

건강한 성인에서 피검사 결과와 환경적 요인에 의한 중증 우울도 데이터 예측

지민준*, 박현희^o

Prediction of Serious Depressive Symptoms by Blood Test and Environmental Factor in Adult Men and Women

Minjun Ji*, Hyunhee Park^o

요 약

기본적인 신체적 검진을 통해서서는 정신적인 위험 요소를 파악할 수 없고, 심각한 우울증이 의심되는 개인조차 정신적 검진을 신체적 검진만큼 시행하지 않고 있기 때문에 기본적인 신체적 검진만으로 정신적 질환에 대한 구체적인 예측과 분석을 시행할 필요가 있다. 따라서 본 논문에서는 신체적 요소와 개개인이 가질 수 있는 환경적 요인으로 인하여 중증 우울도를 예측할 수 있는 모델을 만들고 이에 대한 정확도, 민감도, 특이도에 대한 분석을 시행한다. 특히, 중증 우울도 예측을 위해 neural network을 활용하기 위하여 Artificial Neural Network (ANN) 모델과 Deep Neural Network (DNN) 모델을 사용하여 그 결과를 비교한다. 다양한 최적화 기법을 적용하여 ANN 모델과 DNN 모델을 사용한 결과를 비교해 보았을 때, ANN 모델의 중증 우울도 예측 정확도는 72.27%, DNN 모델의 중증 우울도 예측 정확도는 76.08%를 보임을 알 수 있다. 또한 두 모델에 대한 Area Under Curve (AUC) 결과를 비교함으로써 중증 우울도 예측을 위한 머신러닝 알고리즘의 예측 정확도를 구체적으로 제시한다.

Key Words : Machine learning, ANN, DNN, PHQ-9, depression

ABSTRACT

Severe depression is one of the mental illnesses that attract attention as a serious social issue. These mental illness may be related to physical illness or environmental factors. In this paper, we perform a modeling to predict the severity of depression due to the individual physical factors and the environmental factors, analyze the accuracy, sensitivity, and specificity of the proposed models. In particular, we use the Artificial Neural Network (ANN) and Deep Neural Network (DNN) models of the neural networks for predicting the severity of depression. Comparing the accuracy results of ANN and DNN models using various optimizers, the predictive accuracy of the ANN model is 72.27% and the predictive accuracy of the DNN model is 76.08%, respectively. The accuracy of prediction of machine learning algorithm for predicting severe depression is presented by comparing the Area Under Curve (AUC) values for proposed models.

* 본 연구는 한국연구재단 논문연구과제 (2017R1C1B5017556) 지원으로 수행되었습니다.

• First Author : (ORCID:0000-0002-4168-7895)e-Glue Communications, inciojs@gmail.com, 정회원

o Corresponding Author : (ORCID:0000-0003-3810-7367)Korean Bible University Department of Computer Software, parkhyunhee@gmail.com, 종신회원

논문번호 : 201808-242-C-RN, Received August 13, 2018; Revised August 17, 2018; Accepted August 17, 2018

I. 서 론

우울증은 만성질환에 영향을 미칠 수 있으며, 이는 정신적 장애가 생리학적으로 연관성이 있다는 가능성을 뒷받침 할 수 있는 증거라고 할 수 있다¹⁾. 이를 통해 우리는 정신적 문제와 신체적인 연관성으로 보아 정신적 장애는 결국 심혈관계 질환을 비롯하여 다양한 신체적 질환을 유발할 수 있음을 유추할 수 있다. 예를 들어, 혈중 콜레스테롤 농도와 관련 있는 것으로 보고되는 대표적인 정신과적 질병 중 하나가 우울증이다. 우울증은 중요한 정신과적 요소 중 하나이며, 자살과 같이 사회에 미치는 영향이 매우 크기 때문에 이에 대한 심도 있는 연구가 필요한 것으로 보인다. 혈중 콜레스테롤 농도가 우울증에 미치는 영향에 대한 연구들을 살펴보면 우울증 환자들의 총콜레스테롤 농도와 저밀도 지단백 콜레스테롤의 농도가 낮은 것이 우울 증상과 유의한 상관관계가 있었다는 보고가 있으며, 우울증 환자 중 자살 신고가 있는 환자들의 총콜레스테롤의 양이 유의하게 적다는 분석도 있다^{2,3)}. 이러한 연구 결과를 토대로 콜레스테롤 뿐 아니라 다양한 피검사 및 신체 활동의 여러 요인들이 우울증에 어떤 영향을 미치는 것인지에 대한 분석이 필요하다⁴⁾.

다시 말해, 정신적인 질환이 의심되는 개인이 병원에 내원하여 위와 같은 정신질환 관련 검사를 시행할 수도 있으나, 많은 사람들이 이러한 정신적 질환에 대한 검사를 자발적으로 시행하지 않는 경우가 대다수이다. 1년에 1회 정도의 건강검진을 시행하는 경우에도 정신질환에 대한 검사는 제외되어 있으며 신체적인 검사만을 시행하는 경우가 대부분이다. 이러한 실정 때문에 정신질환에 대한 검사는 대중화되어 있지 않고 질환이 의심되는 중증 우울증이 의심되는 환자들만이 이러한 검사를 진행하는 경우가 많다. 따라서 대부분의 사람들이 정기적으로 시행하는 신체적 건강검진을 통해 정신적인 질환에 대하여 검사를 실시할 필요가 있다. 본 논문에서는 이러한 이유로 인하여 신체적 검사의 일부 요소를 추출하여 정신 질환의 점수를 예측하는 모델을 개발하고자 한다. 특히 중증 우울증에 대한 연구는 보다 시급한 사안으로, 본 논문에서는 이러한 배경을 기반으로 피검사 결과의 일부와 신체적 활동 요소들의 일부의 조합을 바탕으로 이러한 요소들이 중증 우울증에 어떤 영향을 미치는지에 대한 분석을 시행한다.

우울증의 조기 발견을 위해 사용되는 판별 도구로 Patient Health Questionnaire (PHQ)는 1999년에

Spitzer 등이 개발하였고 일차적 임상 장면에서 접하기 쉬운 정신질환들을 감지하고 진단에 도움을 주기 위해 만들어진 자기 보고식 질문지로 주요우울장애의 진단을 위해 9가지 문항으로 이루어져 있다^{5,6)}. 이에 따라 본 논문에서는 우울증을 측정하기 위한 도구로써 PHQ-9의 결과와 신체적 요소들의 연관성에 의하여 우울증을 예측하는 모델을 제시한다^{7,8)}. 이에 따라 머신러닝 알고리즘을 기반으로 한 다양한 regression 알고리즘과 더불어 딥러닝을 적용한 neural networks 모델을 적용하여 PHQ-9의 결과를 예측하는 것을 제안한다.

본 연구는 2007년부터 2014년도의 미국 국민건강영양조사 National Health and Nutrition Examination Survey (NHANES) 데이터 셋을 활용하였으며, 결측치를 제외한 전체 데이터 셋 8301개를 활용하였다. 기본적으로 피검사만으로 우울증을 예측하는 모델을 만든 결과 매우 낮은 우울증 예측 정확도를 보이기 때문에 본 연구에서는 다른 요소를 넣어 예측 정확도를 높이고자 하였다. 미국 국민건강영양조사의 데이터 셋을 조사해 보면 개개인의 수입에 관한 데이터 셋이 존재한다. 따라서 피검사 결과와 같은 신체적 요소에 개인의 환경을 일정 수준 대변할 수 있는 개인 수입 요소를 함께 넣어 모델링을 시도하였다. 또한 본 연구에서는 일반 특성에 따라 중증 우울증의 유병률에 차이가 있는지 분석하였으며, 중증 우울증 여부에 따른 피검사적 요소들과 생활 환경적 요소들에 대한 상관관계를 측정하여 분석하였다.

II. 중증 우울증 예측을 위한 연구방법

우울증 판별 도구 중 그 결과가 입증된 유의한 조사 중 하나는 9개 항목의 질문을 통해 우울 증상을 판단할 수 있도록 개발된 PHQ-9가 있다. 그림 1은 환자 건강 질문지에 대한 파일이다. 각 항목에 대해 ‘전혀 아니다’, ‘여러 날 동안’, ‘1주일 이상’, ‘거의 매일’의 네 단계로 지난 2주간 느낀 정도를 응답하도록 하고 0점 (전혀 아니다) ~ 27점 (매일)으로 점수화하였다. 표 1에 나타낸 바와 같이 0점에서 27점까지의 점수에서 0~4점 사이의 점수 결과의 경우 우울증이 아닌 정상 상태를 나타내며, 5~9점 사이의 점수 결과는 가벼운 우울증상으로 구분된다. 0~9점까지의 결과는 중증 우울이라 진단하지 않을 정도로 표현할 수 있다. 10~19점 사이의 점수 결과는 중간 정도의 우울 상태로 분류되고, 20~27점 사이의 점수 결과는 심각한 우울증으로 분류된다⁶⁾. 10~27점까지의 점수는 중증 우

환자 건강 질문지 -9 (PHQ-9)

지난 2 주일 동안 당신은 다음의 문제들로 인해서 얼마나 자주 방해받았습니까?
(☑으로 답을 나타내시오)

| | 전혀 방해 받지 않았다 | 어질 때 때 | 7 일 이상 때 때 | 거의 매일 때 때 |
|---|--------------------|--------------|---------------------|--------------------|
| 1. 일 또는 여가 활동을 하는 데 얼마나 즐거움을 느끼지 못한 | 0 | 1 | 2 | 3 |
| 2. 기분이 가라앉거나, 우울하거나, 희망이 없음 | 0 | 1 | 2 | 3 |
| 3. 잠이 들거나 계속 잠을 자는 것이 어려우, 또는 잠을 너무 많이 잠 | 0 | 1 | 2 | 3 |
| 4. 피곤하다고 느끼거나 기운이 거의 없음 | 0 | 1 | 2 | 3 |
| 5. 믿음이 없거나 과식을 함 | 0 | 1 | 2 | 3 |
| 6. 자신을 부정적으로 봄 - 혹은 자신이 실망자라고 느끼거나 자신 또는 가족을 실망시킴 | 0 | 1 | 2 | 3 |
| 7. 신문을 읽거나 텔레비전 보는 것과 같은 일에 집중하는 것이 어려움 | 0 | 1 | 2 | 3 |
| 8. 다른 사람들이 주목할 정도로 너무 느리게 움직이거나 말을 함, 또는 반대로 평소보다 많이 움직여서, 너무 안절부절 못하거나 울던 있음 | 0 | 1 | 2 | 3 |
| 9. 자신이 죽는 것이 더 낫다고 생각하거나 어떤 식으로든 자신을 해할 것이라고 생각함 | 0 | 1 | 2 | 3 |

FOR OFFICE CODING 0 + _____ + _____ + _____
=Total Score: _____

만일 당신이 위의 문제 중 하나 이상 "예" 라고 응답하셨으면, 이러한 문제들로 인해서 당신은 일을 하거나
가정일을 돌보거나 다른 사람과 어울리는 것이 얼마나 어려웠습니까?

| | | | |
|--------------------------|--------------------------|--------------------------|--------------------------|
| 전혀 어렵지 않았다 | 약간 어려웠다 | 많이 어려웠다 | 매우 많이 어려웠다 |
| <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |

그림 1. PHQ-9의 건강 설문지.
Fig. 1. PHQ-9 Questionnaire.

표 1. PHQ-9의 결과에 대한 카테고리 분류
Table 1. Scoring categories for PHQ-9 results

| Outlined Depression Severity | PHQ-9 Score |
|---------------------------------|-------------------|
| 1-4 | None |
| 5-9 | Mild |
| 10-19 | Moderately severe |
| 20-17 | Severe |

우도라 판정할 수 있는 정도의 점수라 볼 수 있다. 미국의 경우 전체 인구의 약 8%가 우울증을 앓는 것으로 보고되어 있다. 본 논문에서는 PHQ-9의 측정결과 10점 이상의 중증 우울증에 대한 예측을 위한 모델링을 수행한다.

본 논문에서는 PHQ-9의 측정결과 10점 이상의 중증 우울도에 대한 예측을 위한 모델링을 수행한다. 이를 위해 고려한 input 요소로는, 생활 환경적 요소로써 14가지로 세분화한 자신의 수입과 가족의 수입을 대표적인 요소로 선택하였고, 신체활동과 연관성이 높은 요소 중 하나인 Vitamin D의 요소를 3가지로 세분화하여 선택하였다. 또한 신체 활동과 식이요소가 함께 고려될 수 있는 자가 몸무게를 3가지로 세분화하여 선택하였다. 그리고 신체적 요소를 위하여 성별과

나이 뿐 아니라 피검사 데이터의 일부를 선택하였는데, 이는 BMI, 당화혈색소, HDL 콜레스테롤, 몸무게, LDL 콜레스테롤, WBC 백혈구 수, 혈압, 맥박, 총 콜레스테롤, Hemoglobin (g/dL)를 선택하였다. 최종적으로 예측을 위한 output으로는 PHQ-9의 점수 10점을 cutoff로 하여 10 이상의 중증 우울도를 예측하는 예측 모델에 따른 정확도 결과를 보인다.

III. 우울증 예측을 위한 머신러닝 알고리즘

본 논문에서는 기본적으로 우울증 예측을 위한 머신러닝 모델링을 수행하기 위하여 Optimizer로 Adam, Adagrad, Momentum, RMSProp, GradientDescent 최적화 알고리즘을 사용하여 결과를 분석하였고 Learning rate는 0.0001 ~ 1 까지 적용한 후 가장 최상의 결과를 내는 모델을 사용하였다. 기본적으로 머신러닝 모델링을 수행하기 위하여 hypothesis, loss function (혹은 cost function), optimizer를 적용한다. 기본적으로 머신러닝에서 최적화 알고리즘은 데이터를 학습할 때 실제 결과와 예측 결과 값의 차이를 최소화하는 방법이다. 일반적으로 gradient descent optimizer가 기반인 다양한 최적화 알고리즘들이 존재한다. 그 중에서 본 논문에서는 Adam 알고리즘을 최적화 기법에 사용하고 그 결과를 분석한다. 또한 learning rate는 loss function이 minimum이 되는 최적의 해를 찾아가는 과정을 반복하는 방식이다. Learning rate가 너무 크게 되면, 최적의 값으로 수렴하지 않고 발산하는 overshooting이 발생하게 되고, learning rate가 너무 작게 되면 수렴하는 속도가 너무 느리고 local minimum에 빠질 확률이 증가하게 된다⁷⁾. 따라서 본 논문에서는 적합한 learning rate를 찾아 그 결과를 제시하고 분석한다. 머신러닝 알고리즘을 적용하기 위하여 Scikit-learn을 사용하여 regression 알고리즘을 적용한 결과 예측값이 - 값으로 수렴하는 것을 확인하였다. 이는 input으로 적용한 의료 데이터의 경우 regression과 같은 회귀분석으로는 모델링이 되지 않는 것을 의미한다. 결국 본 논문에서는 이러한 시행착오 후 TensorFlow를 사용하여 neural network 알고리즘을 사용한 모델링을 수행하였다. 특성상 연속적인 raw signal에 강한 특성을 보이는 Neural network는 여러 개의 perceptron으로 하나의 layer를 구성하며 그 layer로 여러 개의 층을 구성하여 network를 구성하는 방법이다⁸⁾. 이는 Multi Layered Perceptrons (MLP)라고 불리고, 훈련된 weight와 bias를 이용하면 non-linear한

문제를 풀 수 있기 때문에 본 논문에서는 neural network 알고리즘을 사용하여 모델링을 수행한다. MLP로 정의되었으나 hidden layer가 1개뿐인 artificial neural network (ANN) 기법과 여러 개의 hidden layer를 적용할수록 높은 정확도를 보일 수 있는 deep neural network (DNN) 기법을 함께 적용하여 중증 우울도 예측의 정확도를 분석한다^{9,10)}. 기본적으로 ANN과 DNN은 perceptron에 트레이닝 데이터 셋을 input으로 넣어 전방향 연산(forward propagation)을 수행한다. 이때 결과로 나온 Neural networks의 prediction의 값과 target value의 값 차이로 에러 값으로 계산하게 되고 이를 back propagation (chain rule) 알고리즘의 적용이라 할 수 있다. MLP에서는 hidden layer를 증가시킴으로써 정확도가 향상될 것을 기대하였으나, back propagation에서 이전 기울기 값을 잃어버리는 vanishing gradient problem이 발생할 수 있다. 이를 위해 RELU 함수를 적용하며 vanishing gradient problem을 해결한다¹¹⁾.

IV. 모델링 적용 및 실험 결과

4.1 모델링을 위한 데이터 전처리

본 논문에서 적용하고자 하는 의료 데이터 상에서 우울증을 예측하기 위한 모델의 성능 평가를 위해 실험한 환경은 표 2와 같다. 또한 본 연구에서는 의료 데이터 중에서도 우울증과 연관된 데이터를 사용하여 환경적 요소와 신체적 요소가 어떤 영향을 미치는지에 대한 모델링을 수행하였다. 특히 이를 위하여 미국 국민건강영양조사 NHANES 데이터의 2007년부터 2014년도까지의 자료를 사용하였다. 미국 국민건강영양조사 NHANES 의료 데이터의 데이터 셋은 약 8천 300여 개의 데이터로 구성되고, 이 중 training data를 위하여 80%를 사용하고 test data를 위하여 20%를 사용한다. 입력 데이터에 대한 구체적 제시를 위하여 개인 수입 요소를 나타내는 코드북을 그림 2에 제시한

표 2. 시뮬레이션 환경
Table 2. Simulation environment

| Parameter | Value |
|-----------|-----------------------------------|
| OS | Windows 10 Pro |
| CPU | Intel Core i5-8400 2.80GHz 64bits |
| GPU | GeForce GTX1070 8GB |
| Memory | 16GB |
| Language | Python 3.6 |
| Library | Tensor Flow 1.9 |

- SEQN - Respondent sequence number
- INQ020 - Income from wages/salaries
- INQ012 - Income from self employment
- INQ030 - Income from Social Security or RR
- INQ060 - Income from other disability pension
- INQ080 - Income from retirement/survivor pension
- INQ090 - Income from Supplemental Security Income
- INQ132 - Income from state/county cash assistance
- INQ140 - Income from interest/dividends or rental
- INQ150 - Income from other sources
- IND235 - Monthly family income
- INDFMMP1 - Family monthly poverty level index
- INDFMMP2 - Family monthly poverty level category
- INQBOX1 - CHECK ITEM
- INQ244 - Family has savings more than \$5000
- IND247 - Total savings/cash assets for the family

그림 2. 개인 수입요소를 나타내는 코드북.
Fig. 2. Component description of income source.

다. 그림 2의 개인 수입 요소를 분류하는 코드북을 보면 모든 요소가 모델링을 위한 input 요소를 들어가지는 않는다. input 요소로 포함하기 위한 개인 수입요소 검증에 의해 의사 결정트리를 위한 분류 분석을 수행하는 대표적인 데이터 마이닝 프로그램으로써 (Waikato Environment for Knowledge Analysis) WEKA를 통해 검증하여 feature selection을 수행한 결과 포함되는 input 요소로는 INQ020, INQ012, INQ030, INQ060, INQ080, INQ090, INQ132, INQ140, INQ150, INQ235이다. 이 중에서 그림 3에서 나타내는 INQ020 (Income from wages/salaries)은

Variable Name: INQ020
SAS Label: Income from wages/salaries
English Text: The next questions are about (your/your combined family) income. When answering these questions, please remember that by {"income/combined family income"}, I mean (your income/your income plus the income of {NAMES OF OTHER NHANES FAMILY MEMBERS} for {LAST CALENDAR YEAR}). Did (you/you and OTHER NHANES FAMILY MEMBERS 16+) receive income in (LAST CALENDAR YEAR) from wages and salaries? [Did (you/you or OTHER FAMILY MEMBERS 16+) get paid for work in (LAST CALENDAR YEAR).]
Target: Both males and females 0 YEARS - 150 YEARS

| Code or Value | Value Description | Count | Cumulative | Skip to Item |
|---------------|-------------------|-------|------------|--------------|
| 1 | Yes | 8142 | 8142 | |
| 2 | No | 1856 | 9998 | |
| 7 | Refused | 48 | 10046 | |
| 9 | Don't know | 6 | 10052 | |
| . | Missing | 123 | 10175 | |

그림 3. 우울증 예측을 위한 개인 수입요소의 값.
Fig. 3. Income from wages/salaries values for prediction model of PHQ-9.

개인의 수입을 대표하는 값이다. 그림 3은 2013년~2014년간의 INQ020의 값을 대표적으로 보여주고 있고, 질문지를 보면 지난 해 본인 혹은 16세 이상의 가족이 직장으로부터 월급을 받았는지에 대하여 묻고 있다. 이에 대한 응답으로 1인 경우 Yes, 2인 경우 No 라고 대답하게 된다. 1년간의 데이터를 보면 8142명이 수입이 있었고, 1856명이 개인 혹은 본인의 가족이 수입이 없는 것으로 응답했다. 특히, INQ020은 우울 증상 모델링을 위한 input으로 포함되면서 그림 4와 같은 형태로 변형되어 입력된다. 그림 4에서 표현된

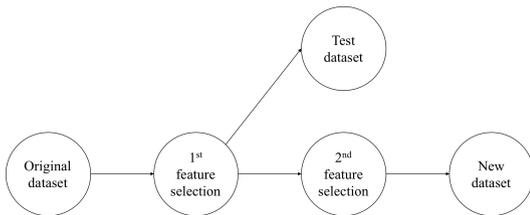


그림 4. Original dataset에서 New dataset까지의 단계.
Fig. 4. Steps from Original dataset to New dataset.

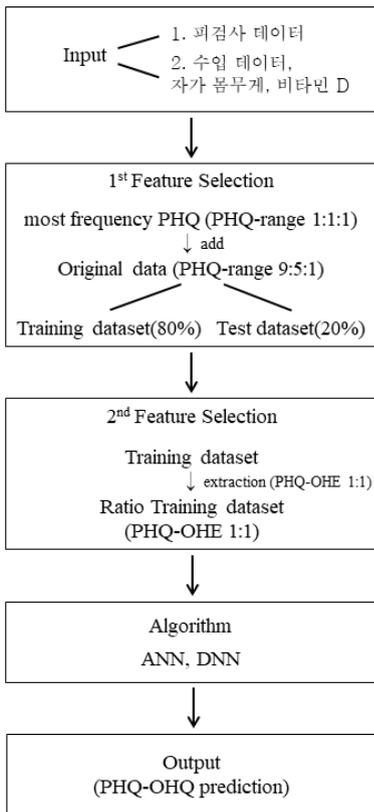


그림 5. 전체적인 모델링의 블록도
Fig. 5. Block diagram of overall modeling

feature engineering의 구체적인 설명은 그림 6과 그림 7을 통해 기술한다.

그림 5의 diagram을 통해 input 데이터가 feature selection을 통해 가공되어 머신러닝 알고리즘에 적용됨으로써 최종 PHQ에 대한 예측을 수행하는 일련의 과정을 알 수 있다. PHQ range 0은 PHQ-9의 측정결과 0점에서 27점까지의 점수에서 0~4점 사이의 점수 결과의 경우 우울증이 아닌 정상 상태를 나타내는 경우로 PHQ-range 0으로 나타내고, 5~9점 사이의 점수 결과는 가벼운 우울증상으로 구분되며 PHQ-range 1로 표현된다. 10~27점까지의 결과는 중간 이상의 우울 상태로 분류되고, 이는 PHQ-range 2로 분류한다. 각각의 PHQ-range의 비율을 살펴보면 PHQ-range 0, 1, 2의 비는 각 약 9:5:1의 비율을 보인다. 그림 6에서

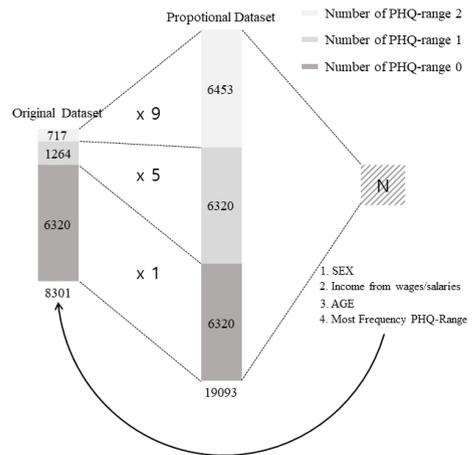


그림 6. 1st feature selection.
Fig. 6. 1st feature selection.

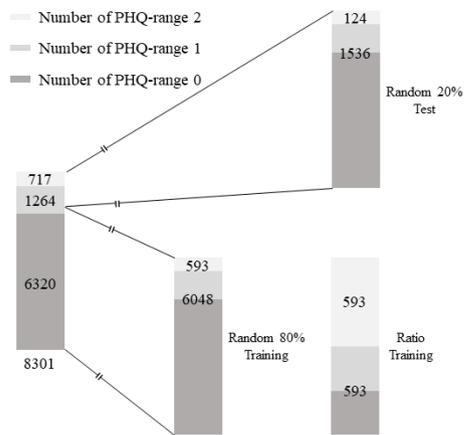


그림 7. 2nd feature selection.
Fig. 7. 2nd feature selection.

나타낸 바와 같이, 1st feature selection의 방법은 각 PHQ-range를 기준으로 0, 1, 2의 값의 x1, x5, x9를 곱해주어 데이터 셋의 규모를 1:1:1로 맞추는 Proportional dataset을 만든 후 Original dataset의 한 사람의 데이터를 성별과, 나이, 수입을 기준으로 Proportional dataset에서 최고 빈도 수의 PHQ-range를 새로운 모델링 값으로 Original dataset에 추가해준다. 이 후 Training dataset과 Test dataset을 8:2의 비율로 나뉜다.

2nd feature selection의 경우 1st feature selection의 결과를 바로 Training dataset으로 사용할 경우 종속변수의 비율이 9:5:1로 맞지 않기 때문에 정확한 예측이 힘들게 된다. 이를 해결하기 위하여 그림 7과 같이, 종속 변수를 중증 우울을 True 그 이외 값을 False로 하여 PHQ_OHE(One - Hot Encoder)를 값으로 하고 True값 개수의 맞춰 Training dataset의 비를 1:1로 구성하는 부분이다. 최종적으로 2nd feature selection까지 끝난 Training dataset을 New dataset이라 표현한다.

모델링을 수행하기 위하여 입력 데이터를 가공하여 사용한다. 기본적으로 입력으로 들어가는 데이터 셋은 다음과 같다.

- 성별, 나이, 키, 몸무게, BMI, 당화혈색소, 총 콜레스테롤, LDL 콜레스테롤, HDL 콜레스테롤, WBC 백혈구 수, 최대혈압, 최소혈압, 맥박, 헤모글로빈, PHQ_total, PHQ_range, PHQ_OHE

본 논문에서는 20개 수치 데이터 클래스를 MinMaxScaler에 적용하였으며, 성별, 수입, PHQ 관련 카테고리형 클래스를 One-Hot Encoder를 사용하여 62개로 사용함으로써 총 82개의 클래스 형태의 Input(독립 변수) 데이터를 구성하였으며 PHQ의 결과 값을 CutOff 10을 기준으로 정상과 우울의 구분을 Output(종속 변수)로 구성하였다.

4.2 머신러닝 모델링 알고리즘

ANN과 DNN 알고리즘을 적용하기 위해 동일하게 Input 데이터의 클래스 값을 151로 설정한다. ANN의 경우 이 데이터가 input layer를 통해 첫 번째 hidden layer로 들어갈 때 256가지의 클래스로 나오게 되고 이 후 ANN의 경우 output layer를 거쳐 1가지의 클래스 값으로 나오도록 neural networks를 구성한다. DNN의 경우 input layer를 통해 첫 번째 hidden layer로 들어갈 때 256가지의 클래스로 나온 이후 다시 두 번째 hidden layer와 세 번째 hidden layer를 거친다. (본 연구에서는 hidden layer의 개수를 5개, 7개로 구성한 DNN의 경우 이 과정을 추가적으로 반복한다.)

이후 output layer를 거쳐 1가지의 클래스 값으로 나오도록 neural networks를 구성한다.

Neural network를 이용하여 모델링을 수행하기 위해서는 regularization method들과 initialization method들, 그리고 overfitting을 최대한 피하기 위한 optimizer method들을 고려해야 한다. 본 논문에서는 앞서 기술한 5가지 optimizer를 고려하여 Adam, Adagrad, Momentum, RMSProp, GradientDescent 최적화 알고리즘을 사용한 결과 Adam optimizer가 가장 좋은 성능을 보이는 것으로 분석하였다. 또한 learning rate는 0.0001부터 1까지의 값을 적용한 결과 0.0001일 때 적합한 것으로 분석하였다. DNN을 위한 hidden layer는 3개, 5개, 7개로 분류하여 적용하였고, 이 때의 hidden dimension의 경우 128, 256, 512가 각 hidden layer의 개수에 따라 약간의 다른 결과를 보이지만 결과 비교를 위해 hidden dimension도 128으로 적용한 결과를 분석하였다. 또한 layer가 많아질 때 발생하는 대표적인 문제점 중 하나인 overfitting 문제를 해결하기 위하여 dropout을 통해 feature들의 개수를 조절할 수 있는 방법을 사용한다. 본 논문에서 dropout의 경우 그 결과값이 영향을 미치게 되는 0.8과 0.7 값을 적용한 결과를 분석하였다. 결과적으로 본 논문에서 사용한 최적화 기법은 표 3과 같다.

표 3. 다양한 최적화 기법의 적용
Table 3. Various optimization methods

| Method | Value | |
|------------------|----------|----------|
| | ANN | DNN |
| Optimizer | Adam | Adam |
| Learning rate | 0.0001 | 0.0001 |
| Hidden layer | 1 | 3, 5, 7 |
| Hidden dimension | 128 | 128 |
| Dropout | 0.7, 0.8 | 0.7, 0.8 |

4.3 머신러닝 모델링 알고리즘의 결과

기존의 우울증의 우울도 예측 알고리즘의 경우 우울증 예측을 위한 resilience data와 같은 우울 척도 기반의 예측이 대부분이다. 예를 들어, 우울과 자살 시도와의 연관성을 예측하는 논문의 경우 우울 척도와 연관된 환자가 1년 동안 자살 시도를 한 경우에 대하여 약 90%의 예측 정확도를 보였다^[2]. 이러한 경우는 우울척도와 자살시도와의 correlation score가 0.9 이상의 연관성에 기반한 연구 결과이므로 높은 예측 정확도를 기대할 수 있다. 본 논문의 경우 피검사 요소

및 환경적 요소가 우울증 척도인 PHQ-9과 correlation score가 0.1 이상의 연관성을 보이는 경우가 없는 예측 실험으로 90% 정도의 예측 정확도 검증에는 미치지 못하나, 순수 피검사의 결과와 환경적 요소만으로 우울도를 예측하는 것에 의미가 있으리라고 본다.

그림 8에서는 표 3에서 제시한 최적화 기법을 적용하여 ANN의 결과를 비교 분석한 것이다. 특히 dropout이 0.8인 경우의 결과이다. Adam 최적화 알고리즘을 적용했을 때 값의 변동이 적어도 epoch이 증가함에 따라 정확도가 단계적으로 증가하는 것을 볼 수 있다. 그러나 adam 최적화 알고리즘에 적절한 learning rate가 적용되어야 최적의 결과를 찾을 수 있음을 알 수 있다. 데이터의 정확도를 위해 learning rate가 0.0001 값일 때 가장 정확도가 높기 때문에 learning rate는 0.0001을 사용한 결과를 보여준다.

그림 8을 통해 epoch이 증가함에 따라 정확도의 값이 어떻게 변화하는지의 추이를 알 수 있다. Training accuracy의 경우 epoch이 증가함에 따라 지속적으로 증가하는 것을 알 수 있고, 이와 달리 test accuracy의 경우 약 90% 정도의 epoch 부터는 정확도의 값이 일정 값으로 saturation 되는 것을 알 수 있다. 즉, ANN을 모델링으로 사용하여 중증 우울도 예측을 한 결과 약 70% 초반의 정확도를 보이며 예측한다.

그림 9는 DNN을 사용했을 때의 정확도 예측 값에 대한 결과를 보여준다. 특히, DNN의 모델링을 위하여 hidden layer를 3개로 셋팅하고 dropout을 0.7로 적용했을 때의 DNN 모델링의 결과를 대표적으로 보여준다. 그림 9를 통해 epoch이 증가함에 따라 정확도의 값이 어떻게 변화하는지의 추이를 알 수 있다. Training accuracy의 경우 epoch이 증가함에 따라 지속적으로 증가하는 것을 알 수 있고, 또한 test accuracy의 경우 약 40% 정도의 epoch 부터는 정확도의

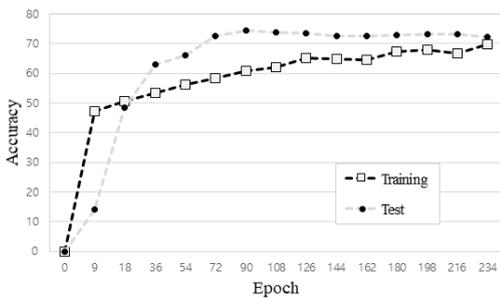


그림 8. 우울증 예측을 위한 ANN 알고리즘의 결과 (Dropout = 0.8)
Fig. 8. Simulation result of ANN algorithm for prediction model of PHQ-9 (Dropout = 0.8)

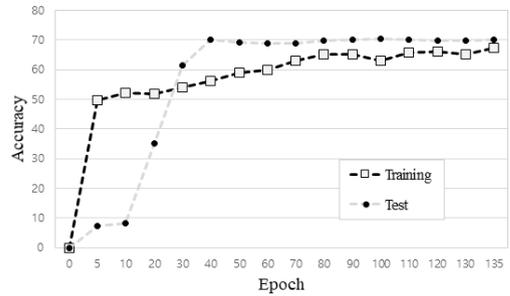


그림 9. 우울증 예측을 위한 DNN 알고리즘의 결과 (Dropout = 0.7, hidden layer = 3)
Fig. 9. Simulation result of DNN algorithm for prediction model of PHQ-9 (Dropout = 0.7, hidden layer = 3)

값이 일정 값으로 saturation 되는 것을 알 수 있다. 즉, ANN을 모델링으로 사용하여 중증 우울도 예측을 한 결과 약 70% 중반의 정확도를 보이며 예측한다. 정확한 결과값에 대한 수치적 비교는 다음 표를 보면서 설명한다.

표 4의 결과에 따르면 ANN의 경우 dropout의 값이 0.7 일 때 71.27%로 0.8의 dropout의 결과보다 약 1% 가량 상승함을 알 수 있다. 이는 hidden layer에 포함되는 많은 수의 뉴런에 의해 상세 모델이 만들어 지면서 임의의 노이즈가 포함됨으로써 overfitting이 발생할 수 있는 것을 dropout을 키워 일부 해결한 것으로 보인다.

결국 그림 10과 표 5를 통해 ANN 모델의 예측 결과와 DNN 모델의 예측 결과를 비교 분석할 수 있다. 대표적인 비교를 위하여 ANN 모델에 dropout 0.8의 값이 적용된 결과와 DNN 모델에 hidden layer 3개에 dropout 0.7의 값이 적용된 값을 비교하였다. 이렇게 특정 모델의 값을 선택하여 비교하는 이유는 ANN 모델과 DNN 모델이 가장 높은 정확도를 보이는 값에서 비교하기 위함이다. 그림 10에서 x축은 ANN 모델과 DNN 모델이 가지는 epoch이 달라 각각의 epoch을 1/10로 축소하여 절대적인 비교를 하기 위한 척도를 적용한 것이다. 그림 10에 의하면 두 모델 모두 epoch이 증가함에 따라 정확도가 일정한 값에서 수렴함을

표 4. 다양한 최적화 기법의 적용의 정확도
Table 4. Accuracy of application of various optimization methods

| Dropout | ANN | DNN (Hidden Layer) | | |
|---------|--------|--------------------|--------|--------|
| | | 3 | 5 | 7 |
| 0.7 | 71.27% | 72.46% | 74.81% | 72.28% |
| 0.8 | 70.24% | 73.19% | 76.08% | 74.93% |

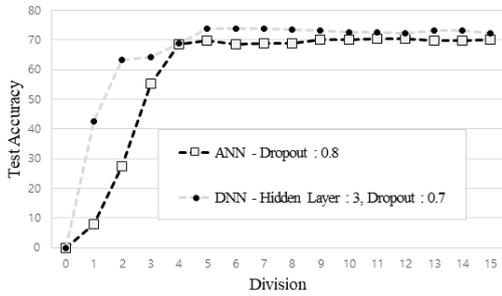


그림 10. 우울증 예측을 위한 ANN 알고리즘과 DNN 알고리즘의 결과 비교
 Fig. 10. Comparison of ANN and DNN algorithms for predicting depression

표 5. 그림 10에 대한 Training accuracy와 Test accuracy의 결과
 Table 5. Training and test accuracy results in Figures 10

| | | Training | Test |
|-----|----------|----------|--------|
| ANN | accuracy | 67.45% | 70.24% |
| | cost | 0.6080 | 0.6108 |
| DNN | accuracy | 69.90% | 72.41% |
| | cost | 0.5957 | 0.6115 |

알 수 있다. 또한 test accuracy의 경우 ANN 모델보다 DNN 모델이 조금 더 높은 값을 가짐을 알 수 있다. 표 5의 결과에 따라 ANN 모델과 DNN 모델을 비교해 보면, 두 모델의 cost 값이 약 0.61로 큰 차이를 보이지 않는 것을 알 수 있다. 즉, cost 값의 유사도의 의해 ANN 모델보다는 정확도가 조금 더 높은 DNN 모델을 사용하는 것이 권장된다.

의료 데이터의 분석을 위하여 본 논문에서는 area under curve (AUC)의 결과를 비교한다. True Positive의 비율은 중증 우울도를 가진 사람을 중증 우울도라고 정확히 예측한 결과이고, False Negative의 비율은 중증 우울증이 아닌 사람을 중증 우울증이라고 잘못 예측한 결과이다. 또한 본 논문에서는 신체적 수치와 환경적 수치가 직접적으로 중증 우울증상에 영향을 미치지 않기 때문에 정확도 측면에서 이주 우수한 예측율을 보이지는 않는다. 그래서 약간의 오류를 허용했을 때 성능이 어떻게 좋아지는지를 평가하기 위한 값으로 민감도 (sensitivity)와 특이도 (specificity)를 측정하였다. 본 연구에서 민감도는 실제로 중증 우울증이 있는 사람을 중증 우울증이라고 정확히 판단하는 비율이며, 특이도는 중증 우울도가 없는 사람을 검사하여 정상이라고 정확히 판단하는 비율이다. 표 6의 결과에 따르면 ANN의 True

Negative와 DNN의 True Negative의 값을 비교해 볼 때 DNN으로 모델링을 수행한 경우 좀 더 중증 우울도를 맞추는 예측이 정확함을 알 수 있다. 또한 표 7에 의해서 민감도와 특이도 모두 ANN 모델보다는 DNN 모델이 좀 더 예측 정확도가 높은 결과를 가짐을 알 수 있다. 이는 각 모델의 AUC 결과 값에 의해 정확히 나타난다.

실제로는 중증 우울증이 아니지만 중증 우울증이라고 잘못 예측하는 지점을 좀 더 느슨하게 판단하여 민감도를 계산하게 된다면 이에 따라 좀 더 높은 민감도를 달성할 수 있을 것이라 예상된다.

표 6. 그림 10에 대한 True/False-Positive/ Negative의 수치 결과
 Table 6. True/False-Positive/Negative values for Figure 10

| | ANN | DNN |
|----------------|---------------|---------------|
| True-Positive | 77 (4.64%) | 81 (4.88%) |
| True-Negative | 1080 (65.60%) | 1121 (67.53%) |
| False-Positive | 447 (26.93%) | 415 (25%) |
| False-Negative | 47 (2.83%) | 43 (2.59%) |

표 7. 그림 10에 대한 민감도, 특이도 및 AUC 결과
 Table 7. Results of Sensitivity, Specificity, and AUC for Figure 10

| | ANN | DNN |
|---------------------------|------|------|
| Sensitivity | 0.62 | 0.65 |
| Specificity | 0.71 | 0.73 |
| Positive predictive value | 0.15 | 0.16 |
| Negative predictive value | 0.96 | 0.96 |
| AUC | 0.66 | 0.69 |

V. 결 론

본 논문에서는 신체적 요소의 기본 검사인 피검사 데이터를 입력 데이터로 추출하여 정신적인 요소에 대한 결과 분석지 중 하나인 PHQ-9의 점수를 예측하는 모델링을 수행하였다. 기본적인 정기 검진을 통해서 정신질환에 대한 검진을 시행할 수 없기 때문에 기본적인 신체적 검진만으로 정신적 질환에 대한 예측을 수행하는 것은 중요한 분석이라 판단된다. 특히, 중증 우울증이 의심되는 개인조차 정신적 검진을 시행하지 않는 경우 신체적 건강에 심각한 위험을 초래할 수 있기 때문에 정신질환은 사회적으로 중요한 이슈로 부각되고 있다. 뿐만 아니라 개인의 수입요소는

개인의 환경을 대변할 수 있는 중요한 요소 중 하나로 입력으로써 함께 고려되었다.

결론적으로 본 논문에서는 기본적인 신체적 요소의 일부인 피검사 결과와 개인 수입을 통해 정신과적 요소의 연관성을 분석하여 중증 우울 증상을 예측하는 모델을 만들고 비교 분석을 수행했다. 특히 ANN 알고리즘과 DNN 알고리즘을 적용하여 중증 우울증을 예측하는 정확도를 판별하고, AUC결과를 기반으로 민감도와 특이도를 판단하는 결과를 분석하였다. ROC 그래프를 기반으로 민감도와 특이도를 판단하는 결과를 분석하였다. 본 논문의 실험 결과에 의하면 ANN 알고리즘에 의한 PHQ-9의 예측 정확도는 70.24%이고, 민감도는 0.62, 특이도는 0.71이고, DNN 알고리즘에 의한 PHQ-9의 예측 정확도는 72.41%이고, 민감도는 0.65, 특이도는 0.73이다. 결국 본 논문에서 제안한 모델링 기법과 머신러닝 알고리즘을 적용하여 정신과적 요소를 제외하고도 개인의 PHQ-9의 점수를 70% 이상 예측함으로써 기본 건강 검진을 통해 정신적 질환을 보다 쉽게 예측하고 증세에 보다 빠르게 대응할 수 있을 것이라 기대된다.

References

- [1] S. H. Lee, M. Park, D. H. Yoon, Y. Lee, and S. S. Kim, "Gender difference in associations between serum cholesterol levels and depression symptoms in healthy general population," *Korean J. Psychosomatic Med.*, vol. 25, no. 1, pp. 27-32, 2017.
- [2] F. M. Sahebzamani, R. F. D'Aoust, D. Friedrich, A. N. Aiyer, S. E. Reis, and K. E. Kip, "Relationship among low cholesterol levels, depressive symptoms, aggression, hostility, and cynicism," *J. Clin. Lipidol.*, vol. 7, pp. 208-216, 2013.
- [3] I. J. Schatz, K. Masaki, K. Yano, R. Chen, B. L. Rodriguez, and J. D. Curb, "Cholesterol and all-cause mortality in elderly people from the Honolulu heart program: a cohort study," *Lancet*, vol. 358, pp. 351-355, 2001.
- [4] H. S. Lee, "Depression and related risk factors in the elderly with a focused on health habits, mental health, chronic diseases, and nutrient intake status," *J. Korean Diet Assoc.*, vol. 24, no. 2, pp. 169-180, 2018.
- [5] K. Kronke, R. L. Spitzer, and J. B. Williams, "The PHQ-9: validity of a brief depression severity measure," *J. Gen. Intern. Med.*, vol. 16, no. 9, pp. 606-613, 2001.
- [6] S. J. Park, H. R. Choi, J. H. Choi, K. Kim, and J. P. Hong, "Reliability and validity of the Korean version of the patient health questionnaire-9 (PHQ-9)," *Anxiety Mood*, vol. 6, no. 2, pp. 119-124, 2010.
- [7] D. Kingma and J. Ba, "ADAM: a method for stochastic optimization," *Int. Conf. Learning Representations*, San Diego, CA, USA, 2015.
- [8] M. A. Sovierzoski, F. I. Argoud, and F. M. Azevedo, "Evaluation of ANN classifiers during supervised training with ROC analysis and cross validation," *Int. Conf. BioMedical Eng. and Informatics*, pp. 274-278, 2008.
- [9] F. Yan, Y. He, O. Ruwase, and E. Smirni, "Efficient deep neural network serving: Fast and furious," *IEEE Trans. Network and Serv. Management*, vol. 15, no. 1, pp. 112-126, 2018.
- [10] C. Xu, C. Lu, X. Liang, J. Gao, W. Zheng, T. Wang, and S. Yan, "Multi-loss regularized deep neural network," *IEEE Trans. Cir. and Syst. for Video Technol.*, vol. 26, no. 12, pp. 2273-2283, 2016.
- [11] Y. Lei, X. Huo, and B. Yan, "Deep neural network for device modeling," *IEEE Electron Devices Technol. Conf.*, pp. 154-156, 2018.
- [12] J. Oh, K. Yun, J. Hwang, and J. Chae, "Classification of suicide attempts through a machine learning algorithm based on multiple systemic psychiatric scales," *Frontiers in Psychiatry*, vol. 192, no. 8, pp. 1-8, 2017.

지 민 준 (Minjun Ji)



2017년 2월 : 한국성서대학교 컴
퓨터소프트웨어학과 졸업
2018년 2월~현재 : 이글루 커뮤니
케이션즈

박 현 희 (Hyunhee Park)



2011년 8월 : 고려대학교 전자
컴퓨터공학과 박사 졸업
2017년 3월~현재 : 한국성서대
학교 컴퓨터소프트웨어학과
조교수
<관심분야> 전자공학, 통신공
학, 머신러닝, 빅데이터