

Original Article

한의 체중 조절 프로그램에 참여한 과체중, 비만 환자에서의 머신러닝 기법을 적용한 체중 감량 예측 연구

김은주¹, 박영배¹, 최가혜^{1,2}, 임영우^{1,2}, 옥지명^{1,2},
노은영³, 송태민⁴, 강지훈⁵, 이향숙⁶, 김서영^{1,2*}

¹누베베 미병연구소, ²누베베 한의원, ³경희대학교 대학원 한의과대학 경락의학과, ⁴삼육대학교 보건관리학과,
⁵한국산업기술대학교 경영학부, ⁶경희대학교 한의과대학 해부학교실

Application of Machine Learning to Predict Weight Loss in Overweight, and Obese Patients on Korean Medicine Weight Management Program

Eunjo Kim¹, Young-Bae Park¹, Kahye Choi^{1,2}, Young-Woo Lim^{1,2}, Ji-Myung Ok^{1,2},
Eun-Young Noh³, Tae Min Song⁴, Jihoon Kang⁵, Hyangsook Lee⁶, Seo-Young Kim^{1,2*}

¹Nubebe Mibyeong Research Institute, ²Nubebe Korean Medical Clinic

³Department of Medical Science of Meridian, College of Korean Medicine, Graduate School, Kyung Hee University

⁴Department of Health Management, Sahmyook University, ⁵Department of Business Administration, Korea Polytechnic University, ⁶Department of Anatomy, College of Korean Medicine, Kyung Hee University

Objectives: The purpose of this study is to predict the weight loss by applying machine learning using real-world clinical data from overweight and obese adults on weight loss program in 4 Korean Medicine obesity clinics.

Methods: From January, 2017 to May, 2019, we collected data from overweight and obese adults (BMI ≥ 23 kg/m²) who registered for a 3-month *Gamitaeumjowi-tang* prescription program. Predictive analysis was conducted at the time of three prescriptions, and the expected reduced rate and reduced weight at the next order of prescription were predicted as binary classification (classification benchmark: highest quartile, median, lowest quartile). For the median, further analysis was conducted after using the variable selection method. The data set for each analysis was 25,988 in the first, 6,304 in the second, and 833 in the third. 5-fold cross validation was used to prevent overfitting.

Results: Prediction accuracy was increased from 1st to 2nd and 3rd analysis. After selecting the variables based on the median, artificial neural network showed the highest accuracy in 1st (54.69%), 2nd (73.52%), and 3rd (81.88%) prediction analysis based on reduced rate. The prediction performance was additionally confirmed through AUC, Random Forest showed the highest in 1st (0.640), 2nd (0.816), and 3rd (0.939) prediction analysis based on reduced weight.

Conclusions: The prediction of weight loss by applying machine learning showed that the accuracy was improved by using the initial weight loss information. There is a possibility that it can be used to screen patients who need intensive intervention when expected weight loss is low.

Key Words : Machine learning, Obesity, Weight loss, Artificial intelligence

• Received : 8 May 2020 • Revised : 22 May 2020 • Accepted : 25 May 2020

• Correspondence to : 김서영

누베베 한의원, 경기도 성남시 분당구 성남대로 926번길 10, 탑빌딩 2층

Tel : +82-31-702-0045, Fax : +82-31-701-8993, E-mail : woori4025@hanmail.net

서론

비만은 에너지 섭취량이 소비량보다 많아서 오는 에너지 불균형 상태가 오랫동안 지속됨으로써 발생하는 질환으로¹⁾ 전세계적으로 유병률이 증가하고 있다²⁾. 비만은 심혈관질환, 당뇨병, 골관절염 등의 발병을 높일 뿐만 아니라 일부 암의 위험을 증가시키는 것으로 알려져 있어^{3,4)}, 비만인의 체중 감량은 건강을 위해 중요하게 인식되어져 왔다.

이에 따라 성공적인 체중 감량을 위해 다양한 선행연구들이 진행되어 왔는데, 지금까지의 연구들은 단순 상관관계나 평균 차 검정, 회귀분석 등을 바탕으로 한 체중감량에 영향을 주는 요인연구가 주를 이루어 왔다⁵⁻¹⁰⁾. 비만도 (초기 체질량지수(Body Mass Index, BMI), 초기 체중)^{5,6)}, 성별^{5,7)}, 연령, 식이요법 준수⁶⁾, 운동⁶⁾, 자기 효능감⁸⁾, 그리고 초기 체중 감량^{5,9,10)} 등을 포함한 많은 요인들이 체중 감량에 영향을 주는 것으로 보고되어 있다.

체중 감량 정도는 중도탈락을 예측할 수 있는 중요한 요인이 되는데¹¹⁾, 비만 프로그램에서 체중 감량을 예측할 수 있다면 환자들이 치료를 완수하는데 도움을 줄 수 있을 것이다. 임상에서 예측되는 체중 감량치가 낮은 환자는 체중 감량을 위한 노력을 더욱 기울일 수 있도록 초기에 독려할 수 있으며, 감량에 대한 높은 기대치로 인해 발생하는 치료 중도탈락의 위험을 미연에 줄일 수 있고, 예측되는 체중 감량치가 높을 경우 자신감을 심어주어 체중 감량에 대한 동기 부여를 할 수 있기 때문이다. 체중 감량에 대한 예측은 전통적으로 선형 회귀의 방식을 사용하여 이루어져 왔다¹²⁾. 선형 회귀분석의 조건(가정)은 매우 엄격한데 대표적으로 표본 가정에서 선형성을 만족시켜야 하지만, 이는 실제 임상에서 만족하기 어려운 경우들이 흔히 존재한다. 머신러닝 기법은 이러한 가정이 만족되지 않을 때 적용해볼 수 있는 좋은 대안이 될 수 있다¹³⁾.

머신러닝 (Machine learning)은 ‘데이터를 수집,

정제, 처리, 분석하고 그로부터 유용한 통찰력을 얻어내는 연구’ 혹은 ‘설명력 있고, 예측 가능한 모델링을 할 뿐만 아니라 통찰력 있고, 흥미롭고, 참신한 패턴을 발견하는 과정’으로, 머신러닝이 기존 전통적인 접근법과 가장 다른 점은 규칙을 프로그래밍 해서 진행되는 것이 아니라 다양한 예시들로 학습을 한다는 점에 있다. 기존 데이터를 통해 높은 적합도로 학습, 훈련이 이루어지면 새로운 데이터가 입력되었을 경우 이를 바탕으로 예측이 가능하게 되고, 입력된 새 데이터로 재학습을 하여 모델의 정확도를 더욱 성장시킬 수 있는 장점을 가지고 있다^{14,15)}.

머신러닝은 이미지 및 음성인식 포함하여 다양한 분야에서 연구가 활발히 진행 중이며, 중앙 및 암 진단^{16,17)}, 심혈관질환 위험도 예측¹⁸⁾, 방사선 판독¹⁹⁾ 등 최근 의학 분야에서의 연구 또한 증가하고 있다. 비만과 관련된 분야에서도 머신러닝을 적용한 시도가 증가하고 있으나²⁰⁻²²⁾ 해외 연구들이 주를 이루고 있다. 국내에서는 비만과 관련된 이러한 연구들이 2010년대 중반부터 비로소 시작되었으며²³⁾, 비만관련 머신러닝 연구는 아직 시작 단계이기 때문에 실제 비만 진료 임상 데이터를 이용한 체중 감량 예측을 진행한 국내 선행 연구는 찾아볼 수 없었다.

선행연구에서 알려진 체중 감량에 영향을 미치는 요인들을 바탕으로 자료를 분석, 정리하여 기존 데이터로부터 특정 값을 예측하는 능력이 뛰어난 머신러닝 기법을 사용하여 체중 감량이 양호하게 이루어질 것인지 아닌지를 판별할 수 있는 예측 모델을 만든다면 임상에서 비만환자들을 진료함에 있어서 초진에서부터 더욱 적절한 조언을 할 수 있고 관리의 강도를 조절해 나가 성공적인 치료 결과를 얻는데 도움을 줄 수 있을 것으로 사료된다.

따라서, 본 연구에서는 한방비만프로그램을 시행한 과체중, 비만 성인 환자들을 대상으로 머신러닝을 적용하여 체중 감량 예측을 연구하고자 하였다. 본 연구의 최종 목표는 체중 감량 예측 모델에 필요한 중요 변수 탐색 및 체중 감량 예측 모델의 비교 분석

으로 이를 위해 크게 두 부분으로 나누어 분석을 진행하였다. 먼저, 선행연구를 바탕으로 본 연구에서 추출 가능한 모든 변수를 사용하여 변수 선택 없이 체중 감량 예측을 진행하였으며, 추가적으로, 중요 변수를 선택하여 이를 바탕으로 체중 감량 예측 모델을 재검토하는 방식으로 진행되었다.

대상 및 방법

1. 연구 대상

이 연구는 다기관, 후향적 연구로서 2017년 1월 4일부터 2019년 5월 30일까지 서울 및 경기도 소재 4개의 N 한방 비만 클리닉에 내원한 환자의 전자 의무기록 데이터를 한차트 (TNH, Gyeonggi-do, Korea)로부터 개인정보를 비식별화된 상태로 제공받아 분석에 사용하였다. 비만 센터에 내원한 모든 환자들은

치료 시작 전 기본적으로 개인정보 제공, 진료 및 연구에 대한 동의서를 작성하였다. 3개월 비만 프로그램에 등록된 66,645명 중에 중도탈락자와 체중 정보가 확인되지 않는 경우, 그리고 연구 방법에 제시된 복용약 이외의 약과 제형을 복용한 경우는 분석에서 제외되었으며, 개인 정보 제공에 동의한 18세 이상의 과체중 및 비만인 (body mass index (BMI) $\geq 23\text{kg/m}^2$)을 분석 대상으로 삼았다. 이 연구 프로토콜은 경희대학교 기관생명윤리위원회 (Institutional Review Board)의 승인을 받아 진행되었으며 (KHSIRB-19-433), 연구 방법에 기반한 Dataset은 1차 예측 분석 25,988명, 2차 예측 분석 6,304명, 3차 예측 분석 833명이다(Figure 1).

2. 연구 방법

가미태음조위탕은 총 3개월간 약 한 달 간격으로

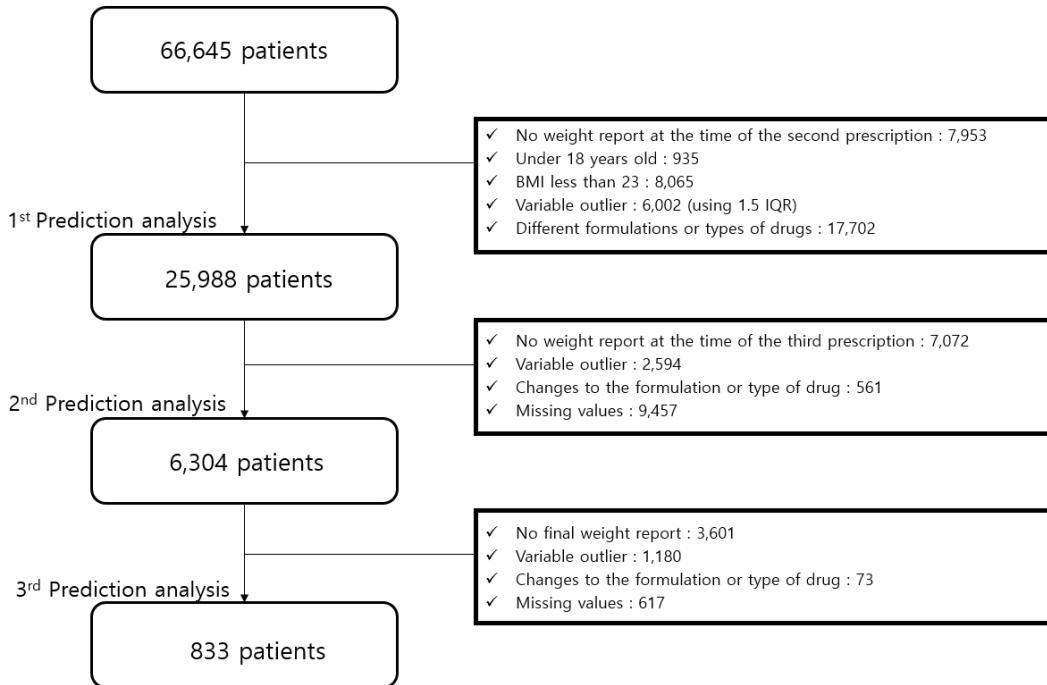


Fig. 1. Flowchart of dataset for analysis

총 3회 투여되었다. 예측 분석은 가미태음조위탕 총 3회의 각 처방 시점에서 이루어졌으며, 그 다음 차수 처방 시 예상되는 체중 감량률 및 감량 체중을 머신러닝 기법을 적용하여 분류의 기준점이 되는 부분을 넘길 것인지, 넘기지 못할 것인지 예측이 이루어지는 이분형으로 예측 (binary classification)하는 방식으로 이루어졌다. 본 연구는 분류의 기준점 및 변수 선택 기법 적용 유무에 따라 크게 두 가지 종류의 분석으로 이루어졌다. 첫 번째 분석에서는 변수 선택 없이 모든 입력변수 (독립변수)를 사용하여 본원에서 가미태음조위탕을 복용한 모집단의 데이터를 바탕으로 체중 감량 정도를 감량률과 감량 체중을 기준으로 각각 4분위 (상위 25%, 중위 50%, 하위 25%)로 구분하여 이를 분류의 기준점을 삼아 예측이 진행되었으며, 이에 사용된 각 모델별 예측 정확도를 비교 분석하였다. 두 번째 분석에서는 분류의 기준을 체중 감량 중위 50%로 두었을 경우를 바탕으로 중요 변수를 선택하여 각 모델별 성능을 추가 분석하였다.

큰 줄기인 두 가지 분석 내에서 각각 1차, 2차, 3차 예측 분석이 세부적으로 시행되었다.

전체적인 데이터 분석의 과정은 데이터 수집, 데이터 정제 및 전처리, 탐색적 데이터 분석, 모델링 기법 선택 및 모델링 수행, 검증 및 평가의 과정으로 이루어졌다(Figure 2).

1) 약물 구성과 투여 방법

총 3개월 간의 비만 프로그램 동안 가미태음조위탕 (1일 복용량. 마황 16~20g, 숙지황 8g, 의이인 8g, 생강 4g, 석창포 3.3g, 산조인 3.3g, 택사 2.6g, 황금 1.3g, 오미자 1.3g, 갈근 1.3g, 천문동 1.3g, 고본 1.3g, 용안육 1.3g, 건울 1.3g, 맥문동 1.3g) 처방은 약 한 달 간격으로 총 3회에 걸쳐 이루어졌다. 마황의 용량은 체중 및 발현 증상 변화에 따라 조절되었으며, 초진 방문 시 한식 위주의 균형 있는 식단을 바탕으로 한 저열량식이 및 산책, 스트레칭 등을 통한 신체 활동 유지 권장 등 체중 감량을 위한 일반적

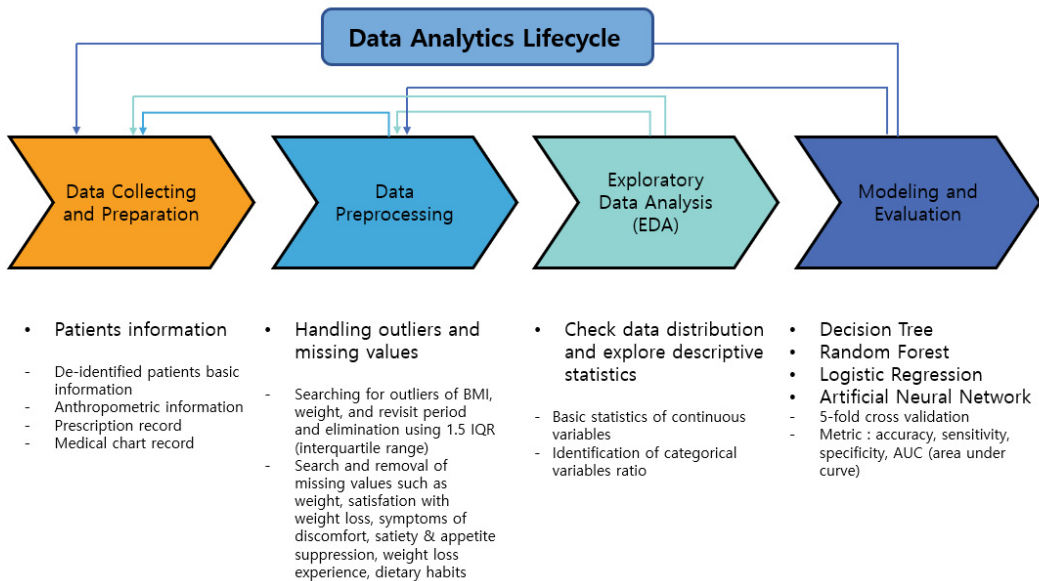


Fig. 2. Data analytics lifecycle

인 조인이 함께 이루어졌다. 처방 약물은 1일 복용량 한약을 달여서 동결건조한 정제 (550 mg/정)의 형태로 1회 3정, 1일 3회 복용하였다.

2) 데이터 수집 및 처리

(1) 분류 기준점에 따른 분석 대상

① 체중 감량 4분위 기준: 가미태음조위탕을 복용한 모집단의 체중 감량 정도를 각각 감량률, 감량 체중 기준으로 4분위 (상위 25%, 중위 50%, 하위 25%)로 나누어 이를 기준으로 삼아 입력변수의 선택 없이 모든 변수를 사용하여 분석

② 체중 감량 중위 50% 기준: 체중 감량 4분위 중에서 중위 50%를 분류 기준점으로 삼아 변수 선택 후 추가 분석

(2) 데이터 분석 시점과 용어

체중 감량 4분위 기준 분석과 체중 감량 중위 50% 기준 추가 분석 내에서 각각 1, 2, 3차 예측 분석이 세부적으로 이루어졌다. 1차 예측 분석은 1차 처방 시점에서 이루어졌으며, 2차 처방이 이루어지는 시점의 체중 감량 (1~2차 처방 사이에 발생하는 체중 감량률, 감량 체중)을 초진 정보를 활용하여 예측하였다. 2차 예측 분석은 2차 처방 시점에서 이루어

졌으며, 3차 처방이 이루어지는 시점의 체중 감량 (1~3차 처방 사이에 발생하는 체중 감량률, 감량 체중)을 초진 정보와 1-2차 약 복용 사이에 발생한 추가 정보를 활용하여 예측하였다. 3차 예측 분석은 3차 처방 시점에서 이루어졌으며, 최종 체중 보고가 이루어지는 시점의 체중 감량 (1차 처방~최종 체중 보고 사이에 발생하는 체중 감량률, 감량 체중)을 초진 정보, 1-2차 및 2-3차 약 복용 사이에 발생한 추가 정보를 활용하여 예측하였다(Figure 3).

(3) 측정

체중 정보는 Inbody 370 ((주)인바디, 서울, 대한민국)를 통해 이루어졌으며, 신장의 측정은 BSM 330 ((주)인바디, 서울, 대한민국) 신장계를 사용하여 측정되었다. BMI는 측정된 체중 (kg)을 키 (m)의 제곱으로 나눈 값으로 계산하여 사용하였다.

(4) 전처리

전체 데이터는 비식별화된 환자의 기본 정보, 계측 정보, 약 처방, 진료 노트가 모두 존재하는 사람을 대상으로 데이터를 추출했다. 제외 기준에 속하는 BMI가 23 kg/m^2 미만인 사람과 18세 미만의 환자는 제외하였다. 1차 예측 분석에 사용된 변수는 그대로

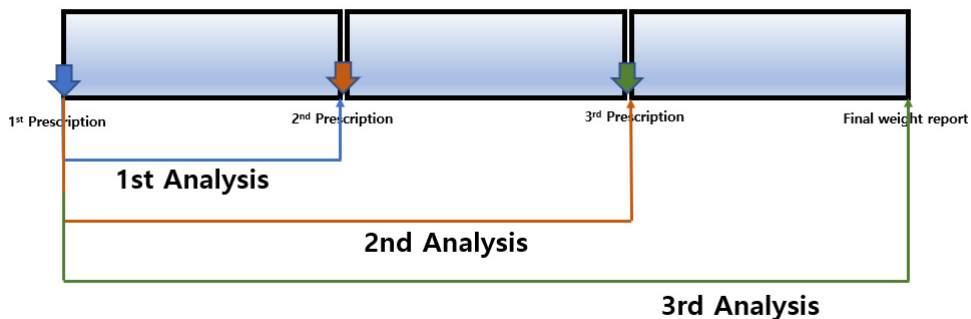


Fig. 3. A schematic diagram of prediction analyses of weight loss.

The analysis for predicting weight loss was divided into three parts, and the weight loss at each time point refers to the change from the initial point of treatment to the point of weight report.

2차 예측에 사용되었고, 1-2차 약 복용 사이에 추가된 변수는 3차 예측에 사용되었다.

연속형 변수 중 체중은 계측 정보와 진료 노트를 통하여 추출하였고, 그 과정에서 1.5 IQR (interquartile range)을 통해 구해진 이상치를 제거하고 결측치는 제외하였다. 재방문 기간은 각 차수 날짜 간의 차이를 구하여 추출하였고, BMI는 키와 방문 차수의 몸무게를 통하여 계산하여 구하였다. 감량 체중과 감량 BMI는 각 차수 때의 체중 혹은 BMI 간의 차이를 통하여 추출하였고, 감량률은 각 차수 간의 체중을 통해 아래와 같이 계산하여 추출하였다.

$$\text{감량률 (\%)} = \{(\text{초기 체중-예측 대상 시점 체중}) / \text{초기 체중}\} * 100 (\%)$$

3) 변수 선정

(1) 입력변수 (Feature vectors)

선행연구²⁴⁾를 통해 체중 감량에 영향을 미치는 인자들을 바탕으로 변수들이 선택되었으며, 1차 예측 분석에서는 초진 정보 중에서 연령, 초기 체중, BMI 등은 연속 변수로 사용되었으며, 약의 용량, 성별, 식이 습관, 다이어트 경험, 병력 등은 진료 노트 부분에서 텍스트 마이닝 형태로 추출되어 범주형 변수로 사용되었다. 2, 3차 예측 분석에서는 다이어트를 진행하면서 발생하는 추가 정보들을 변수로 활용하였으며, 각 처방 시점에서의 체중 및 BMI, 약 처방 사이의 기간, 감량률, 감량 체중, 감량 BMI 등이 연속형 변수로 분석에 추가되었으며, 추가 약접수시 방문의 유무, 마황 증량의 유무, 약 복용 후 발생한 기관계별 불편증상, 감량 만족도, 포만감 및 식욕억제는 범주형 변수로 추가되었다. 불편증상, 포만감, 감량만족도는 용언(用言)의 활용이 많아 형태소 분석을 위해 MeCab을 사용하였으며, 불편증상은 선행연구²⁵⁾를 참고하여 위장기관계, 중추 및 말초신경계, 신경정신계, 자율신경계, 그 외 기타로 범주화 하여 추출하였다(Table 1).

(2) 출력변수 (Labels) 및 분류의 기준점 (Classification bench mark)

본 연구의 체중 감량 예측은 이분형 분류 (binary classification) 형태로 시행되었다. 이분형 분류는 출력 변수 (labels)를 1 또는 0, 혹은 1 또는 -1로 예측하게 되는데, 본 연구에서는 예측 결과 분류의 기준점을 넘기면 1을 출력 (return)하고, 넘기지 못하면 0을 출력 (return)하도록 하였다. 연구는 분류의 기준점을 바탕으로 다음과 같이 크게 두 부분으로 진행되었다. 첫째, 모든 변수를 사용하여 가미태음조위탕을 복용한 모집단의 감량률 (reduced rate), 감량 체중 (reduced weight)을 각각 4분위 (분류 기준점: 상위 25%, 중위 50%, 하위 25%)로 나누어서 1차 예측 분석, 2차 예측 분석, 3차 예측 분석을 진행하였다(Table 2). 체중 감량이 양호한 군과 저조한 군을 비교하기 위해 1차 예측 분석에서 1분위와 4분위에 해당하는 환자들의 신체 정보, 감량 정도를 초진 정보와 비교해 보았다(Table 3). 둘째, 중요 변수를 탐색하기 위하여 체중 감량 4분위 중에서 감량률, 감량 체중의 중위 50%를 기준으로 추가 분석을 시행하였다. ‘속성 선택 기법 (feature selection method)’의 하나인 ‘보루타 알고리즘 (Boruta Algorithm)’²⁶⁾을 통해 의미 있는 변수를 선택한 후, 변수들의 선택이 잘 이루어져서 안정적으로 예측 정확도가 향상되는지 확인하기 위해 ‘변수 중요도 (feature importance)’에 따라 중요도가 낮은 변수에서 높은 변수 순서로 추가하면서 모델 성능이 변화하는 과정을 비교 검토하였다(Table 4).

4) 예측 모델

본 연구에서는 비만 진료를 통해 축적된 임상 빅데이터 분석을 위한 머신러닝 알고리즘으로 의사결정나무 (Decision Tree, DT), 랜덤 포레스트 (Random Forest, RF), 로지스틱 회귀분석 (Logistic Regression, LR), 인공신경망 (Artificial Neural Network, ANN)을 사용하였다.

DT는 불순도 (entropy)를 최소화하는 방식으로 이진 분할을 통해 유의미한 독립변수를 나누어 나가는데 그 모양이 나무의 가지가 갈라지는 모양과 유사하여 의사결정나무라고 불린다. 공선성에 강하며, 시각화가 가능하여 해석이 용이한 화이트 박스 (white box) 모델에 속하며, 수행을 빠르게 할 수 있고 컴퓨터 자원을 적게 소모하는 장점이 있어 본 연구에서 사용되었다. 한편, 과적합이 잘 되는 경향이 있고, 탐욕적 알고리즘을 기반으로 하기 때문에 최적해를 찾을 수 없는 경우도 발생한다는 단점이 있다^{13,27}.

RF는 훈련 자료에서 변수에 대하여 무작위 복원 추출 (bootstrap) 하여 DT를 만드는 과정을 반복해서 여러 개의 DT를 만든 후, 새로운 데이터 입력하여 예측을 시행할 때 회귀의 경우는 평균을, 분류의 경우에는 다수결 (voting)을 이용하는 방법을 이용한다. 이러한 방법을 앙상블 (ensemble) 기법 중에서 배깅

(bagging) 기법이라고 한다. 앙상블 방법을 통하여 예측의 변동성이 줄어들어 정밀도가 좋고 예측 성능이 우수하며, 변수의 중요성을 파악할 수 있다는 장점을 가지고 있어 본 연구의 분석에 활용되었다. 하지만, DT에 비해 속도가 크게 떨어지며 내부를 들여다볼 수가 없어 해석이 어려운 단점이 존재한다^{27,28}.

LR은 범주형 종속변수에 대하여 다중선형회귀를 적용하기 위한 방법으로, 선형적인 모델을 로그형태로 바꾼 뒤 범주에 대한 경계 값을 정하여 분류를 수행한다. 데이터의 차원이 크지 않을 때 과적합되지 않는 경향이 있고, 회귀계수를 구하여 독립변수가 가지는 음양의 방향성과 계수 값 비교를 통해 변수의 상대적인 중요도를 확인할 수 있는 장점을 가지고 있어 본 연구에 사용하였다. 그러나 차원이 큰 경우에는 성능이 좋지 않고, 종속변수가 범주형일 때만 사용 가능한 단점이 있다²⁷.

Table 1A. Independent Variables Used in the First Analysis (n=25,988)

Variables		
Age (years)		36.34 ± 10.5
Weight (kg)		72.69 ± 10.50
BMI (kg/m ²)		27.60 ± 3.04
Patients with medication dose change (n, %)	Stable	21,590 (83.08)
	Increased	4,398 (16.92)
Gender (n, %)	Male	2,292 (8.82)
	Female	23,696 (91.18)
Dietary habits (n, %)*	Light eating	749 (2.88)
	Binge eating	10,817 (41.62)
	Nighttime eating	8,197 (31.54)
Weight loss experience (n, %)	None	6,575 (25.3)
	Diet, exercise only or weight loss drug for less than 3 months	10,247 (39.43)
	Weight loss drug over 3 months	9,166 (35.27)
Diseases (n, %)*	High blood pressure	1,790 (6.89)
	Anemia	1,922 (7.4)
	Diabetes	676 (2.6)
	Hypothyroidism	598 (2.3)
	Hyperthyroidism	293 (1.13)
	Gastritis	846 (3.26)
	Reflux esophagitis	490 (1.89)
	Hyperlipidemia	610 (2.35)
	Low back pain	794 (3.06)

*: multiple choices allowed
Data are expressed as n (%) for categorical variables and mean ± SD for continuous variables.

Table 1B. Independent Variables Used in the 2nd Analysis (n=6,304)

Variables	Added Variables in 2 nd Analysis	
Age (years)	37.09±10.25	
Weight (kg)	72.51±10.36	
BMI (kg/m ²)	27.58±3.02	
Patients with medication dose change (n, %)	Stable	Weight2 (kg) 69.81±10.01
	Increased	BMI2 (kg/m ²) 26.55±2.94
		Prescription Period1-2 (days) 25.72±5.19
Gender (n, %)	Male	Reduced BMI1-2 (kg/m ²) 1.02±0.46
	Female	Reduced Rate1-2 (%) 3.71±1.6
Dietary habits (n, %)*	Light eating	Reduced Weight1-2 (kg) 2.7±1.25
	Binge eating	Symptoms of discomfort 1-2 (n, %)*
	Nighttime eating	Gastro-intestinal system 673 (10.68)
	None	Central and peripheral nervous system 497 (7.88)
Weight loss experience (n, %)	Diet, exercise only or weight loss drug for less than 3 months	Psychiatric Symptoms 204 (3.24)
	Weight loss drug over 3 months	Autonomic nervous system 117 (1.86)
Diseases (n, %)*	High blood pressure	Others 205 (3.25)
	Anemia	Satisfaction with weight loss 1-2 (n, %)
	Diabetes	Good 3,143 (49.86)
	Hypothyroidism	Fair 2,296 (36.42)
	Hyperthyroidism	Poor 865 (13.72)
	Gastritis	Good 2,691 (42.69)
	Reflux esophagitis	Fair 1381 (21.91)
	Hyperlipidemia	Poor 2232 (35.41)
	Low back pain	Attendance2 (n, %) 3,793 (60.17)

*: multiple choices allowed
Data are expressed as n (%) for categorical variables and mean ± SD for continuous variables.

Table 1C. Independent Variables Used in the 3rd Analysis (n=833)

Variables					
Age (years)			38.19 ± 10.42		
Weight (kg)			73.96 ± 10.26		
BMI (kg/m ²)			28.3 ± 3.14		
Patients with medication dose change (n, %)	Stable	648(77.79) (1 st)→481(57.74) (2 nd)→324(38.90) (3 rd)			
	Increased	185(22.21) (1 st)→352(42.26) (2 nd)→509(61.10) (3 rd)			
Gender (n, %)	Male		51 (8.82)		
	Female		782 (91.18)		
Dietary habits (n, %)*	Light eating		23 (2.76)		
	Binge eating		349 (41.9)		
	Nighttime eating		242 (29.05)		
Weight loss experience (n, %)	None		215 (25.81)		
	Diet, exercise only or weight loss drug for less than 3 months		305 (36.61)		
	Weight loss drug over 3 months		313 (37.58)		
Diseases (n, %)*	High blood pressure		72 (8.64)		
	Anemia		49 (5.88)		
	Diabetes		26 (3.12)		
	Hypothyroidism		30 (3.6)		
	Hyperthyroidism		8 (0.96)		
	Gastritis		18 (2.16)		
	Reflux esophagitis		10 (1.2)		
	Hyperlipidemia		27 (3.24)		
	Low back pain		28 (3.36)		
Added Variables in the 2nd Analysis		Added Variables in the 3rd Analysis			
Variables		Variables			
Weight2 (kg)	71.21 ± 9.81	Weight3 (kg)	68.97 ± 9.51		
BMI2 (kg/m ²)	27.25 ± 3.02	BMI3 (kg/m ²)	26.39 ± 2.94		
Prescription Period1-2 (days)	25.49 ± 5.24	Prescription Period1-3 (days)	58.57 ± 8.17		
		Prescription Period2-3 (days)	33.09 ± 6.53		
Reduced BMI1-2 (kg/m ²)	1.05 ± 0.46	Reduced BMI1-3 (kg/m ²)	1.91 ± 0.71		
		Reduced BMI2-3 (kg/m ²)	0.86 ± 0.48		
Reduced Rate1-2 (%)	3.70 ± 1.54	Reduced Rate1-3 (%)	6.71 ± 2.31		
		Reduced Rate2-3 (%)	3.13 ± 1.74		
Reduced Weight1-2 (kg)	2.75 ± 1.25	Reduced Weight1-3 (kg)	4.99 ± 1.93		
		Reduced Weight2-3 (kg)	2.24 ± 1.28		
Symptoms of discomfort 1-2 (n, %)*	Gastro-intestinal system	79 (9.48)	Symptoms of discomfort 2-3 (n, %)*	Gastro-intestinal system	72 (8.64)
	Central and peripheral nervous system	80 (9.6)		Central and peripheral nervous system	38 (4.56)
	Psychiatric Symptoms	27 (3.24)		Psychiatric Symptoms	24 (2.88)
	Autonomic nervous system	14 (1.68)		Autonomic nervous system	14 (1.68)
	Others	42 (5.04)		Others	29 (3.48)
Satisfaction with weight loss 1-2 (n, %)	Good	463 (55.58)	Satisfaction with weight loss 2-3 (n, %)	Good	472 (56.66)
	Fair	284 (34.09)		Fair	264 (31.69)
	Poor	86 (10.32)		Poor	97 (11.64)
Satiety and appetite suppression 1-2 (n, %)	Good	377 (45.26)	Satiety and appetite suppression 2-3 (n, %)	Good	343 (41.18)
	Fair	186 (22.33)		Fair	181 (21.73)
	Poor	270 (32.41)		Poor	309 (37.09)
Attendance2 (n, %)	586 (70.35)	Attendance3 (n, %)	616 (73.95)		

*: multiple choices allowed

Data are expressed as n (%) for categorical variables and mean ± SD for continuous variables.

ANN은 사람의 뉴런-시냅스 신경망 구조를 모방한 것으로 입력층 (input layer), 은닉층 (hidden layer), 출력층 (output layer)의 3개 층으로 이루어져 있다. 입력되는 독립변수와 출력되는 종속변수 사이에 복잡하고 다차원, 비선형 관계가 있을지라도 유연하게 모델링 할 수 있는 수학적인 알고리즘이다. 데이터가 많을 때 사용하기에 적절하고, 복잡한 비선형 관계를 찾아내는 능력과 변수들이 많을 때 접근할 수 있는

신뢰할 수 있는 모델이어서 본 연구 분석에 사용되었다. 한편, 각각의 독립변수가 종속변수에 어떤 영향을 미치는지 알 수 없는 블랙박스 (black box) 모델이며, 분석에 상당한 컴퓨터 자원과 시간이 요구된다는 단점이 있다²⁷⁾.

3. 분석 및 평가

데이터 분석에는 Jupyter-notebook, version 6.0.1

Table 2. Model Performance for Quartile Classification Bench Mark

		Reduced Rate					Reduced Weight				
1 st Bench mark		DT1	RF1	LR1	ANN1	1 st Bench mark		DT1	RF1	LR1	ANN1
Upper 25%	4.87%	74.91%	74.91%	74.91%	74.78%	Upper 25%	3.54kg	75.68%	75.70%	75.70%	76.12%
50%	3.75%	53.41%	54.73%	54.26%	55.01%	50%	2.68kg	59.47%	60.18%	59.05%	60.02%
Lower 25%	2.66%	74.80%	74.80%	74.70%	75.08%	Lower 25%	1.90kg	75.27%	75.27%	75.29%	75.41%
2 nd Bench mark		DT2	RF2	LR2	ANN2	2 nd Bench mark		DT2	RF2	LR2	ANN2
Upper 25%	8.21%	78.81%	80.50%	80.39%	80.17%	Upper 25%	5.94kg	82.51%	82.51%	81.13%	82.46%
50%	6.53%	71.04%	71.51%	71.67%	73.59%	50%	4.64kg	72.15%	73.36%	73.89%	75.79%
Lower 25%	4.87%	79.81%	79.39%	80.39%	80.96%	Lower 25%	3.44kg	80.13%	80.34%	80.13%	81.22%
3 rd Bench mark		DT3	RF3	LR3	ANN3	3 rd Bench mark		DT3	RF3	LR3	ANN3
Upper 25%	10.66%	86.00%	90.00%	86.80%	86.31%	Upper 25%	8.02kg	89.20%	90.00%	88.00%	87.51%
50%	8.68%	80.00%	83.20%	84.80%	83.91%	50%	6.22kg	86.00%	85.20%	84.40%	85.23%
Lower 25%	6.68%	84.00%	84.40%	85.60%	86.67%	Lower 25%	4.72kg	87.60%	86.00%	86.40%	86.55%

DT: Decision Tree; RF: Random Forest; LR: Logistic Regression; ANN: Artificial Neural Network
 Reduced rate of 1st Bench mark = (initial weight - weight at 2nd prescription)/ initial weight *100
 Reduced rate of 2nd Bench mark = (initial weight - weight at 3rd prescription)/ initial weight *100
 Reduced rate of 3rd Bench mark = (initial weight - weight at last weight report)/ initial weight *100

Table 3. Features of First and Fourth Quartile based on First Prediction Analysis

	Reduced Rate		Reduced Weight	
	More Than Upper 25% (n=6,554)	Less Than Lower 25% (n=6,527)	More Than Upper 25% (n=6,499)	Less Than Lower 25% (n=6,507)
Age (years)	34.92 ± 9.58	37.52 ± 10.59	34.59 ± 9.37	37.84 ± 10.67
Gender (n, %)	Female 5,930 (90) Male 624 (10)	Female 5,918 (91) Male 609 (9)	Female 5,400 (83) Male 1,099 (17)	Female 6,118 (94) Male 389 (6)
Height (cm)	162.28 ± 6.67	161.95 ± 6.71	164.22 ± 7.32	160.94 ± 6.26
Weight (kg)	72.76 ± 10.78	72.74 ± 10.5	77.68 ± 12.01	70.06 ± 9.26
BMI (kg/m ²)	27.56 ± 3.01	27.68 ± 3.07	28.71 ± 3.23	27.01 ± 2.81
Diet 1 (n, %)	1,908 (29)	1,426 (22)	1,902 (29)	1,393 (21)
Diet 2 (n, %)	2,689 (41)	2,436 (37)	2,593 (40)	2,482 (38)
Diet 3 (n, %)	1,957 (30)	2,665 (41)	2,004 (31)	2,632 (40)
RR (%)	5.89 ± 0.82	1.67 ± 0.79	5.76 ± 0.98	1.71 ± 0.84
RW (kg)	4.29 ± 0.87	1.22 ± 0.61	4.42 ± 0.77	1.18 ± 0.56

Data are expressed as n (%) for categorical variables and mean ± SD for continuous variables.
 Diet 1: Weight Loss Experience_None; Diet 2: Diet, exercise only or weight loss drug for less than 3 months; Diet 3: Weight Loss Experience_Weight Loss Drug over 3 Months; RR: Reduced Rate; RW: Reduced Weight

(Python, version 3.7.4; Python Software Foundation (PSF), Wilmington, Delaware, USA)을 사용하였다. 모델링 수행에 사용된 DT, RF, LR은 Scikit-learn을 ANN은 Keras를 메인 패키지로 사용하였다. 속성 선택 기법에 사용된 Boruta algorithm은 BorutaPy 패키지를 이용하였으며, 분산팽창지수 (Variance Inflation Factor, VIF)는 Statsmodels로 분석되었다. 속성 선택 기법을 통해 변수 선택이 이루어진 후 구한 변수 중요도는 RF 기반^{29,30)}으로 결정되었다.

편향 (Underfitting)과 과적합 (overfitting)의 균형을 맞추기 위해 검증 및 평가에 5-fold cross validation을 사용하였다³¹⁾. 5-fold cross validation은 사용할 전체 데이터를 5개의 블록으로 나누고, 한 블록을 테스트 세트 (test set)로 나머지는 훈련 세트

(train set)가 되도록 하고, 이 과정을 5번 반복하여 모든 데이터가 테스트 데이터로 쓰이는 경우가 존재하도록 만든 검증 기법이다. 각 과정 (iteration)을 통해 나온 정확도 (accuracy)를 평가지표로 삼았으며 5번의 결과를 평균 내어 나온 값을 모델을 평가하는데 사용하였다. 매개변수 (Hyperparameter)의 선택은 매개 변수를 튜닝할 때 설정해 둔 각 범위에 대하여 가능한 모든 조합을 시도하여 모델링한 뒤 모델링 결과가 가장 좋은 매개 변수를 고르는 방법인 그리드 서치 (grid search)를 이용하였다.

분류의 기준점을 중위 50%로 하여 추가 분석 시행 후 예측력을 평가함에 있어서 5-fold cross validation을 통해 얻어진 정확도 (accuracy) 이외에, receiver operating characteristics (ROC) 커브를 통해 area

Table 4. Model Performance according to Variables Ranking Based on Feature Importance

		Reduced Rate in 1 st Analysis				Reduced Weight in 1 st Analysis				
Ranking	Variables	DT (%)	RF (%)	LR (%)	ANN (%)	Variables	DT (%)	RF (%)	LR (%)	ANN (%)
7						Diet 1	52.19	52.19	52.19	52.44
6						MD_S	54.51	54.51	54.51	54.71
5						Diet 3	54.51	54.51	54.51	54.71
4						Gender	55.07	55.07	55.07	55.22
3	Diet3	52.39	52.39	52.39	53.44	Age	55.91	55.61	55.46	56.48
2	Weight	52.53	52.64	52.39	53.44	BMI	58.61	59.20	58.56	58.90
1	Age	54.06	54.05	53.79	54.69	Weight	58.78	60.06	59.05	59.95
		Reduced Rate in 2 nd Analysis				Reduced Weight in 2 nd Analysis				
Ranking	Variables	DT (%)	RF (%)	LR (%)	ANN (%)	Variables	DT (%)	RF (%)	LR (%)	ANN (%)
8						Gender	54.65	54.65	54.65	53.44
7						SAS 1-2_G	57.14	57.14	57.14	57.63
6	SAS1-2_G	54.44	54.44	54.44	55.50	MD_S	57.77	57.77	57.77	58.60
5	SWL1-2_B	56.40	56.40	56.40	56.76	SWL1-2_B	59.73	59.73	59.83	60.01
4	Weight	57.24	56.50	56.50	57.12	Age	58.56	58.67	60.15	61.02
3	Age	55.97	57.03	57.40	59.18	SWL1-2_G	61.31	62.21	62.58	63.52
2	SWL1-2_G	60.94	61.52	62.42	62.21	Weight	64.16	64.64	64.11	66.23
1	RR 1-2	72.04	71.83	70.51	73.52	RR 1-2	73.41	73.15	72.20	75.33
		Reduced Rate in 3 rd Analysis				Reduced Weight in 3 rd Analysis				
Ranking	Variables	DT (%)	RF (%)	LR (%)	ANN (%)	Variables	DT (%)	RF (%)	LR (%)	ANN (%)
4	Age	54.80	52.80	53.60	58.95	Age	59.20	59.20	58.00	61.95
3	Weight3	52.80	56.80	50.40	57.63	Weight	59.20	64.80	60.80	67.71
2	RR 1-2	70.80	73.60	69.20	71.55	RW 2-3	71.60	76.80	72.00	76.11
1	RR 1-3	80.00	81.20	81.60	81.88	RW 1-3	86.00	86.00	82.00	83.67

DT: Decision Tree; RF: Random Forest; LR: Logistic Regression; ANN: Artificial Neural Network; Diet 3: Weight Loss Experience_Weight Loss Drug over 3 Months; MD_S: Patients with Medication Dose Change_Stable; Diet 1: Weight Loss Experience_None; SAS 1-2_G: Satiety and Appetite Suppression 1-2_Good; SWL1-2_B: Satisfaction with Weight Loss 1-2_Bad; SWL1-2_G: Satisfaction with Weight Loss 1-2_Good; RR: Reduced Rate; RW: Reduced Weight

under the curve (AUC)의 값을 확인하여 추가적으로 모델성능을 평가하는데 사용하였으며, AUC의 값은 1에 가까울수록 더 좋은 성능을 의미한다. 체중 감량 4분위를 분류 기준점으로 했을 때 보다 AUC를 통해 추가적인 예측력 평가를 진행한 이유는 4분위 중에서 중위 50%가 가장 설명력 있다고 판단하여 이를 본 연구의 최종 분류 기준점으로 삼았기 때문이며, 더불어 ROC 커브의 시각화를 통해 모델 성능의 변화를 한 눈에 살펴보기 위한 목적이었다.

결 과

1. 연구 대상의 일반적 특성

1차 체중 감량 예측 분석 대상자는 25,988명으로 평균 연령은 36.34 ± 10.5 세, 체중은 72.69 ± 10.50

kg, BMI는 $27.60 \pm 3.04 \text{ kg/m}^2$ 이었으며 남녀 비율은 남성은 8.82%, 여성은 91.18%로 여성의 비율이 우세하며 2차, 3차 분석에서도 근사한 비율을 유지하였다. 1차 분석에서 사용된 독립변수의 경우 총 22개이다. 2차 체중 감량 예측 분석 대상자는 6,304명으로 평균 연령은 37.09 ± 10.25 세, 초진 시 체중은 $72.51 \pm 10.36 \text{ kg}$, BMI는 $27.58 \pm 3.02 \text{ kg/m}^2$, 1차 처방과 2차 처방 사이의 기간은 25.72 ± 5.19 일이었으며, 다이어트를 진행하면서 발생한 정보를 추가하여 분석에 사용된 총 독립변수는 42개이다. 3차 체중 감량 예측 분석대상자는 833명으로 평균 연령은 38.19 ± 10.42 세, 초진 시 체중은 $73.96 \pm 10.26 \text{ kg}$, BMI는 $28.3 \pm 3.14 \text{ kg/m}^2$, 2차 처방과 3차 처방 사이의 기간은 33.09 ± 6.53 일이었으며, 사용된 독립변수의 총 개수는 66개이다. 1차에서 3차 예측 분

Table 5. Prediction Model Performance Results based on AUC

		Algorithm	Sensitivity	Specificity	AUC
1 st Analysis	Reduced Rate	DT	0.557	0.524	0.551
		RF	0.591	0.489	0.557
		LR	0.576	0.499	0.546
		ANN	0.574	0.497	0.550
	Reduced weight	DT	0.709	0.465	0.630
		RF	0.609	0.591	0.640
		LR	0.582	0.602	0.631
		ANN	0.426	0.748	0.620
2 nd Analysis	Reduced Rate	DT	0.722	0.719	0.791
		RF	0.748	0.690	0.789
		LR	0.734	0.677	0.785
		ANN	0.805	0.606	0.785
	Reduced weight	DT	0.784	0.687	0.802
		RF	0.764	0.700	0.816
		LR	0.761	0.685	0.801
		ANN	0.880	0.557	0.798
3 rd Analysis	Reduced Rate	DT	0.762	0.836	0.890
		RF	0.787	0.836	0.890
		LR	0.852	0.781	0.897
		ANN	0.828	0.789	0.880
	Reduced weight	DT	0.873	0.848	0.937
		RF	0.873	0.848	0.939
		LR	0.847	0.795	0.905
		ANN	0.788	0.856	0.920

DT: Decision Tree; RF: Random Forest; LR: Logistic Regression; ANN: Artificial Neural Network;

AUC: Area Under the Curve (0.90 - 1.00: excellent, 0.80 - 0.90: good, 0.70 - 0.80: fair, 0.60 - 0.70: poor, 0.50 - 0.60: fail)

석 대상자로 갈수록 초진 체중과 BMI가 비교적 높아지는 경향을 보이는데, 3개월 프로그램 진행 후 획득되는 체중 정보는 추가적인 감량을 상담할 때 얻어지는 경우가 많아 비교적 감량할 범위가 많은 사람들이 대상자에 포함되었기 때문으로 생각된다. 약의 용량을 증량하는 경우는 1차 처방 시점에서 2, 3차로 갈수록 증가하였다(Table 1).

체중 감량이 양호한 군과 저조한 군을 비교하기 위해 1차 예측 분석에서 1분위와 4분위에 해당하는 환자들의 신체 정보, 감량 정도를 초진 정보와 비교해 본 결과 감량률, 감량 체중 기준 모두에서 감량 양호군(1분위)에서 감량 저조군(4분위)보다 상대적으로 연령이 낮고 다이어트 경험이 적은 경향이 있었으며, 감량 체중 기준으로는 감량 양호군에서 초기 비만도가 상대적으로 높은 것을 확인할 수 있었다(Table 3).

2. 모델 성능 비교

1) 체중 감량 4분위 분류 기준점에 따른 모델 성능 비교

체중 감량 4분위 분류 기준점(classification benchmark)이라 함은 가미태음조위탕을 복용한 모집단의 감량률, 감량 체중의 상위 25%, 중위 50%, 하위 25%을 의미한다. 모델 성능은 5-fold cross validation으로 확인하였다(Table 2).

(1) 1차 예측 분석

1차 처방 약 복용 후 2차 처방 시점에서 확인된 평균 체중 감량률은 3.77 ± 1.66 %, 평균 감량 체중은 2.74 ± 1.29 kg으로 4분위로 나누어 살펴본 하위 25%, 중위 50%, 상위 25%에 해당하는 감량률은 2.66%, 3.75%, 4.87%이며, 감량 체중은 1.9kg, 2.68kg, 3.54kg이다. 상위 25%와 하위 25% 기준으로는 모델링의 결과가 감량률, 감량 체중 기준에 관계없이 모두 약 75% 전후로 나왔으며, 감량률 50%를 기준으로는 모델별로 53.41% (DT)에서 55.01% (ANN)이, 감량 체중 50%를 기준으로는 59.47% (DT)에서

60.18% (RF)로 나왔다. 1차 예측 모델에서는 분류 기준점 체중 감량 50%에서 감량 체중을 기준으로 한 결과가 감량률 기준으로 한 결과보다 각 모델별로 약 5%정도 높게 나왔다.

(2) 2차 예측 분석

2차 처방 약 복용 후 3차 처방 시점에서 확인된 평균 체중 감량률은 6.54 ± 4.76 %, 평균 감량 체중은 4.76 ± 2.00 kg으로 4분위로 나누어 살펴본 하위 25%, 중위 50%, 상위 25%에 해당하는 감량률은 4.87%, 6.53%, 8.21%이며, 감량 체중은 3.44kg, 4.64kg, 5.94kg이다. 상위 25%와 하위 25% 기준으로는 모델링의 결과 값은 약 80% 전후로 나왔다. 중위 50% 기준으로는 71.04 % (감량률 기준, DT)에서 75.79% (감량체중 기준, ANN)까지 분포하고 있으며, 1차 모델에서 감량률 기준과 감량체중 기준 결과의 차이가 약 5%로 나타났는데 2차 모델에서는 그 차이가 2% 정도로 줄어 들었다. 체중 감량 4분위 기준별 모델링의 결과 값들이 1차 분석 결과 값과 비교 시 모두 상승하였으며, 2차 예측 분석부터는 1차에서 나왔던 상하위 25% 기준으로 한 가지 정답만 계속 고르더라도 나올 수 있는 정확도인 75%를 벗어나 약 80% 전후의 결과 값들이 나왔다. 이로부터 2차 예측 분석에서는 설명력 있는 변수가 추가되어 1차 분석 때와는 다른 방법으로 예측이 진행되고 있음을 추정해 볼 수 있다.

(3) 3차 예측 분석

3차 처방 약 복용 후 최종 체중 보고 시점에서 확인된 평균 체중 감량률은 8.74 ± 3.07 %, 평균 감량 체중은 6.50 ± 2.59 kg으로 4분위로 나누어 살펴본 하위 25%, 중위 50%, 상위 25%에 해당하는 감량률은 6.68%, 8.68%, 10.66%이며, 감량 체중은 4.72kg, 6.22kg, 8.02kg이다. 상위 25% 기준으로는 86% (감량률 기준, DT)에서 90%까지의 예측 정확도가 나왔는데, 특히 감량률, 감량 체중 기준 모두에서 RF를

사용하였을 경우 90%의 결과 값이 나왔다. 하위 25% 기준으로는 결과 값이 84% (감량률 기준, DT)에서 87.60% (감량 체중 기준, DT), 중위 50% 기준으로는 80% (감량률 기준, DT)에서 86% (감량 체중 기준, DT)까지 분포하였다. 3차 분석에서 예측 정확도는 2차 분석과 비교 시 모두 상승되었음을 확인할 수 있었다.

2) 체중 감량 중위 50% 분류 기준점에서 변수 선택을 통한 모델 성능 추가 비교

(1) 변수 선택과 변수 중요도

속성 선택기법 사용 후 감량률, 감량 체중 기준으로 1차 예측 분석에서는 22개의 독립 변수 중에서 각각 3개, 8개, 2차 예측 분석에서는 42개의 변수 중에서 각각 12개, 15개, 그리고 3차 예측 분석에서는 66개의 변수 중에서 각각 13개, 16개의 변수가 선택되었다. 1차 분석에서는 감량률, 감량 체중 기준 분석 모두에서 연령, 초기 체중, 다이어트 경험 3개월 이상이 중요한 변수로 선택되었으며, 2차 분석에서는 1차 처방 복용 후 2차 처방 시점까지의 감량률과 감량 체중 등이 높은 순위를 차지 하였으며, 약을 복용하면서 추가로 획득된 변수인 감량 만족도와 식욕억제 및 포만감에 대한 변수 또한 중요 변수로 선택되었음을 확인할 수 있었다. 3차 분석에서 순위가 높은 변수로 선택된 것은 1차 처방 시점에서 3차 처방 시점까지의 감량률, 감량 체중으로 나타났으며, 나머지 선택된 변수들을 살펴보면 감량 결과로 얻게 되는 체중 정보와 관련된 변수들이 대부분임을 확인할 수 있었다.

(2) 모델 성능 비교

① 변수 중요도 기반 변수 추가에 따른 모델능의 변화: 5-fold cross validation

선택된 변수들을 변수 중요도^{29,30)}가 낮은 순서에서 높은 순서로 추가하면서 모델 성능이 개선되는지 확인하는 과정을 거쳤다. 이 과정에서 변수가 하나씩

추가될 때 마다 반복적인 학습을 되풀이하게 되는데, 이 때 발생하는 모델 학습 소요 시간을 합리적으로 단축하기 분산팽창지수 (Variance Inflation factor, VIF)를 확인하여 값이 10이 넘는 변수들 중 중요도가 낮은 변수들부터 차례로 제거하여 모든 변수들의 VIF가 10이 넘지 않도록 했다. 그 결과 감량률, 감량 체중 기준으로 1차 예측 분석에서는 각각 3개 중에 3개, 8개 중에 7개의 변수가, 2차 예측 분석에서는 각각 12개 중에 6개, 15개 중에 8개의 변수가, 그리고 3차 예측 분석에서는 각각 13개 중에 4개, 16개의 중에 4개의 변수가 남았다. 중요도 순위가 낮은 변수들부터 순위가 높은 변수들을 차례로 하나씩 추가하면서 5-fold cross validation으로 모델 성능을 확인한 결과 큰 폭으로 정확도 (accuracy)가 하락하는 경우 없이 대부분 순차적으로 정확도가 증가하는 것을 볼 수 있었으며, 감량률 기준으로는 ANN이 1차 (54.69%), 2차 (73.52%), 3차 (81.88%) 예측 분석 모두에서 가장 높은 결과 값을 보였다. 변수를 선택해서 진행한 3차 예측 분석에서 감량률 기준으로는 RF (81.20%), LR (81.60%), ANN (81.88%) 모델에서와 감량 체중 기준으로 LR (82.00%), ANN (83.67%) 모델에서, 변수를 선택하지 않고 모든 변수를 넣고 돌린 분류 기준점 체중 감량 4분위 중 중위 50%의 3차 예측 분석 결과와 비교해 봤을 때, 변수 선택 및 VIF를 사용하여 변수를 줄이고 돌린 정확도가 오히려 상대적으로 더 낮게 나온 것을 확인할 수 있었다(Table 4).

② AUC 값을 바탕으로 한 모델 예측 성능 추가 비교

시각화를 통해 모델 성능 변화를 한눈에 파악하기 위한 목적으로 체중 감량 중위 50% 분류 기준점에서 변수 선택 후 진행한 분석에서는 ROC 커브를 통해 나오는 AUC의 값을 확인하여 추가적으로 모델 성능을 비교하였으며, 각 차수의 예측 분석에서 모델별로 민감도, 특이도, AUC 값을 획득하였다(Table 5).

감량 체중 기준으로 RF에서 1차 (0.640), 2차 (0.816), 3차 (0.939) 예측 분석 모두 가장 좋은 결과 값을 보였다. 1~3차 예측 분석의 모든 모델을 통틀어 3차 예측분석, 감량 체중 기준으로 RF로 분석을 진행했을 경우 민감도 (0.873), 특이도 (0.848), AUC (0.939)가 가장 높게 나왔다. 각 분석 차수별로 가장 높은 AUC 값을 가진 ROC 커브들을 감량률, 감량 체중 기준으로 시각화 하였으며, 차수가 지날수록 AUC 값이 개선됨을 한 눈에 확인할 수 있었다. ROC 커브는 AUC의 값에 따라 0.90 - 1.00는 excellent, 0.80 - 0.90는 good, 0.70 - 0.80는 fair, 0.60 - 0.70는 poor, 0.50 - 0.60는 fail로 해석하는데, 3차 예측 분석에서 감량률 기준으로는 가장 높은 AUC값이 0.90 (LR), 감량 체중 기준으로는 0.94 (RF)로 모두 excellent 범주에 들었다(Figure 4).

지고 있는데 임상상의 지도하에 체중 감량 증재가 이루어지는 방법 뿐만 아니라 상업용 시판 프로그램, 셀프(self) 감량 시도 및 모바일 앱을 이용하는 방법 등이 그 예이다³²⁾. 이러한 다양한 접근들로 인하여 관련 데이터가 빠르게 쌓이고 있으나, 빅데이터가 가지는 복잡한 속성 때문에 기존의 고전적인 통계 방법 만으로는 분석에 어려움이 있다. 특히 electronic health records (EHR) 데이터는 구조화되어 있는 상태로 존재하는 경우도 있지만 병력에 대한 부분은 대부분 비구조화 되어 있어 정제를 위해 전처리가 필요한 경우가 많은데, 그 양이 방대하므로 연구자가 일일이 모든 내용을 확인하는데 어려움이 있을 수 있다³³⁾. 인공지능의 한 분야인 머신러닝은 명시적인 지시를 위해 사람이 모든 상황을 세밀하게 코딩하지 않아도 컴퓨터 시스템이 데이터 내에 존재하는 관계, 규칙, 패턴들을 탐색하여 스스로 배우고 개선해 나가는 알고리즘 및 통계 모델로 특히 대용량의 자료를 다루는데 유용하다¹⁵⁾. 따라서, 본 연구에서는 한방 비만 임상 빅데이터 분석을 위해 DT, LR, RF, ANN

고 찰

최근 체중 감량 시도는 다각화된 방법으로 이루어

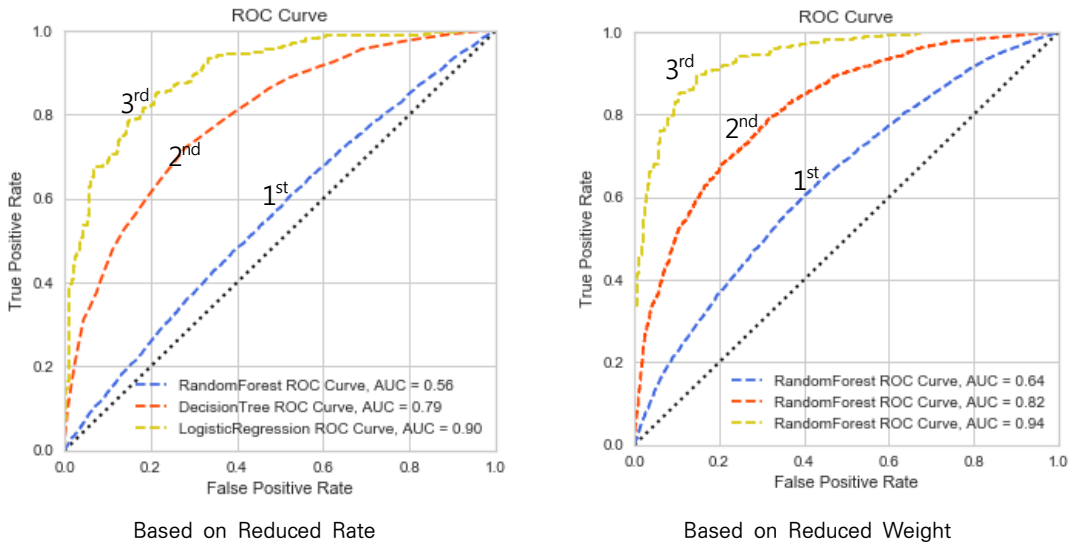


Fig. 4. Receiver operating characteristics (ROC) curves

등의 머신러닝 기법을 사용하여 체중 감량 예측 연구를 진행하였다.

본 연구의 체중 감량 예측은 정해진 체중 감량 성공 기준에 따라 감량 성공과 감량 실패를 예측한 선행연구들^{13,22)}을 바탕으로 이분형 분류 (binary classification) 형태로 시행되었다. 다양한 선행연구^{9,10,13,34)}에서 체중 감량의 성공 기준을 초기 체중의 5%를 감량한 것으로 설정하여 연구를 진행하였는데, 초기 체중의 5% 감량은 다양한 인체 기관에서 대사 기능을 개선시키는 것으로 알려져 있어³⁵⁾ 체중 감량 성공 기준으로 연구에 많이 적용된다. 체중 감량 시도 후 3-6개월³⁴⁾ 혹은 5개월²²⁾때 초기 체중의 5% 이상 체중 감량을 감량 성공 기준으로 본 연구도 있는가 하면, 다수의 연구는 12개월 동안 5% 감량을 성공 기준으로 보고 있었다^{9,10,13)}. 본원 비만 프로그램을 이용한 환자들의 약 1달 뒤 평균 감량률은 3.77%, 약 2달 뒤 평균 감량률은 6.54%로 평균 감량률이 2달 이내에 벌써 초기 체중의 5%를 뛰어 넘기 때문에 체중 감량의 성공 기준을 초기 체중의 5% 감량으로 설정하는 것은 임상 적용 시 활용도가 떨어진다는 판단 하에, 분류의 기준점 (classification benchmark) 설정을 위해 다양한 고찰과 시도를 통해 최종 기준을 설정하였다.

Annesi 등⁸⁾은 체중 감량률을 4분위로 나누어서 분석 그룹 내 상위 25%를 감량 성공으로, 하위 25%를 감량 실패로 분석하였는데, 이러한 방법에 착안하여 본 연구에서는 체중 감량률 및 감량 체중을 4분위로 나누어 각각을 분류의 기준점으로 하여 다양한 머신러닝 기법을 사용해 이분형 분류 (binary classification)를 시행하였다. 체중 감량 4분위를 기준으로 먼저 분석을 시행한 이유는 선행연구들에서 체중 감량 성공 기준을 초기 체중의 5% 감량으로 잡은 것처럼 본 연구에서도 최종적으로 가장 의미 있는 하나의 분류 기준점을 찾기 위해 선행적으로 분석을 시도해보기 위함이었다. 최종적으로 의미를 두고 추가 분석을 시행한 분류의 기준점은 체중 감량 중위 50%이었다.

모집단 내에서 중위 50%에 해당하는 감량률, 감량 체중을 추가 분석을 위한 분류의 기준점으로 선택한 이유는 본원 비만 프로그램 시행 시 예상되는 중등도 감량 정도 보다 양호하게 감량이 될 지, 저조하게 감량이 될 지를 제시할 수 있으면서도, 상대적으로 다른 차수에 비해 예측이 어려울 수밖에 없는 1차 예측 분석에서 다른 기준 값들 (그룹 내 상하위 25%)보다 추가적인 설명력이 있었기 때문이다. 1차 예측분석 결과를 살펴보면, 종속변수를 중위 50% 감량률, 감량 체중을 기준으로 하여 분석할 경우 예측 정확도 (accuracy) 수치 자체는 다른 결과 값들 보다 낮아 보인다. 그러나, 종속변수를 상위 25%, 하위 25%로 설정할 경우 컴퓨터가 학습 (training) 후 테스트 (test) 과정에서 한 가지 정답만 똑같이 계속 고르더라도 나올 수 있는 정확도인 75%에 가까운 값들이 대부분이었다. 따라서, 체중 감량 중위 50% 감량률, 감량 체중을 분류의 기준점으로 설정할 경우가 모델 분석에서 추가적인 설명력이 더 있다고 판단하였다 (Table 2).

속성 선택 기법 시행 후 1차 예측 분석에서 체중 감량 기준으로는 더 많은 변수들이 선택되었으며 변수들이 추가될 때 마다 설명력이 높아지게 되고 정확도는 상승하여 감량률 기준으로 분석했을 때보다 결과 값이 상대적으로 높게 나왔다. 하지만 1차 예측 분석의 전체적인 정확도는 54.05% (감량률 기준, RF)부터 60.06% (감량 체중 기준, RF)까지로 높지 않았으며, AUC 기준으로도 0.550 (감량률 기준, ANN)에서 0.640 (감량 체중 기준, RF)까지 나와 fail에서 poor 범주에 속했다. 체중 감량에 영향을 주는 인자들에 대하여 고찰한 선행연구를 살펴보면 1차 예측 분석에서 속성 선택 기법을 통해 선택된 변수들인 초기비만도, 성별, 연령, 다이어트 경험 등은 모두 그 영향에 있어서 방향성이 일관되지 않은 부분이 있는 인자들이며 상충된 연구 결과들이 존재한다²⁴⁾. 약의 용량이라는 변수 또한 초진 시에는 초기 비만도에 의해 결정되는 경우가 많다. 일관성 있게

작용하는 예측 인자가 1차 예측 분석의 입력(독립)변수에서는 부재해 있다고 볼 수 있다. 또한 1차 예측 분석이 이루어지는 시점은 초진 내원 시인데 초진 시에는 다이어트에 권장되는 식사 습관에 대한 간략한 안내가 이루어지고, 처방되는 약 또한 복용 후 식욕억제를 경험하게 되어 초진 시점에서 환자가 가지고 있는 정보와 실제로 다이어트를 시작하고 나서 가지게 되는 식이 습관 등은 달라질 수 밖에 없는데 초진 시 가지고 있는 정보만으로 향후 감량 예측을 진행한다는 것이 어려운 부분이 있기 때문에 1차 예측 정확도는 낮게 나왔을 가능성을 생각해 볼 수 있겠다.

2차 예측 분석부터는 다이어트를 진행하면서 얻을 수 있는 변수들이 추가되었으며 1-2차 처방 시점 사이의 체중 감량률, 감량만족도, 식욕억제 및 포만감이 중요 변수로 선택되었다. 감량만족도가 좋은 경우가 변수로 추가되면서 정확도가 1차 예측 분석 결과를 넘어섰으며, 최종적으로 변수 중요도가 가장 높은 1-2차 처방 시점 사이의 체중 감량률 변수가 추가되면서 감량률 기준, 감량 체중 기준 모두에서 예측 정확도가 상당한 비율로 상승하는 것을 확인할 수 있었다. 선행연구^{5,9,10,24)}에 따르면 초기 체중 감량(initial weight loss)은 상승되지 않고 일관되게 체중 감량을 예측할 수 있는 중요 인자로 거론되고 있으며, 이러한 강력한 입력(독립)변수가 추가됨에 따라 2차 예측 분석의 정확도가 상승한 것으로 생각된다. 초기 체중 감량이라는 입력(독립)변수는 체중 감소에 영향을 미치는 요인들은 다양하지만 적용되고 있는 약, 식이, 운동 등의 중재 요인들의 영향으로 실질적으로 체중이 얼마나 잘 감소되고 있는지를 가장 잘 반영하기 때문에 예측력에 중요한 영향을 미치는 것으로 생각된다.

3차 예측 분석에서는 2차 예측 분석에서 선택된 변수들과 달리 감량만족도, 식욕억제 및 포만감과 같은 변수들은 선택되지 않았으며, 실제 감량이 일어난 정도와 관련된 변수들 위주로 선택되었다. 가미태음

조위탕의 주성분인 마황은 식욕억제 효과를 나타내지만 반응급감현상을 보여 반복적으로 사용하면 효과가 줄어들게 되는데³⁶⁾, 처방 차수가 지날수록 마황 용량이 높은 처방을 복용하는 사람이 증가하는 것 또한 이러한 반응급감현상과 관련이 있다(Table 1). 따라서, 체중 감소가 양호하게 진행되고 있는 사람 또한 반응급감현상 때문에 식욕억제 정도가 줄어들면서 의식적인 식사 조절이 증가하고 이로 인한 만족도가 줄어들 수 있는데, 이런 경우 실제 체중 감량 정도와 식욕억제나 감량만족도 변수와는 일치되지 않는 결과를 보일 수 있어 3차 예측 분석에서는 중요 변수로 선택되지 않았을 가능성을 고려해볼 수 있다. 2차 예측 분석과 유사하게 3차 예측 분석에서도 가장 중요한 변수는 1-3차 처방 시점 사이의 체중 감량률과 감량 체중이 선택되었으며, 이와 더불어 1-2차, 2-3차 사이의 부분적인 감량률과 감량 체중도 중요 변수로 선택되었다(Table 4). 체중 감량의 궤적(trajectories)은 개인마다 차이가 있을 수 있는데, 분석 직전까지의 체중 감량과 더불어 그 사이의 부분적인 감량 정도가 추가적인 변수로 모두 들어가 있어 향후 변화에 대한 예측이 더 용이 해졌을 가능성을 생각해 볼 수 있으며, 이에 따라 3차 예측 분석의 정확도는 2차에 비하여 상승한 것으로 생각된다.

본 연구는 한방 비만 진료를 통해 축적된 비교적 통제된 프로그램에 참여한 환자들의 실제 임상 빅데이터를 사용해 기존의 전통적인 예측 방법이 아닌 머신러닝 기법을 사용하여 분석을 시도했다는데 의미가 있다. 한방 임상 빅데이터를 이용하여 머신러닝 기법을 적용한 연구들은 많지 않으며, 특히 한방 비만 진료 분야에서는 아직까지 선행연구를 찾아보기는 쉽지 않다. 검색일 2020년 2월 20일 기준으로 검색 DB는 Pubmed에서 (obesity OR weight loss) AND (artificial intelligence OR big data OR data mining OR machine learning OR deep learning OR artificial neural network) 을 검색식으로 하여 검색을 진행한 결과 총 170편의 논문이 검색되었으며, 제

목과 초록을 탐색하여 관련 주제와 상관 없는 연구들을 배제하고, 총 83편의 논문을 리뷰한 결과 비만 관련 예측 (prediction) 선행연구들을 확인할 수 있었다^{20,21}. 성인을 대상으로 하여 체중 감량을 예측을 목표로 하는 연구들은 Batterham 등¹³)과 Aswani 등²²)의 연구를 찾아볼 수 있었다. 두 연구 모두 이분형 분류 (binary classification)로 감량을 예측하는 것이 본 연구와 유사한 점이었고, 가장 큰 차이점은 분석 대상자 수에 있었는데, Batterham 등¹³)의 연구는 93명, Aswani 등²²)의 연구는 61명의 임상연구 (trial) 참가자가 분석 대상자였으며, 모두 임상연구였기 때문에 통제된 환경 설정을 바탕으로 적은 대상자 수를 바탕으로 분석이 이루어졌다. 국내 연구와 마찬가지로 해외 연구에서도 비만관련 실제 임상 진료 빅데이터를 활용하여 체중 감량 예측을 위한 머신러닝 기법을 활용한 예는 많지 않음을 확인할 수 있었다.

본질적으로 실제 임상 빅데이터는 누락되거나 잘못된 기록된 데이터 등 잡음 (noise)이 많고, 복잡성을 가지고 있는데, 머신러닝 기법은 전통적으로 체중 감량의 예측을 위해 사용된 선형 회귀 및 기타 통계적 방법보다 이러한 문제를 더욱 포괄적으로 처리할 수 있는 장점을 가지고 있어 이러한 접근 방식이 앞으로 비만 관련 예측에 접근하는 효과적인 방법이 될 수 있으며²⁰), 본 연구는 그 시도의 첫걸음을 내딛는 과정을 진행했다는 데 의의를 찾을 수 있겠다.

본 연구의 제한점은 다음과 같다.

첫째, 분석에 사용된 변수 이외에도 사회경제적 요인 및 복약 순응도 등 많은 변수가 체중 감소에 영향을 미칠 수 있는데, EHR data의 한계상 진료 차트에서 추출될 수 있는 정보만 사용 가능하다는 점 때문에 현재 가지고 있는 정보만으로 분석이 시행된 점이다. 특히, 복약 순응도는 선행연구에서 중요한 비만 치료 영향 인자로 나타났으나³⁴) 본 연구에서는 정보 획득의 어려움으로 이를 변수로 사용되지 못하였는데, 향후 차트 개선을 통해 이러한 정보를 추가하여 모델을 개선해 나간다면 예측도 상승에 기여할

수 있을 것으로 기대된다.

둘째, 기존 선행연구들의 회귀 모델을 통한 예측과 마찬가지로 본 연구의 예측 분석 모델 또한 이 분석 그룹 안에서만 유용하다는 점이다. 따라서, 다른 비만 프로그램을 적용하고 있는 환자들에게 이 분석 모델을 바로 적용하기 어려우며, 분석 방법을 응용하여 적용하려고 하더라도 머신러닝 기법에 능숙한 분석 전문가가 없으면 임상 현장에서 쉽게 만들어서 적용하기 어렵다는 점이다.

셋째, 체중 감량의 예측에서 분류를 적용하는 부분 자체의 한계점도 생각해 볼 수 있다. 분류는 구분의 기준이 명확한 경우 사용되는 것이 좋으나 체중 감량의 성공군이라는 기준은 정해진 cut point를 넘기느냐 아니냐로 정해지는데, 예를 들어, 성공 기준이 초기 체중의 5% 체중 감량이라면 5.1%의 체중 감량은 성공이고, 4.9%의 체중 감량은 실패로 이진 분류되기 때문이다. 본 연구 또한 이분형 분류를 통한 예측을 진행하였기 때문에 이러한 문제에서 자유로울 수는 없다. 다만, 본 연구는 단순 이분형 분류 외에 체중 감량 4분위 감량 기준을 통해 예측 분석 그룹 내에서 다이어트를 진행하면 환자 본인이 어떤 분위 내에 위치할 수 있을지 가능할 수 있는 분석이 가능하여, 그룹 내에서 평균적인 감량보다 양호할 지 저조할 지 보다 세부적인 분석을 원하는 환자에게 추가적인 예측을 제공할 수 있어 그 결과에 따라 좀 더 세분화된 감량 목표 설정 및 중재를 위한 정보를 획득할 수 있겠다.

넷째, 본 연구에서 분석에 사용된 알고리즘 중에서 독보적으로 뛰어난 성능을 나타낸 것은 없었으나 일관되게 비교적 높은 예측 결과 값을 나타낸 것을 살펴보면, 체중 감량 중위 50% 기준으로 속성 선택기법 사용 후 감량을 기준으로 봤을 때 1, 2, 3 차 예측 분석 모두에서 ANN의 정확도가 가장 높았으며 (Table 4), 감량 체중 기준으로는 1, 2, 3차 예측 분석 모두에서 RF의 AUC 값이 가장 높았다 (Table 5). 그러나 ANN과 RF 모두 블랙 박스 알고리즘으로 그

내부를 들여다볼 수가 없어 해석이 어렵다는 한계점을 가지고 있다^{27,37)}. 속성 선택 기법에 사용된 보르타 알고리즘을 통해 선택된 변수들은 통계적 유의미함을 기준으로 하는 것이 아니며, 제시되어 있는 한 변수 중요도 값이 다른 변수의 2배라고 해서 이 변수가 정확히 2배의 영향력을 가지는 것과 별개의 문제일 수 있다. 또한 영향을 미치는 변수의 방향성을 알 수가 없다. 예를 들어, 연령이 중요 변수로 선택되었다고 하더라도 나이가 어릴수록 감량이 잘 되는지, 나이가 많을수록 감량이 잘 되는지에 대한 부분은 여전히 알 수 없는 상태로 남아있게 된다. 추후 체중 감량에 영향을 미치는 변수들에 대한 구체적인 분석을 위해서 LR에서 표준화 회귀계수를 구하여 변수의 음양 방향성을 살펴보거나 분류가 아닌 회귀의 형태로 머신러닝을 진행하여 회귀계수를 살펴보는 방법을 통해 변수들의 영향력의 크기나 방향성에 대하여 분석해보는 추가적인 연구가 필요할 것으로 생각된다.

끝으로, 향후 머신러닝 기법을 적용한 이상 반응 및 중도 탈락 예측 연구를 진행한다면 종합적인 비만 관리 프로그램을 만드는 데 도움을 줄 수 있을 것으로 생각된다. 본 연구에서 초기에는 체중 감량 예측이 어려운 부분을 확인할 수 있었는데, 초진 시에는 예측의 주안점을 향후 발생할 수 있는 이상 반응 예측에 둔다면 약의 용량을 조절해서 선제적 대응을 할 수 있어 중도 탈락을 줄이고 체중 감량 프로그램을 완수하는데 도움을 줄 수 있을 것이다. 뿐만 아니라, 스마트 기기를 통하여 내원을 하지 않더라도 다빈도로 환자의 식이, 운동 정보들을 추적하여 체중 감량 예측의 정확도를 업그레이드해 나간다면 통합 비만 관리에 한 걸음 더 다가설 수 있을 것으로 생각된다.

결론

본 연구에서는 DT, RF, LR, ANN 등의 머신러닝

기법을 적용하여 비만 프로그램을 진행한 과체중 및 비만 성인 환자들의 체중 감량 예측을 총 3회의 가미태음조위탕 약 처방 시점에서 이분형 분류 (binary classification) 형태로 시행하였다. 본 연구는 크게 두 부분으로 나누어 분석을 진행하였는데, 첫 번째 분석은 변수 선택 없이 선행연구를 바탕으로 본 연구에서 추출 가능한 모든 변수를 사용하여 체중 감량 예측을 진행하였으며, 추가로 진행한 두 번째 분석은 중요 변수를 추출하여 이를 바탕으로 체중 감량 예측 모델을 재검토하는 방식으로 진행되었다.

1. 체중 감량의 예측은 가미태음조위탕을 복용한 모집단의 감량률과 감량 체중의 4분위인 상위 25%, 중위 50%, 하위 25%를 분류의 기준점 (classification bench mark)으로 삼았으며, 각 기준점을 넘을 것인가를 맞추는 예측을 시행하여 정확도 (accuracy)를 5-fold cross validation으로 구하여 비교 분석하였다. 이 분석을 통하여, 체중 감량 중위 50%를 분류의 기준점으로 삼았을 때가 상하위 25%를 분류의 기준점으로 삼았을 때 보다 모델의 설명력이 더 좋음을 확인할 수 있었다. 그리고, 1차 (55.01%, 감량률 50%기준 ANN /60.18%, 감량 체중 50%기준 RF)에서 2차 (73.59%, 감량률 50%기준 ANN /75.79%, 감량 체중 50% 기준 ANN), 3차 (84.80%, 감량률 50%기준 RF /86.00%, 감량 체중 50%기준 DT) 예측 분석으로 갈수록 정확도가 상승하는 것을 확인할 수 있었다.
2. 체중 감량 중위 50%를 분류 기준으로 변수 선택 후 변수 중요도를 확인하여 추가 예측 분석을 시행한 결과, 감량률 기준으로 ANN의 정확도가 1차 (54.69%), 2차 (73.52%), 3차 (81.88%) 예측 분석 모두에서 가장 높게 나왔다. AUC를 통해 추가적으로 예측 성능 확인을 진행 하였으며 감량 체중 기준으로 RF의 AUC가 1차 (0.640), 2차 (0.816), 3차 (0.939) 예측 분석 모두에서 가장 높았다.

머신러닝 기법을 적용한 체중 감량 예측은 초기

체중 감량 정보를 활용하여 정확도가 개선됨을 확인할 수 있었으며, 향후 임상에서 본 연구 결과는 가미 태음조위탕에 반응하지 않거나 예측되는 체중 감량치가 낮은 환자들을 선별하여 선제적 증재를 실시하는데 활용할 수 있을 것으로 생각된다.

참고문헌

- Hill JO, Wyatt HR, Peters JC. Energy balance and obesity. *Circulation*. 2012;126(1):126-32.
- Haidar YM, Cosman BC. Obesity epidemiology. *Clin Colon Rectal Surg*. 2011;24(4):205-10.
- Wall KC, Politzer CS, Chahla J, Garrigues GE. Obesity is associated with an increased prevalence of glenohumeral osteoarthritis and arthroplasty: A cohort study. *Orthop Clin N Am*. 2020;51(2):259-264.
- Kolb R, Sutterwala FS, Zhang W. Obesity and cancer: inflammation bridges the two. *Curr Opin Pharmacol*. 2016;29:77-89.
- Handjieva-Darlenska T, Handjiev S, Larsen TM, Baak MA, Jebb S, Papadaki A, et al. Initial weight loss on an 800-kcal diet as a predictor of weight loss success after 8 weeks: the Diogenes study. *Eur J Clin Nutr*. 2010;64(9):994-9.
- Hollis JF, Gullion CM, Stevens VJ, Brantley PJ, Appel LJ, Ard JD, et al. Weight loss during the intensive intervention phase of the weight-loss maintenance trial. *Am J Prev Med*. 2008;35(2):118-26.
- Reed JR, Yates BC, Houfek J, Briner W, Schmid KK, Pullen CH. Motivational Factors Predict Weight Loss in Rural Adults. *Public Health Nurs*. 2016;33(3):232-241.
- Annesi JJ, Whitaker AC, Psychological factors discriminating between successful and unsuccessful weight loss in a behavioral exercise and nutrition education treatment. *Int J Behav Med*. 2010;17(3):168-75.
- Fabricatore AN, Wadden TA, Moore RH, Butryn ML, Heymsfield SB, Nguyen AM. Predictors of attrition and weight loss success: Results from a randomized controlled trial. *Behav Res Ther*. 2009;47(8):685-91.
- Hadziabdic MO, Mucalo I, Hrabac P, Matic T, Rahelic D, Bozиков V. Factors predictive of drop-out and weight loss success in weight management of obese patients. *J Hum Nutr Diet*. 2015;28(2):24-32.
- Batterham M, Tapsell LC, Charlton KE. Predicting dropout in dietary weight loss trials using demographic and early weight change characteristics: Implications for trial design. *Obes Res Clin Pract*. 2016;10(2):189-96.
- Kang EY, Park YB, Kim MY, Park YJ. A Study on Factors Associated with Weight Loss by 'Gamitaeumjowee-Tang'. *J Korean Med Obes Res*. 2017;17(2):68-72.
- Batterham M, Tapsell L, Charlton K, O'Shea J, Thorne R. Using data mining to predict success in a weight loss trial. *J Hum Nutr Diet*. 2017;30(4):471-478.
- Rajkomar A, Dean J, Kohane I. Machine Learning in Medicine. *N Engl J Med*. 2019;380:1347-1358.
- Kim H, Yang SB, Kang Y, Park YB, Kim JH. Machine learning approach to blood stasis pattern identification based on self-reported symptoms. *Korean J Acupunct*. 2016;33(3):102-113.
- Sharma K, Kaur A, Gujral S. Brain tumor

- detection based on machine learning algorithms. *Int J Comput Appl.* 2014;103(1):7-11.
17. Kourou K, Exarchos TP, Exarchos KP, Karamouzis MV, Fotiadis DI. Machine learning applications in cancer prognosis and prediction. *Comput Struct Biotechnol J.* 2015;13:8-17.
 18. Wu CC, Hsu WD, Islam MM, Poly TN, Yang HC, Nguyen PA, et al. An artificial intelligence approach to early predict non-ST-elevation myocardial infarction patients with chest pain. *Comput Methods Programs Biomed.* 2019;173:109-117.
 19. Wang S, Summers RM. Machine learning and radiology. *Med Image Anal.* 2012;16(5):933-51.
 20. Dugan TM, Mukhopadhyay S, Carroll A, Downs S. Machine Learning Techniques for Prediction of Early Childhood Obesity. *Appl Clin Inform.* 2015;6(3):506-520.
 21. Hammond R, Athanasiadou R, Curado S, Aphinyanaphongs Y, Abrams C, Messito MJ, et al. Predicting childhood obesity using electronic health records and publicly available data. *PLoS ONE* 14(4): e0215571. <https://doi.org/10.1371/journal.pone.0215571>
 22. Aswani A, Kaminsky P, Mintz Y, Flowers E, Fukuoka Y. Behavioral Modeling in Weight Loss Interventions. *Eur J Oper Res.* 2019; 272(3):1058-1072.
 23. Kim YM, Cho DG, Kang SH. Analysis of Factors associated with Geographic Variations in the Prevalence of Adult Obesity using Decision Tree. *Health Soc Sci.* 2014;36(1):157-181.
 24. Nam SH, Kim SY, Lim YW, Park YB. Review on predictors of weight loss in obesity treatment. *J Korean Med Obes Res.* 2018;18(2):115-127.
 25. Yoon NR, Yoo YJ, Kim MJ, Kim SY, Lim YW, Lim HH, et al. Analysis of adverse events in weight loss program in combination with 'Gamitaeumjowee-Tang' and low-calorie diet. *J Korean Med Obes Res.* 2018;18(1):1-9.
 26. Kurs MB, Rudnicki WR. Feature Selection with the Boruta Package. *J Stat Softw.* 2010;36(11):1-13.
 27. Jung D, Kim G, Park J, Lee H, Kim H, Choi H, et al. Prediction of rehospitalization of patients and finding causes of it with data mashup and bigdata analysis. *Entrue J Inf Technol.* 2015;14(3):133-149.
 28. Díaz-Uriarte R, Andrés SA. Gene selection and classification of microarray data using random forest. *BMC Bioinformatics.* 2006;7:3. <https://doi.org/10.1186/1471-2105-7-3>
 29. Ortega Hinojosa AM, MacLeod KE, Balmes J, Jerrett M. Influence of school environments on childhood obesity in California. *Environ Res.* 2018;166:100-107.
 30. Munger E, Choi H, Dey AK, Elnabawi YA, Groenendyk JW, Rodante J, et al., Application of machine learning to determine top predictors of noncalcified coronary burden in psoriasis: An observational cohort study". *J Am Acad Dermatol.* 2019. Article in press. <https://doi.org/10.1016/j.jaad.2019.10.060>
 31. Scheinker D, Valencia A, Rodriguez F. Identification of factors associated with variation in US county-level obesity prevalence rates using epidemiologic vs machine learning models. *JAMA Netw Open.* 2019;2(4):e192884.
 32. Forman EM, Kerrigan SG, Butryn ML, Juarascio AS, Manasse SM, Ontañón S, et al. Can the artificial intelligence technique of reinforcement learning use continuously

- monitored digital data to optimize treatment for weight loss? *J Behav Med.* 2019;42(2): 276-290.
33. Hong N, Wen A, Stone DJ, Tsuji S, Kingsbury PR, Rasmussen LV, et al. Developing a FHIR-based EHR phenotyping framework: A case study for identification of patients with obesity and multiple comorbidities from discharge summaries. *J Biomed Inform.* 2019;99:103310 <https://doi.org/10.1016/j.jbi.2019.103310>
34. Han JY, Park YJ. Analysis of factors influencing obesity treatment according to initial condition and compliance with medication. *J Korean Med Obes Res.* 2019;19(1):31-41.
35. Magkos F, Fraterrigo G, Yoshino J, Luecking C, Kirbach K, Kelly SC. Effects of Moderate and Subsequent Progressive Weight Loss on Metabolic Function and Adipose Tissue Biology in Humans with Obesity. *Cell Metab.* 2016; 23(4):591-601.
36. Jo GW, Ok JM, Kim SY, Lim YW. Review on the Efficacy and Safety of Mahuang and Ephedrine in the Treatment of Obesity -Focused on RCT-. *J Korean Med.* 2017;38(3): 170-184.
37. Disse E, Ledoux S, Bétry C, Caussy C, Maitrepierre C, Coupaye M, et al. An artificial neural network to predict resting energy expenditure in obesity. *Clin Nutr.* 2018;37(5): 1661-1669.

ORCID

- 김은주 <https://orcid.org/0000-0003-3708-3690>
박영배 <https://orcid.org/0000-0001-6025-1731>
최가혜 <https://orcid.org/0000-0002-7758-7888>
임영우 <https://orcid.org/0000-0002-3039-7307>
옥지명 <https://orcid.org/0000-0002-5453-2948>
노은영 <https://orcid.org/0000-0001-8111-7785>
송태민 <https://orcid.org/0000-0002-2078-4119>
강지훈 <https://orcid.org/0000-0003-3784-5958>
이향숙 <https://orcid.org/0000-0002-1107-9389>
김서영 <https://orcid.org/0000-0002-6031-850X>