​형광 투시된 주요 혈관의 수 등등

           \* 분석데이터 경로→ <https://www.kaggle.com/ronitf/heart-disease-uci>

* 그 후 데이터의 타입을 하나로 맞추어 계산에 용이하기 위해   raw 데이터와 모델링하는 데이터'를 연결짓는 브릿지 역할로 feature\_columns = []를 생성 하여  모든 feature들을 숫자로 처리한다.
* 예) Age 와 같은 특정 Bucketized column은 수치값을 구간을 나누어 범주형으로 변환, 성별, 가슴통증 유형과 같은 Categorical column은 특정 문자열을 수치형으로 매핑하여 전달.. 등등
* 그 후 전체데이터를 4:1 의 비율로Training set과 Test set를 나눠서 모델에 적용 하여 학습 .
* 복잡한  데이터들로부터 여러 속성들 간의 관계를 통합적으로 파악해야 하고 다층의 은닉층으로부터 계산하는 과정이 필요하기 때문에 DNN(Deep Neuron Network) 모델을 사용한다.
* 하이퍼 파라미터를 여러 조합을 사용하면서 조정하고, 다양한 optimizer 기법들을 적용해 보면서 최적의 값을 찾는다.

**4. 기대효과**

* 나이,성별, 콜레스테롤 수치, 혈당과 같은 기본적인 데이터 뿐만 아니라 안정 심전도 결과, 형광 투시 된 주요 혈관의 수와 같은 전문적인 의료데이터를 접목시켜 보다 정확한 심장병 유무의 예측을 진행 할 수 있다.
* 주어진 모델을 사용하는 것에서 벗어나 여러 최적화 기법들을 사용해 모델의 성능을 개선시킴으로써 기존에 제시된 모델의 성능보다 더 향상된 결과를 얻을 수 있다.
* 수업 시간에 배웠던 내용에서 추가적으로 직접 새로운 데이터를 가지고 응용해보고, activation funcion이나 complie에 필요한 optimizer, loss function 등 여러 요소를 다르게 하는 것이 예측에 어떠한 영향을 미치는지 파악할 수 있다.

**[2]. 프로젝트 작업파일**

프로젝트 작업 파일 : [Team\_8.zip](/redirect/s3?bucket=uploads&prefix=paste%2Fkec8y8is6ly2gf%2F34c4de8d869c96396f13336058febb3a1e9663541ba0a8dd9291c8c5564c3587%2FTeam_8.zip)

Source Code : <https://bioinformaticsandme.tistory.com/m/262>

<https://githubplus.com/themrzmaster/keras-tuner>

< 프로젝트 작업 zip 파일 세부 항목 >

* Code : [team8\_code.ipynb](/redirect/s3?bucket=uploads&prefix=paste%2Fkec8y8is6ly2gf%2F4786a817884077abaa47ac447b3ccceef1d304255ef43c33a081b2abf274607e%2Fteam8_code.ipynb)
* PPT :  [Team8\_ppt.pptx](/redirect/s3?bucket=uploads&prefix=paste%2Fkec8y8is6ly2gf%2F09500df0623f730e9d4ea49681c6e35bf7a35d57f5b2c86bbe7ba730ee60c3c4%2F_________8___ppt.pptx)
* Data : [heart.csv](/redirect/s3?bucket=uploads&prefix=paste%2Fkec8y8is6ly2gf%2F7c3014365675306819510a49ff289efbec1d1a6a666a2dc7652f1547b383d859%2Fheart.csv)

           \*  Data 다운 경로 : <https://www.kaggle.com/ronitf/heart-disease-uci>

**[3] 동영상 발표**

* 동영상 링크

<https://youtu.be/lCyJymF7Udo>