

국민건강영양조사를 활용한 대사증후군 유병 예측모형 개발을 위한 융복합 연구: 데이터마이닝을 활용하여

김한결*, 최근호**, 임성원***, 이현실****

고려대학교 대학원 보건과학과 BK21 플러스 인간생명-상호작용 융복합 사업단*, 근로복지공단 근로복지연구원 조사연구부 통계분석2팀장 최근호 (책임연구원)**, 고려대학교 일반대학원 보건과학과***, 보건과학대학 보건정책관리학부****

Development of Prediction Model for Prevalence of Metabolic Syndrome Using Data Mining: Korea National Health and Nutrition Examination Study

Han-Kyoul Kim*, Keun-Ho Choi**, Sung-Won Lim, Hyun-Sill Rhee****

Department of Public Health Science, Graduate School BK21 Plus Program in Public Health Science, Korea University*

Korea Worker's Compensation & Welfare Service, Labor Welfare Research Institute, Research Department**
School of Health Policy and Management, College of Health Science, Korea University***

요약 이 연구의 목적은 국민건강영양조사 2012년 자료 중 40세 이상 성인의 대사증후군 유병 여부를 예측에 영향을 미치는 변수를 확인하고 이를 예측하는 모형 개발하는데 있다. 선행연구를 통해 모델 생성에 필요한 투입변수를 선정하였다. 연구결과 투입변수 중 사회경제적 요인이 상위 순위에 해당하였으며, 건강행위 요인의 경우 하위 순위로 나타났다. 또한, 최종 예측모형은 의사결정나무 (Decision Tree)일 경우 90.32%의 가장 높은 예측력을 나타내고 있었다. 이 연구의 결과는 다음과 같은 시사점을 나타낸다. 먼저, 대사증후군에 대한 예방 및 관리에 있어 건강행위에 대한 접근과 함께 사회경제적 요인에 대한 접근도 병행해 고려해야 한다. 또한, 의사결정나무 알고리즘의 경우 결과해석의 용이성이 있어 보건의료분야에서 많이 사용되며, 선행연구의 결과와 마찬가지로 높은 예측정확도를 나타내고 있다.

주제어 : 대사증후군, 속성선택, 데이터 마이닝, 의사결정나무, 로지스틱 회귀분석, 인공신경망, 국민건강영양조사

Abstract The purpose of this study is to investigate the attributes influencing the prevalence of metabolic syndrome and develop the prediction model for metabolic syndrome over 40-aged people from Korea Health and Nutrition Examination Study 2012. The researcher chose the attributes for prediction model through literature review. Also, we used the decision tree, logistic regression, artificial neural network of data mining algorithm through Weka 3.6. As results, social economic status factors of input attributes were ranked higher than health-related factors. Additionally, prediction model using decision tree algorithm showed finally the highest accuracy. This study suggests that, first of all, prevention and management of metabolic syndrome will be approached by aspect of social economic status and health-related factors. Also, decision tree algorithms known from other research are useful in the field of public health due to their usefulness of interpretation.

Key Words : Metabolic syndrome, Feature selection, Data mining, Decision tree, Logistic regression, Artificial neural network, KHNES

* 본 논문은 2013년 고려대학교 보건과학대학 특별연구비에 의하여 지원되었음

Received 1 January 2016, Revised 30 January 2016

Accepted 20 February 2016, Published 28 February 2016

Corresponding Author: Hyunsill Rhee(Department of Public Health Science, Graduate School BK21 Plus Program in Public Health Science, Korea University)

Email: pridehyun@korea.ac.kr

ISSN: 1738-1916

© The Society of Digital Policy & Management. All rights reserved. This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. 서론

지난 1세기 동안 공중보건 분야는 커다란 변화가 겪었다. 산업화와 근현대화는 인류의 영양 및 주거 상태의 향상, 첨단의료기술을 발달시켰으며 전 인류의 평균수명은 급격하게 증가시켰다. 기대수명의 증가는 인구집단의 질병양상에도 변화를 가져왔다. 과거 전염성이 높고 급성기적인 특성을 보인 질환은 현대사회로 접어들며 비전염적이지만 만성적으로 존재하는 특징을 보이고 있다. 비전염적 만성질환은 갑작스런 치료 및 관리가 필요한 급성 질환과 달리 지속적이며 장기적인 관리와 치료가 필요한 질환이다. 이런 비전염성 만성질환 중 가장 대표적인 질환은 대사증후군 (metabolic syndrome)이다.

대사증후군이란, 주로 심혈관계 질환을 유발하는 것으로 알려진 5가지의 위험요인들이 한 개인에게 동시 다발적으로 발생하는 것을 말한다. American Heart Association [16]에서는 대사증후군을 신진대사의 위험요인들의 집합이며 한 개인에게 다음의 위험요인들이 함께 발현될 때 하나의 위험요인에 노출되는 경우에 비해 심혈관계에의 문제가 있을 가능성이 높아진다고 정의하고 있다. 또한, 대사증후군은 하나의 만성적 질환이 독립적으로 존재하지 않는 특성을 보이고 있다. 즉, 개인에게 있어 2개 이상의 질환이 동시에 보이는 복합적인 만성적 상태 (multiple chronic condition)을 보이는 특성이 있다.

우리나라의 경우 최근 5년간 대사성 질환으로 진단 환자의 수가 지속적으로 증가하고 있는 추세이며, 이에 대한 진료비 및 급여비도 증가하고 있다. 또한, 대사성 질환은 관련된 5개 질환 외에 심혈관계 및 내분비계 질환을 추가적으로 야기할 수 있는 위험을 지니고 있다. 수술 후 환자에게 발생할 수 있는 합병증에 관해 대사증후군이 미치는 영향에 대해 추적 연구를 진행한 연구에 따르면, 대사증후군을 보유한 집단에서 더딘 회복과 함께 수술 후 감염의 문제를 더 많이 보이는 것으로 나타났다 [12]. 정은순 [1]의 연구에 따르면, 비만으로 인한 대사성질환은 여성의 경우에 우울과 연관성이 있을 수 있다는 결과를 도출했다. 이는 대사증후군의 영향이 의료비용 측면 이외에 개인의 생활 및 심리적 상태에 까지 영향을 미칠 수 있다는 것을 시사한다. 이처럼 대사증후군은 보건의료비의 지출 증가뿐 아니라 유병 환자의 신체적, 정신적인 건강에의 위협을 가하고 있다. 즉, 대사증후군의 위험

에 놓인 사람을 미리 예측하여 미리 예방하는 것은 2차적인 질환으로의 전개를 미연에 방지하고 보건의료비 지출을 경감시키는데 매우 효과적인 방법이다[15]. 그러나 현재 우리나라의 대사증후군 유병과 관련된 연구들은 관련 요인을 탐색하는 연구에 그치고 있다. 따라서 본 연구는 대사증후군 유병 여부를 예측하는 모형을 개발하고자한다.

2. 연구방법

2.1 자료원 및 연구대상

본 연구는 국민건강증진법 16조에 근거하여 국민의 건강 및 영양 상태를 파악하기 위해 실시되고 있는 국민건강영양조사 자료를 활용하였다. 국민건강영양조사의 목적은 국민의 건강수준, 건강관련 의식 및 형태, 식품 및 영양섭취 실태에 대한 국가 및 시·도 단위의 대표성과 신뢰성을 갖춘 통계를 산출하는 것이다. 본 연구의 연구대상은 대사성 질환에 노출되기 시작하는 40세 이상의 성인남녀로 결측값 및 이상값을 제거한 후 최종 11,424명을 연구 대상으로 선정하였다.

2.2 변수

2.2.1 타겟 변수

American Heart Association [16]에 따르면, 대사증후군을 진단하기 위한 기준은 복부비만의 경우 남성은 40인치, 여성은 35인치 이상인 경우를 이상으로 진단하며, 이상지혈증은 triglycerides 수치가 150mg/dL 이상, HDL-콜레스테롤수치는 남자는 40mg/dL, 여성의 경우 50mg/dL 이상인 경우를 이상으로 진단한다. 또한, 혈압은 수축기 혈압이 130/85mg/dL보다 높으며, 공복 시 혈당은 100mg/dL 이상인 경우를 신진대사 이상으로 진단한다. 이 중 대사증후군은 3가지 부분에서 이상 진단을 받았을 때를 말한다. 따라서 본 연구의 목표 변수는 대사증후군 관련 5개 질환 중 2개 이하를 보유한 사람의 경우 '0', 3개 이상의 질환을 보유한사람을 '1'로 코딩하였다.

2.2.2 투입 변수

이 연구에서 사용된 투입변수는 다음과 같다. 첫째, 사회경제적 수준을 반영하는 성별, 연령, 교육수준, 결혼상

(Table 1) Definition of Variables

Variables		Definition	Type of variable
Target variable	Metabolic syndrome	Normal, Metabolic syndrome	Nominal
	Sex	Male, Female	Nominal
	Age	Age of participants	Continuous
	Education level	Under elementary Middle school High school Over university	Nominal
	Marital status	Single Married Etc (brevement, divorced)	Nominal
	Household income	4-quartile	Nominal
	Smoking	Non-smoker Past smoker Current smoker	Nominal
	Drinking	Non-drinker Sometimes per month Sometimes per 2-weeks Sometimes per week	Nominal
Input variables	Type of insurance	National Health Insurance Assistance	Nominal
	Private health insurance	No Yes	Nominal
	Economic activity	No Yes	Nominal
	Physical activity	Never exercise Few per 2-weeks Sometimes per 2-week	Nominal
	Time of sleep		continuous
	Perceived body shape	Very skinny Somewhat skinny Average Somewhat heavy Very heavy	Nominal

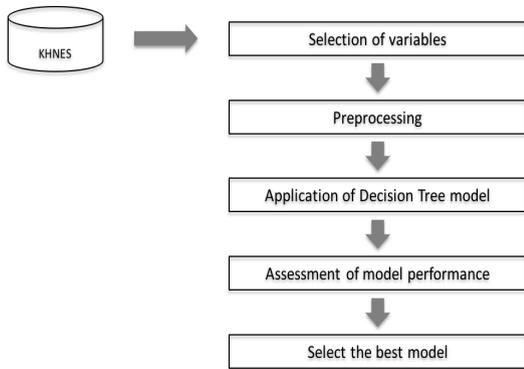
태 및 가구의 소득수준을 포함하였다. 또한, 대사증후군의 경우 건강행위와 밀접한 관련을 보이기 때문에 흡연 및 음주, 지난 2주간 신체활동 정도, 수면시간, 본인이 인지하는 체형을 투입하였다. 투입변수들은 국민건강영양조사에서 제공하는 변수의 형태를 변형없이 사용하였으며, 흡연 및 음주, 2주간 신체활동정도의 경우 노출 및 실행의 정도에 따라 명목형 변수로 변환하여 사용하였다. 본 연구에서 투입한 변수 및 설명은 <Table 1>과 같다.

2.3 분석방법

2.3.1 연구절차

본 연구의 목적인 대사증후군 유병 여부를 예측하기 위한 모형을 개발하기 위한 연구의 절차는 [Fig. 1] 과 같다. 국민건강영양조사 2012년 자료에서 목표 변수인 대

사증후군유병 및 투입 변수인 11개의 변수를 선택하였다. 다음으로 데이터 전처리 과정을 통해 결측값 및 이상치를 제거하였으며, 분석의 목적에 맞게 목표변수와 독립변수를 각각 변환하였다. 또한, 전처리과정의 일환으로 데이터 마이닝의 기법 중 하나인 균형 작업(balancing)을 진행하였다. 균형 작업은 목표 변수를 분포가 하나의 값에 편중되어 있는 경우 정확한 모델의 생성을 위해 값의 분포를 균일하게 만들어주는 기법이다. 다음으로 예측모형 개발 및 평가에는 일반화를 위해 전체 데이터를 5개로 구분하여 모델 생성 및 검증 데이터로 사용하는 5-fold 교차타당도 (5-fold cross validation) 방법을 사용하였다. 다음으로, 만들어진 모델들의 정확도를 평가한 후 최적의 모델을 선택하였다.



[Fig. 1] The process of study

2.3.2 속성선택

문헌고찰을 통해 본 연구에 투입된 독립변수는 최종 13개 이다. 그러나 모든 투입 변수가 목표 변수에 동일한 예측력을 보이지 않는다. 따라서 본 연구는 목표 변수에 대한 투입 변수의 영향력을 순위로 확인하기 위한 속성 선택 기법을 적용하였다. 속성 선택 기법은 다수의 투입 변수들 중 목표 변수의 예측에 중요한 역할을 하는 변수를 선정하는 방법이다. 이를 통해 투입 변수들 중에서 목표변수를 예측하거나 평가할 때 도움을 주지 못하는 변수는 제거한 후 예측 모형을 개발하는 것이 바람직하다 [14]. 왜냐하면, 위 변수들은 목표 변수와 큰 상관이 없음에도 불구하고 예측 모형에 포함될 경우 모형의 정확도 및 성능에 부정적인 영향을 줄 수 있기 때문이다[11].

2.3.3 분류모델

본 연구의 목적인 대사증후군 유병 여부 예측 모형을 개발하기 위한 데이터 마이닝 도구로서 Wakito university에서 오픈 소스로 제공하는 Weka 3.6을 사용하였다. 다음으로 예측 모형을 생성하기 위한 알고리즘으로 의사결정나무, 로지스틱 회귀분석, 인공신경망을 사용하였다. 의사결정나무는 각각의 노드에 변수 및 속성에 대한 값을 제시함으로써 규칙을 도식화 해주는 기법으로 사용자가 의사결정 방법에 대해 쉽게 이해할 수 있는 장점을 지닌다. 로지스틱 회귀분석은 목표 변수의 형태가 범주화된 이분형인 자료에서 많이 사용되는 알고리즘으로 0/1의 값만을 취하는 자료에서 이를 보완해주는 장점이 있다. 마지막으로, 인공 신경망은 불안정한 형태의 자료에서 가중치를 적용하여 보완해주는 장점이 존재

하는 반면 은닉층의 존재로 인해규칙에 대한 설명이 불가능하다는 단점이 있다.

3. 연구결과

3.1 연구대상자의 일반적 특성

연구 대상자의 일반적 특성을 확인하기 위한 기술통계분석의 결과는 <Table 2>와 같다. 여성이 남성에 비해 많이 포함되어 있었으며, 평균 연령은 58.4세로 비교적 높은 편이었다. 또한, 초졸 이하의 학력을 지닌 사람이 35.2%로 가장 많았으며, 고졸 및 대졸 이상이 49.7%를 차지하고 있었다. 대부분의 대상자들이 기혼 (83.3%)이었으며, 가구소득은 비교적 균일하게 분포되어 있었다. 또한, 우리나라의 특성상 대부분의 연구대상자가 국민건강보험에 가입되어있는 반면, 민간의료보험의 경우 65.8%만이 가입되어 있었으며, 연구대상자의 약 60%는 경제활동에 참여를 하고 있었다. 흡연 및 음주를 살펴보면, 음주의 경우 평생 음주경험이 없는 사람이 17.7%에 그치는 반면, 흡연의 경우 평생 흡연경험이 없는 사람이 58.8%로 매우 많았다. 운동의 경우 2주에 약간 운동을 하는 집단이 67.5%로 많았으며, 평균 수면시간은 약 6.7시간 정도로 나타났다. 마지막으로, 연구 대상자가 인지하는 자신의 체형에 대해서는 일반적으로 보통 (41.5%)이라고 응답한 사람이 가장 많았으며, 다소 뚱뚱하다고 응답한 집단도 33.9%로 다음으로 많이 나타났다.

3.2 속성선택 결과

본 연구의 목적인 대사증후군 유병여부를 예측하기 위한 투입변수는 13개에 대한 속성선택 결과는 <Table 3>과 같다. 먼저, 조사 대상자가 인지하는 자신의 체형인식이 가장 순위가 높은 변수로 나타났다. 다음으로, 교육 수준 및 결혼상태, 성별, 연령과 같은 사회경제적 지위와 관련된 변수들이 상위권에 위치하였다. 반면, 음주 및 흡연, 수면시간, 운동여부와 같은 건강행태와 관련된 요인은 사회경제적인 지표들에 비해 낮은 순위에 위치하였다. 본 연구는 목표변수에 대한 투입변수들의 다음과 같은 결과를 토대로 가장 관련이 낮은 변수를 차례로 제거하며 예측모형을 개발하였다. 위 방법은 모든 변수들이 투입된 모형으로부터 단계적으로 관련이 낮은 변수를 제

거해가며 최적의 모형을 찾아나가는 방법으로 Wrapper의 단계적 제거방법 (wrapper's stepwise backward elimination)이라고 한다.

〈Table 2〉 General characteristics of research population

Input variables		N	(%)
Sex	Male	4,944	(43.6)
	Female	6,486	(56.4)
Age		58.4	(11.5)
Education level	Under elementary	3,985	(35.2)
	Middle	1,708	(15.1)
	High	3,368	(29.7)
Marital status	Over graduate	2,269	(20.0)
	Single	171	(1.5)
	Married	9,438	(83.3)
Household income	Etc	1,721	(15.2)
	Low	2,704	(23.9)
	Middle low	2,863	(25.3)
Type of insurance	Middle high	2,732	(24.1)
	High	3,032	(26.8)
	NHI	10,997	(97.1)
PHI	Assistance	333	(2.94)
	Yes	7,450	(65.8)
Economic activity	No	3,880	(34.2)
	Yes	6,709	(59.2)
Drinking	No-drinker	4,621	(40.8)
	Sometimes/month	2,003	(17.7)
	Sometimes/2weeks	3,859	(34.1)
Smoking	Sometimes/week	3,007	(26.5)
	Non-smoker	2,461	(21.7)
	Past smoker	6,666	(58.8)
Physical activity	Current smoker	2,631	(23.2)
	Never exercise	2,033	(17.9)
	Few/2-weeks	1,556	(13.7)
Time of sleep	Sometimes/2weeks	7,648	(67.5)
	Very skinny	2,126	(18.8)
Perceived body shape	Somewhat skinny	6.7	(1.4)
	Average	597	(5.3)
	Somewhat heavy	1,373	(12.1)
	Very heavy	4,701	(41.5)
Household income		3,839	(33.9)
		820	(7.2)

〈Table 3〉 The Result of Feature Selection

Variables	Ranking	Variables	Ranking
Perceived body shape	1	Type of insurance	8
Education level	2	Drinking	9
Marital status	3	Time of Sleep	10
Sex	4	Physical activity	11
Age	5	Economic activity	12
Private health insurance	6	Smoking	13
Household income	7	-	-

3.3 분류기법 결과

본 연구에서 사용한 분류기법은 의사결정나무, 로지스틱 회귀분석, 인공신경망이다. <Table 4>는 Wrapper의 단계적 제거방법을 적용한 분류분석의 결과이다. 분석결과 의사결정나무 알고리즘이 다른 두 알고리즘에 비해 전반적으로 더 정확한 예측력을 보이고 있었다. 로지스틱 회귀분석의 경우 예측정확도가 65% 정도로 미미하였으며, 인공 신경망의 경우 변수의 수가 높을수록 예측정확도가 상승하는 경향을 보이고 있으나, 13개의 변수가 모두 투입되었을 때에도 약 70% 정도의 적중률을 보이고 있다. 이에 반해, 의사결정나무의 경우 투입된 변수의 수가 적은 경우는 다른 두 알고리즘과 큰 차이를 보이고 있지 않으나, 투입변수의 순위가 5위인 변수까지 포함되는 지점에서 예측 정확도가 약 70% 정도로 상승한다. 또한, 투입변수의 수가 늘어남에 따라, 지속적으로 증가하는 추세를 보이며 13개의 변수가 모두 투입되었을 경우 약 89%의 높은 예측 정확도를 나타내고 있다.

〈Table 4〉 Accuracy of the models

No.	DT	ANN	LR	No.	DT	ANN	LR
1	63.3	63.3	63.3	8	77.9	66.0	64.9
2	63.8	63.6	64.0	9	83.2	67.3	65.0
3	65.3	64.2	63.9	10	87.1	67.8	64.9
4	66.1	65.6	64.5	11	88.2	69.6	64.8
5	70.3	65.1	64.8	12	88.5	68.8	64.9
6	72.5	65.6	65.2	13	88.8	70.2	65.1
7	77.4	65.5	64.9	-	-	-	-

4. 논의

본 연구는 국민건강영양조사 자료를 바탕으로 대사증후군의유병에 관련이 있는 변수들을 파악하고, 이를 기반으로 유병 여부를 예측하는 모형을 개발하고자 했다.

이론적 고찰을 토대로 선정된 투입변수들을 토대로 속성 선택 (Feature selection)을 시행한 결과, 교육수준 및 결혼상태, 성별 등의 사회경제적 지위를 반영하는 요인들이 우선적으로 나타났으며 단계적 제거방법 (wrapper's stepwise backward elimination)며, 음주 및 흡연, 운동 등의 건강 행위와 관련된 변수들은 하위의 순위에 나타났다.

일반적으로 질환의 이환과 관련하여 개인 및 집단의 사회적 계급 및 지위가 더 많이 연관되어 있다. 사회경제적 지위에 따른 건강 행동 참여에 관한 Haught, Rose, and Brown [6]의 연구를 보면, 소득 및 교육수준이 높을수록 건강 행동에의 참여가 더 나타나는 결과를 나타냈다. 또한, 브라질 도시 인구에 대한 대사증후군 연구에 따르면, 대사증후군을 보유한 사람의 경우 경제적으로 빈곤층에 속하며 고연령 추세가 나타나고 있음을 보였다 [4]. 한국의 성인 남성과 여성에 대한 교육수준과 대사 질환 간의 관련성에 대한 연구에 따르면, 남성과 여성에서 교육수준에 따른 대사 질환의 유병이 다르게 나타난다는 결과를 보였다[10]. 대사증후군의 유병에 있어 사회적 지위가 건강행동보다 선행된다는 것은 생애주기 관점의 맥락에서도 나타난다. Ramsay, Whincup, Morris, Lennon, and Wannamethee [13]의 연구에 따르면, 어린 시절의 낮은 지위가 대사증후군의 유병과 높은 관련성을 보이고 있었으며 성인기의 사회경제적 지위는 부분적으로 건강 행동과 함께 대사증후군과의 연관성을 보이고 있었다. 이는 건강 행동이 기타 질환에 있어 일차적인원인으로 작용할 수 있으나, 건강을 지키기 위한 행동에 앞서 사회경제적인 요인들이 더 작용하고 있다는 것으로 생각해 볼 수 있다. 즉, 본 연구의 속성 선택 결과에서 사회경제적 요인에 해당하는 교육수준, 결혼상태, 성별, 연령이 상대적으로 높은 순위에 위치한 반면, 음주 및 흡연, 수면시간, 운동 여부의 건강 행위 관련 요인이 낮은 순위에 위치한 것은 사회경제적 지위가 건강 행위보다 선행하는 영향이 반영된 것을 보인다. 본 연구의 속성 선택 결과는 대사증후군을 포함한 만성질환 관리에 대한 접근에 있어 다른 관점을 제공한다. 현재 만성질환의 관리에 관한 연구는 대부분 개인의 건강 행위 실천과 관련하여 이루어지고 있다 [7]. 그러나 본 연구의 속성 선택 결과와 선행 연구는 개인의 건강 행위에 앞서 사회경제적 지위가 영향을 미치는 것으로 볼 수 있다. 따라서 개인의 건강증진

행위를 장려하는 정책과 함께 장기적으로 대사증후군 집단의 사회경제적인 지위를 향상시키기 위한 노력이 병행되어야 한다. U-health system을 통해 대사증후군에 대한 체계적이고 지속적인 관리를 모색한 연구를 살펴보면, 신체의 실제수치에 대한 즉각적인 피드백과 함께 필요로 하는 신체활동 등에 대한 알람서비스를 통해 관리가 될 수 있다고 보고하고 있다[5, 9].

다음으로, 속성 선택의 순위 결과를 기반으로 분류기법을 적용한 결과 의사결정나무 알고리즘을 사용하였을 경우, 로지스틱 회귀분석 및 인공 신경망 알고리즘을 사용한 경우에 비해 더 높은 예측정확도를 보이고 있었다. 기존의 선행연구도 데이터 마이닝의 의사결정나무 알고리즘은 보건의료분야의 자료를 예측하는데 있어 유용한 자료임을 나타내고 있다[2, 3, 8]. 따라서 본 연구의 결과는 기존의 연구결과와 동일한 결과를 나타내고 있다고 할 수 있다. 그러나 기존의 대사증후군에 관한 연구와 다르게 본 연구는 사회경제적 요인과 함께 건강 행위 관련 요인을 투입 변수로 사용하였으며, 다양한 알고리즘을 적용함으로써 모델의 예측정확도를 비교하였는데 그 의의가 있다. 또한, 속성선택의 결과를 기반으로 변수를 제거해 나가는 단계적 후진 제거법을 사용하여 예측도의 변화를 확인하였다.

본 연구는 다음과 같은 한계점을 지니고 있다. 먼저, 국민건강영양조사는 인구 총조사를 기반으로 표본을 추출하여 전 국민의 대표성을 지닌다는 장점이 존재하는 반면 조사대상자의 응답을 요구하기 때문에 회상편의 오류가 존재한다. 두 번째로 국민건강영양조사는 단면조사 연구로서 대사증후군의 유병을 야기하는 영향요인에 대해 시간의 순서를 반영할 수 없다. 향후 연구는 추적 연구를 통한 잠재적인 영향요인을 규명할 필요가 있다.

5. 결론

본 연구는 선행연구 고찰을 통해 대사증후군과 관련된 요인을 사회경제적 요인과 건강 행위 요인을 선별하고, 이를 토대로 대사증후군 유병 여부를 예측하는 모형을 개발하였다. 연구 결과 대사증후군에 영향을 미치는 변수로는 건강 행위 관련 인자에 비해 사회경제적 지위가 더 높은 순위로 나타났다. 이는 현재 대사증후군을 포

함한 만성질환관리에 있어 개인의 건강 행위에 초점을 둔 접근에서 사회경제적인 분야에서의 접근이 필요함을 시사한다.

ACKNOWLEDGMENTS

This study was funded by the special research grant of the Korea University Research Institute of Health Science.

REFERENCES

- [1] E. S. Chung, An analysis of the Relationship between Depression and Menopausal Syndromes in Women at Mid life. *Korean J Women Health Nurs*, Vol. 3, No. 2. pp.230-240, 1997.
- [2] F. Firouzi, M. Rashidi, S. Hashemi, M. Kangavari, A. Bahari, N. E. Daryani, M. M. Emam, N. Naderi, H. M. Shalmani, A. Farnood, M. Zali, A decision tree-based approach for determining low bone mineral density in inflammatory bowel disease using PWEKA software. *European journal of gastroenterology & hepatology*, Vol. 19, No. 12, pp. 1075-1081, 2007.
- [3] F. S. de Edelenyi, L. Goumidi, S. Bertrais, C. Phillips, R. MacManus, H. Roche, R. Planells, D. Lairon, Prediction of the metabolic syndrome status based on dietary and genetic parameters, using Random Forest. *Genes&nutrition*, Vol. 3, No. 3-4, pp. 173-176, 2008.
- [4] G. F. Marquezine, C. M. Oliveira, A. C. Pereira, J. E. Krieger, J. G. Mill, Metabolic syndrome determinants in an urban population from Brazil: social class and gender-specific interaction. *International journal of cardiology*, Vol. 129, NO. 2, pp.259-265, 2008.
- [5] G. J. Kim, J. S. Han, Chronic Disease Management using Smart Mobile Device, *JOURNAL OF DIGITAL CONVERGENCE*, Vol. 12, No. 4, pp. 335-342, 2014.
- [6] H. M. Haught, J. P. Rose, J. A. Brown, Social-class indicators differentially predict engagement in prevention vs. detection behaviours. *Psychology&Health*, Vol. 31, No. 1, pp. 21-39, 2016.
- [7] H. S. Chang, A Study on Weight Control Behaviour, Eating Habits and Health-related Life Habits According to Obesity Degree of Teacher in Jeonbuk Province, Korea. *Journal of The Korean Society of Dietary Culture*, Vol. 30, No.1 , pp.105-117. 2015.
- [8] J. Quentin-Trautvetter, P. Devos, A. Duhamel, R. Beuscart, Assessing association rules and decision trees on analysis of diabetes data from the DiabCare program in France. *Studies in health technology and informatics*, Vol. 90, pp.557-561. 2001.
- [9] J. W. Chang, S. C. Sung, The effect of applying u-health system on metabolic syndrome management of elderly, *JOURNAL OF DIGITAL CONVERGENCE*, Vol. 11, No. 11, pp. 553-560, 2013.
- [10] K. D. Ko, B. Cho, W. C. Lee, H. W. Lee, H. K. Lee, & B. J. Oh, Obesity Explains Gender Differences in the Association Between Education Level and Metabolic Syndrome in South Korea The Results From the Korean National Health and Nutrition Examination Survey 2010. *Asia-Pacific Journal of Public Health*, Vol. 27, No. 2, pp.630-639, 2015.
- [11] M. Dash, H. Liu, Feature selection for classification. *Intelligent data analysis*, Vol. 1, No. 3, pp. 131-156, 1997.
- [12] M. J. Gage, R. Schwarzkopf, M. Abrouk, J. D. Slover, Impact of metabolic syndrome on perioperative complication rates after total joint arthroplasty surgery. *The Journal of arthroplasty*, Vol. 29, No. 9, pp.1842-1845, 2014.
- [13] S. E. Ramsay, P. H. Whincup, R. Morris, L. Lennon, S. Wannamethee, Is socioeconomic position related to the prevalence of metabolic syndrome? Influence of social class across the life course in a population-based study of older men. *Diabetes care*, Vol. 31, No. 12, pp.2380-2382, 2008.
- [14] S. H. Chung, Y. M. Suh, Development of a Medial

Care Cost Prediction Model for Cancer Patients Using Case-Based Reasoning, Asia Pacific Journal of Information Systems, Vol. 16, No. 2, pp. 69-84, 2006.

- [15] Y. S. Seo, S. H. Kang, A Convergence Study in the Severity-adjusted Mortality Ratio on inpatients with multiple chronic conditions, JOURNAL OF DIGITAL CONVERGENCE, Vol. 13, No. 12, pp. 245-257, 2015.
- [16] American Heart Association, About Metabolic Syndrome, American Heart Association Home page, 2014, http://www.heart.org/HEARTORG/Conditions/More/MetabolicSyndrome/About-Metabolic-Syndrome_UCM_301920_Article.jsp, May 14.
- [17] Donghyun Kim, Seoksoo Kim, "Design of Key Tree-based Management Scheme for Healthcare Information Exchange in Convergent u-Healthcare Service ", Journal of the Korea Convergence Society, Vol. 6, No. 6, pp. 81-86, 2015.
- [18] Young-Sook Kwon, "Necessity of the Development of a Web-based Obesity Management Program to Prevent Metabolic Syndrome of the Workers", Journal of the Korea Convergence Society, Vol. 5, No. 4, pp. 121-127, 2014.

김 한 결(Kim, Han Kyoul)



- 2012년 2월 : 건양대학교 병원관리학과 (경영학사)
- 2014년 6월 : 고려대학교 일반대학원 보건과학과 보건행정학 전공 (보건학 석사)
- 2014년 9월 ~ 현재 : 고려대학교 보건정책관리학 전공 박사과정
- 관심분야 : 보건학, 데이터마이닝, 간병서비스
- E-Mail : collkhk@korea.ac.kr

최 근 호(Choi, Keun Ho)



- 2008년 2월 : 인하대학교 일반대학원 경영학과 MIS전공 (경영학 석사)
- 2013년 2월 : 고려대학교 일반대학원 경영학과 MIS전공 (경영학 박사)
- 2014년 4월 ~ 현재 : 근로복지공단 근로복지연구원 책임연구원
- 관심분야 : 경영학, 데이터마이닝
- E-Mail : ckh0515@hanmail.net

임 성 원(Lim, Sung Won)



- 2010년 2월 : 경희대학교 간호학사
- 2013년 6월 : 고려대학교 일반대학원 보건과학과 보건행정학 전공 (보건학 석사)
- 2013년 7월 ~ 2015년 2월 : 한국보건 의료연구원 연구원
- 관심분야 : 의료정보, 건강증진, 지역사회 참여적 건강증진
- E-Mail : staroot1@gmail.com

이 현 실(Rhee, Hyun Sill)



- 1990년 2월 : 서울대학교 보건대학원 (보건학 석사)
- 2005년 2월 : 경상대학교 보건대학원 (보건학 박사)
- 1994년 3월 ~ 현재 : 고려대학교 보건과학대학 보건행정학 교수
- 관심분야 : 보건정보관리, 의료질관리, 개인건강기록관리
- E-Mail : pridehyun@korea.ac.kr