

# 상황인지 환경 기반 유헬스 서비스의 추천 요인 식별 및 의사결정 모델 생성

김재권\*, 이영호\*\*

인하대학교 컴퓨터정보공학과, 가천대학교 IT대학\*\*

## Context Aware Environment based U-Health Service of Recommendation Factors Identity and Decision-Making Model Creation

Jae-Kwon Kim\*, Young-Ho Lee\*\*

Dept of Computer and Information Engineering, Inha University\*

Dept. of IT College, Gachon University\*\*

**요약** 상황인지 환경의 유헬스 서비스는 환자가 실생활에 접촉할 수 있는 여러 상황에 대해 컴퓨터가 인지하여 건강 서비스를 제공하는 것이다. 상황인지 환경의 서비스를 추천하기 위해서는 상황 데이터의 정의와 서비스 추천 요인과 관련이 있는지를 식별해야 한다. 본 논문에서는 상황인지 환경의 유헬스 서비스를 제공하기 위해 상황 데이터에 대한 추천 요인들을 다변량 분석기법을 이용하여 식별하며, 의사결정 트리 및 연관성 규칙 기반의 의사결정 모델을 생성한다. 추천 요인의 식별을 통해서 건강 서비스 제공에 유의한 상황 데이터를 판별할 수 있다. 또한 선호도 의사결정 모델을 통해 환자의 상황 데이터에 따라 선호 요인을 알 수 있다.

**주제어** : 유헬스, 상황인지, 추천 요인 식별, 추천 의사결정 모델, 의사결정 트리, 연관성 규칙

**Abstract** Context aware environment u-health service is to provide health service with recognition of a computer. The computer recognizes that a patient can contact real life in many context. Context aware environment service for recommend have to definition of context data and service recommendations related to factors shall be identified. In this paper, Context aware environment of u-health service will be provide context data related to identifies recommendations factors using multivariate analysis method and recommendations factors creation to decision tree, association rule based decision model. health service recommend for significantly context data can be distinguish through recommendation factors of identify. Also, context data of patient can know preference factors through preference decision model.

**Key Words** : U-Health, Context Aware, Recommendation Factors Identity, Recommendation Decision Model, Decision Tree, Association Rule

\* 본 논문은 Ministry of Health and Welfare, Republic of Korea (A112020)의 일환으로 수행하였음

Received 15 March 2013, Revised 25 April 2013

Accepted 20 May 2013

Corresponding Author: Young-Ho Lee(Dept. of IT College, Gachon University)

Email: lyh@gachon.ac.kr

© 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

유헬스 서비스는 언제, 어디서든지 환자의 건강상태와 생체 신호정보를 수집하여 서비스를 제공하는 개념으로서, 유헬스 서비스의 환경에 맞는 시스템 및 미들웨어의 개발이 활발하다[1]. 유헬스 서비스는 환자의 질병의 치료를 목적으로 개발되고 있으며, 건강증진을 위해서는 여러 상황에 맞는 건강 서비스를 제공해야 한다.

최근에는, 유헬스 서비스 구현을 위한 상황인식 컴퓨팅 기반의 질 높은 의료서비스 제공에 대한 관심이 높아지고 있으며, 환자의 여러 상황에 대응할 수 있는 상황인식 컴퓨팅은 지능적인 상황정보의 관리로 서비스의 만족도를 높일 수 있다[2]. 또한 유헬스 환경에서의 상황정보의 활용은 환자와 관련된 모든 상황데이터를 추론하여 환자에게 질 높은 건강 서비스를 제공하는 것을 목적으로 하기 때문에 능동적이고 지능적인 서비스의 제공이 가능하다[3]. 상황인지 컴퓨팅은 예외적인 상황에도 적절하게 대응이 가능하기 때문에 환자에게 맞춤형된 서비스 환경을 제공해 줄 수 있다[4]. 하지만 맞춤형된 추천을 위해서는 그 범위가 상당히 반대하며, 환자가 선호하는 서비스 또한 가변적으로 바뀌기 때문에 이를 해결하기에는 매우 어려우며, 현재까지도 맞춤형된 추천 서비스의 성공적인 사례가 전무하다[5]. 또한, 건강관리 서비스를 이용하는 환자들은 개인마다의 건강 증진을 위한 처방방법이 서로 다르며, 선호하는 서비스의 종류 또한 다르다.

상황인지 기반 유헬스 건강 서비스의 제공을 하기 위해서는 상황 데이터를 통해서 환자의 선호도 정보에 대한 판별이 필요하다. 따라서 상황 인지 환경의 서비스를 추천하기 위해서는 상황 데이터의 정의가 필요하며, 상황 데이터가 건강 서비스의 추천 요인에 대해 관련 있는지를 식별해야 한다.

본 논문에서는 상황인지 환경의 유헬스 건강서비스를 제공하기 위해 상황 데이터에 대한 선호도의 추천 요인들을 다변량 분석(Multivariate analysis)을 통하여 식별하며, 추천 요인에 따르는 규칙 기반의 추천 의사결정 모델을 생성한다. 규칙기반의 추천 의사결정 모델을 생성하기 위해 의사결정 트리 기법과 연관성 규칙 기법을 이용한다.

선호도 추천 요인을 분석하기 위해서 상황데이터를 정의하며, 제 5기 국민건강 영양조사[6]데이터를 통해서 상황 데이터가 대해 건강관리에 대한 두 가지의 추천 요

인(운동, 식사량 감소)에 대해 유의한 데이터를 판별을 한다. 또한, 판별된 추천 요인을 통해서 선호도 추천 의사결정 모델을 생성하여 환자의 상황 정보에 따라 선호 요인을 알 수 있으며 맞춤형 서비스를 제공할 수 있게 된다.

본 논문의 구성은 다음과 같다. 제2장에서는 관련연구에 대해 설명을 하며, 제3장에서는 추천 요인을 식별하기 위한 실험 및 결과, 제4장에서는 결론을 설명 한다.

## 2. 관련연구

### 2.1 상황인지 기반 유헬스 서비스

유헬스 서비스는 현대인들의 식생활의 변화, 활동량의 감소, 고령화 사회의 진입으로 인해 점차 중요성이 높아지고 있다. 질병에 따른 만성질환자가 급증하고 있는 가운데 고혈압, 당뇨병 등의 합병증이 발생할 가능성이 높으며, 환자는 자신의 관리가 소홀해 질 수 있기 때문에 큰 고통과 회복이 불가능한 상태도 초래할 수 있다[1]. 따라서 언제, 어디서나 관리할 수 있는 방법이 필요하다.

상황인식 서비스는 환자의 이벤트와 관련 있는 것을 상황으로 정의하여 환자가 언제, 어디서든지 여러 상황에 따라 컴퓨터가 이를 인식하여 환자에게 맞춤형 서비스를 제공하는 지능적인 서비스이다[7]. 상황인식 기반 서비스를 제공하기 위해서 IF-Then, 온톨로지, 필터링 방법 등 여러 가지 기법이 존재하며 이를 통해 적절한 서비스의 제공이 가능하다[8]. 하지만, 건강 서비스 제공에 대한 상황 데이터의 정의는 매우 어려우며, 어떠한 데이터가 건강 서비스와 관련이 있는지에 대한 식별이 필요하다. 이는 환자의 추천 요인에 대해 필요한 데이터이며, 이를 통해서 환자에게 맞춤형 서비스를 제공할 수 있게 된다. 식별된 데이터를 통해서 환자에게 맞춤형 서비스를 제공하기 위해 추천 요인을 제공할 수 있는 의사결정 모델을 생성해야한다.

### 2.2 추천 의사결정 모델

건강 서비스의 제공을 위해서는 상황데이터에 따르는 추천 방법이 필요하며, 어떤 건강서비스를 제공해야 할지 판별하는 의사결정 모델이 필요하다. 추천 의사결정 모델은 건강 서비스 제공에 앞서, 상황 데이터를 통해서 어떤 추천 요인을 선택할지에 대한 의사결정을 지원한다.

상황 데이터와 추천 요인에 대해 유의한 데이터를 판

별한 후, 판별된 데이터를 통해 추천 의사결정 모델을 생성한다. 의사결정 모델을 생성하기 위해서는 데이터 마이닝 기법을 통해 숨겨진 규칙이나 패턴을 발견할 수 있으며, 의사결정을 위한 룰을 생성할 수 있다[9].

룰을 생성하는 데이터 마이닝 알고리즘은 대표적으로 의사결정 나무(Decision Tree)[10]와 연관성 규칙(Association Rule)[11]이 있다.

의사결정 나무는 Entropy(평균 정보량)를 통해 노드를 생성하여 규칙을 제공할 수 있다. 연관성 규칙은 지지도(Support)와 신뢰도(Confidence)기반으로 연관성 규칙을 제공할 수 있다.

본 논문에서는 추천 요인 데이터를 이용하여 의사결정 나무와 연관성 규칙 알고리즘을 이용하여 추천 의사결정 모델을 생성한다.

### 3. 실험 및 결과

#### 3.1 상황 데이터 선택

건강 서비스를 위한 추천 요인 식별을 위해서 상황 데이터를 정의한다. 본 논문에서는 추천 요인 식별을 위해서 제5기 국민건강영양조사[1]를 이용한다. 제5기 국민건강 영양조사는 2010년부터 2012년까지 질병관리본부에서 수행한 조사이다. 국민건강 영양조사는 국민의 건강 수준, 건강관련 의식 및 행태, 식품 및 영양섭취 실태에 대한 국가 및 시도 단위의 대표성과 신뢰성을 갖춘 통계를 산출하는 것이며, 국민건강증진종합계획의 목표 설정 및 평가, 건강 프로그램 개발 등 보건정책 수립 및 평가에 필요한 기초자료이다.

건강 서비스에서 가장 대표적으로 운동 및 식이서비스가 있으며, 이 두 서비스는 윌헬스 서비스 모델에서 가

〈Table 1〉 Context Data

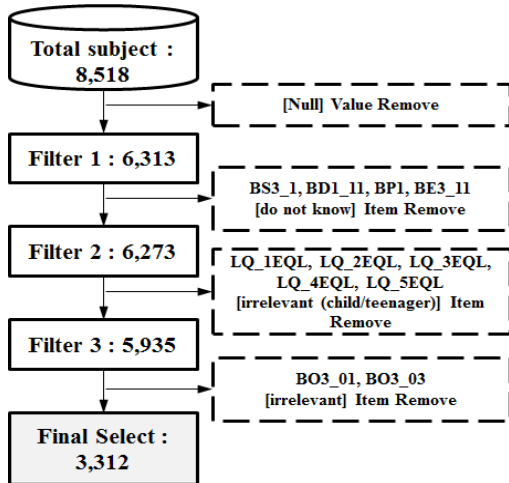
	Data	Range
Input	Sex(sex)	[1: male, 2: female]
	Income quartile-individual (incm)	[1: low, 2: medium low, 3: medium high, 4: high]
	교육수준 재분류 코드(edu)	[1: below elementary-school graduate, 2: middle-school graduate, 3: high-school graduate, 4: above university graduate]
	Education level reclassification code (occp)	[1: Supervisor · Expert, 2: Clerks, 3: Service · Sales, 4: Agriculture and forestry · fishing industry, 5: Equipment · machine · assembly, 6: Simple works, 7: Unemployed]
	Married (marri_1)	[1: Married, 2: Not Married, 8: do not know]
	Smoking status (BS3_1)	[1: smoking, 2: nonsmoking, 3: smoking in the past but currently nonsmoking, 4: irrelevant, 8: do not know]
	Frequency of drinking for 1 year (BD1_11)	[1: did not drink at all for the past 1 year, 2: less than once a month, 3: once a month, 4: 2-4 times a month, 5: 2-3 times a week, 6: over 4 times a week, 7: irrelevant, 8: do not know]
	Status of Hypertension(DI1_pr)	[0: yes, 1: no, 8: irrelevant (never suffered from high blood pressure)]
	Status of diabetes (DE1_pr)	[0: yes, 1: no, 8: irrelevant (never suffered from high blood pressure)]
	Days of intense physical activities a week (BE3_11)	[0: never, 1: 1 day, 2: 2 days, 3: 3 days, 4: 4 days, 5: 5 days, 6: 6 days, 7: 7 days (every day), 8: irrelevant, 9: do not know]
	Level of general stress awareness (BP1)	[1: extremely high, 2: considerably high, 3: considerably low, 4: almost none, 8: do not know]
	Motor ability (LQ_1EQL)	[1: no trouble walking, 2: some trouble walking, 3: must lie down all day, 8: irrelevant (child/teenager), 9: do not know]
	Self-control (LQ_2EQL)	[1: no trouble bathing or dressing, 2: some trouble bathing or dressing by myself, 3: cannot bathe or dress by myself, 8: irrelevant (child/teenager), 9: do not know]
	Daily activities (LQ_3EQL)	[1: no trouble doing daily activities, 2: some trouble doing daily activities, 3: cannot do daily activities, 8: irrelevant (child/teenager), 9: do not know]
	Output	Pain/discomfort (LQ_4EQL)
Anxiety/depression(LQ_5EQL)		[1: no anxiety/depression, 2: some anxiety/depression, 3: severe anxiety/depression, 8: irrelevant (child/teenager), 9: do not know]
Output	Excercise(BO3_01)	[0: yes, 1: no, 8: irrelevant]
	Decrease of meal(BO3_03)	[0: yes, 1: no, 8: irrelevant]

장 많이 이용되고 있는 서비스이다. 따라서 운동 및 식이 서비스에 대한 추천 요인의 식별이 필요하다.

본 논문에서는 위와 같은 데이터를 바탕으로 건강 서비스를 제공할 수 있는 요인들을 선택하여 상황데이터로 정의하며, 선택된 상황데이터에 대해 추천 요인에 유의한지에 대해 식별을 한다. 또한 식별된 데이터를 바탕으로 추천 요인 의사결정 모델을 생성한다.

추천 요인에 유의한지 판별을 위한 상황 데이터의 선택은 <표 1>과 같다. 총 18개의 변수 중에서 16개의 입력 변수와 2개의 출력변수로 구성한다. 입력변수로 구성된 항목은 상황 데이터로 정의하며, 이에 출력변수로 구성된 추천 요인항목과 유사한지 판별을 한다.

상황 데이터의 판별을 위해서 선택된 16개의 데이터에 대해서 전처리가 필요하다. 전처리는 입력이 누락된 항목과 불필요한 항목을 제거하는 과정이며 [그림 1]과 같이 상황 데이터에 대해서 전처리를 한다.



[Fig. 1] Data Preprocessing

총 8,518개의 레코드 중에서 Null값을 제거하고, 요인 분석에 부적합한 모름 항목 및 비해당 항목을 제거하여 5,935의 레코드가 선택 되었다. 또한, 추천 결과인 출력변수의 비 해당 항목을 제거하여 최종적으로 3,312개의 레코드를 이용한다. 3,312개의 데이터를 이용하여 추천 요인을 식별한다.

### 3.2 추천 요인 식별

상황 데이터가 추천 선호 요인에 적합한지 판단하기

위해 출력변수인 운동 및 식사량 감소 정보에 대해서 유의한지 판단한다. 3,312개의 데이터의 유의확률을 판단하여 추천 요인을 식별하기 위해서 SPSS의 Clementine 12.1을 이용하며, 다변량 분석(Multivariate analysis)을 사용하여 식별한다.

첫 번째로, 상황데이터에 대해 운동 요인에 대한 회귀 결과는 <표 2>와 같다. 유의 수준은 0.05로 정의하며, 운동 요인에 대한 상황데이터 중에서 유의한 상황 데이터는 총 8개로서, 1년간 음주 빈도(BD1\_11, 0.015), 1주일간 격렬한신체활동일수(BE3\_11, 0.000), 평소 스트레스인지 정도(BP1, 0.000), 현재흡연여부(BS3\_1, 0.027), 당뇨여부(DE1\_pr, 0.018), 자기관리(LQ\_2EQL, 0.040), 소득사분위수-개인(incm, 0.045), 성별(sex, 0.000)이다. 위와 같은 상황 데이터 8개는 운동 추천 요인에 관련이 있다고 할 수 있다.

<Table 2> Regression result of exercise factor

Variable	Understandized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error			
BD1_11	9.15E-03	0.004	0.042	2.434	0.015
BE3_11	4.74E-02	0.005	0.176	10.314	0.000
BP1	4.87E-02	0.011	0.077	4.346	0.000
BS3_1	7.96E-03	0.004	0.052	2.207	0.027
DE1_pr	-9.64E-03	0.004	-0.042	-2.363	0.018
DII_pr	-3.97E-03	0.003	-0.026	-1.395	0.163
LQ_1EQL	-1.29E-02	0.028	-0.01	-0.454	0.650
LQ_2EQL	-8.70E-02	0.042	-0.04	-2.052	0.040
LQ_3EQL	1.43E-02	0.029	0.011	0.487	0.626
LQ_4EQL	-2.67E-02	0.02	-0.027	-1.332	0.183
LQ_5EQL	1.85E-02	0.024	0.014	0.764	0.445
edu	9.96E-03	0.009	0.024	1.169	0.243
incm	1.42E-02	0.007	0.035	2.003	0.045
marri_1	1.73E-02	0.022	0.014	0.791	0.429
occp	6.07E-03	0.004	0.031	1.7	0.089
sex	-0.146	0.022	-0.158	-6.556	0.000

두 번째로, 상황데이터에 대해 식사량 감소에 대한 회귀 결과는 <표 3>과 같다. 유의 수준은 0.05로 정의하며, 식사량 감소 요인에 대한 상황 데이터 중에서 유의한 상황데이터는 6개로서, 1년간 음주 빈도(BD1\_11, 0.032), 1

주일간 격렬한신체활동일수(BE3\_11, 0.006), 운동능력(LQ\_1EQL, 0.008), 교육수준(edu, 0.000), 결혼여부(marri\_1, 0.000), 성별(sex, 0.000)과 같다. 위와 같은 상황 데이터 6개는 식사량 감소 추천 요인에 관련이 있다고 할 수 있다.

(Table 3) Regression result of decrease meal factor

Variable	Understandized Coefficients		Standar dized Coefficients	t	Sig.
	B	Std. Error	Beta		
BD1_11	-9.03E-03	0.004	-0.037	-2.146	<b>0.032</b>
BE3_11	-1.41E-02	0.005	-0.047	-2.734	<b>0.006</b>
BP1	-1.44E-02	0.013	-0.021	-1.148	0.251
BS3_1	2.56E-03	0.004	0.015	0.634	0.526
DE1_pr	1.63E-03	0.005	0.006	0.357	0.721
DI1_pr	-4.15E-03	0.003	-0.025	-1.302	0.193
LQ_1EQL	8.45E-02	0.032	0.06	2.648	<b>0.008</b>
LQ_2EQL	1.71E-02	0.047	0.007	0.36	0.719
LQ_3EQL	-1.67E-02	0.033	-0.012	-0.511	0.61
LQ_4EQL	3.98E-02	0.022	0.037	1.775	0.076
LQ_5EQL	8.15E-03	0.027	0.006	0.301	0.764
edu	4.11E-02	0.01	0.088	4.31	<b>0.000</b>
incm	4.28E-03	0.008	0.01	0.54	0.589
marri_1	0.12	0.024	0.087	4.897	<b>0.000</b>
occp	-1.02E-03	0.004	-0.005	-0.256	0.798
sex	0.137	0.025	0.135	5.492	<b>0.000</b>

### 3.3 추천 의사결정 모델 생성

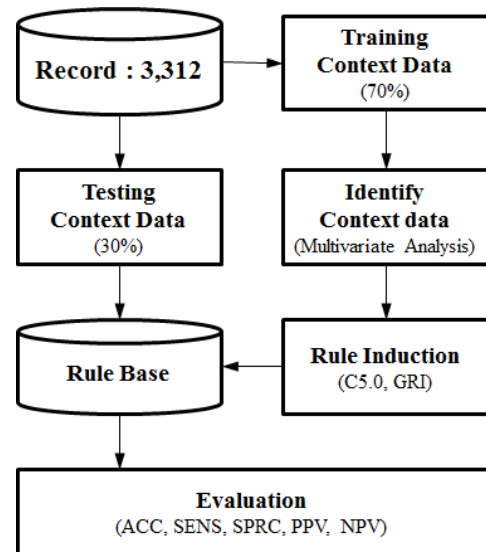
추천 요인과 관련된 상황데이터의 식별을 통해서 새로 입력받은 상황 데이터에 대해 어느 서비스를 제공할지 추천할 수 있다.

상황 데이터에 입력에 대해 추천 항목을 제시할 수 있으며 본 논문에서는 의사결정 트리의 C5.0 알고리즘과 연관성 분석의 GRI 알고리즘을 통한 룰 베이스 생성을 위해서 추천 요인과 관련된 데이터를 이용하며, SPSS의 Clementine 12.1을 이용하여 규칙을 생성한다.

추천 의사결정 모델의 규칙을 생성 및 평가하기 위한 구성은 [그림 2]와 같다.

총 3,312개의 데이터 중에서, 본 논문에서는 룰 베이스의 생성을 목적으로 하기 때문에 트레이닝 셋은 70%로 구성하였으며, 룰 베이스를 평가하기 위한 테스트 셋은 30%로 구성하였다. 트레이닝 셋은 추천 식별 상황 데이

터를 이용하여 룰을 생성을 한다.



[Fig. 2] Experiment Scenario

운동(BO3\_01)추천 요인에 대해 룰을 만들기 위한 상황데이터는 1년간 음주 빈도(BD1\_11, 0.015), 1주일간 격렬한신체활동일수(BE3\_11), 평소 스트레스인지 정도(BP1), 현재흡연여부(BS3\_1), 당뇨여부(DE1\_pr), 자기관리(LQ\_2EQL), 소득사분위수-개인(incm), 성별(sex)를 사용한다.

식사량 감소(BO3\_03)추천 요인에 대한 룰을 만들기 위한 상황데이터는 1년간 음주 빈도(BD1\_11), 1주일간 격렬한신체활동일수(BE3\_11), 운동능력(LQ\_1EQL), 교육수준(edu), 결혼여부(marri\_1), 성별(sex)을 이용한다.

70%의 트레이닝 데이터와 C5.0 알고리즘으로 생성된 이용한 운동, 식사량 감소의 연관성 룰의 구성은 <표 4>와 같다. C5.0 알고리즘의 옵션은 boosting 10, Cross-validation 10, Pruning Sensitive 50, Minimum records per child branch 2로 설정하였다.

70%의 트레이닝 데이터와 GRI 알고리즘으로 생성된 이용한 운동, 식사량 감소의 연관성 룰의 구성은 <표 5>와 같다. GRI 알고리즘의 옵션은 Minimum Confidence 30, Max Number of antecedents 3, Maximum number of rules 10으로 지정하였다.

<Table 4> Rule base using C5.0(Training set 70%)

Consequent	Rule
BO3_01 = 1	sex=2 and BE3_11 = 1 and DE1_pr=1 and LQ_2EQL = 1, 2
BO3_01 = 1	sex=2 and BE3_11 = 1 and DE1_pr=8 and BS3_1 = 2, 3
BO3_01 = 1	sex=2 and BE3_11 = 1 and DE1_pr=8 and BS3_1 = 8 and BP1 = 1 and incm = 3, 4
BO3_01 = 1	sex=2 and BE3_11 = 1 and DE1_pr=8 and BS3_1 = 8 and BP1 = 2 and incm = 2, 3
BO3_01 = 1	sex=2 and BE3_11 = 1 and DE1_pr=8 and BS3_1 = 8 and BP1 = 2 and incm = 4 and BDI_11 = 1,2,3,4,5,6,8
BO3_01 = 1	sex=2 and BE3_11 = 1 and DE1_pr=8 and BS3_1 = 8 and BP1 = 3,4
BO3_01 = 1	sex=2 and BE3_11 = 1 and DE1_pr=8 and BS3_1 = 9
BO3_01 = 1	sex=2 and BE3_11 = 2,3,4,5,6,7,8
BO3_01 = 1	sex = 1
BO3_03 = 1	sex = 2 and marri_1 = 2 and LQ_1EQL = 1
BO3_03 = 1	sex = 2 and marri_1 = 2 and
BO3_03 = 1	sex = 2

<Table 5> Rule base using GRI(Training set 70%)

Consequent	Rule
BO3_01 = 1	sex = 1 and LQ_2EQL = 1 and BS3_1 = 3
BO3_01 = 1	sex = 1 and BS3_1 = 3
BO3_01 = 1	LQ_2EQL = 1 and BS3_1 = 3
BO3_01 = 1	BS3_1 = 3
BO3_01 = 1	BS3_1 = 8 and BE3_11 = 1
BO3_01 = 1	DE1_pr = 8 and BS3_1 = 8 and BE3_11 = 1
BO3_01 = 1	sex = 2 and LQ_2EQL = 1 and BE3_11 = 1
BO3_01 = 1	sex = 2 and BS3_1 = 8 and BE3_11 = 1
BO3_01 = 1	sex = 2 and BE3_11 = 1
BO3_01 = 1	sex = 2 and DE1_pr = 8 and BE3_11 = 1
BO3_03 = 1	sex = 2 and marri_1 = 2 and LQ_1EQL = 1
BO3_03 = 1	sex = 2 and marri_1 = 2
BO3_03 = 1	sex = 2 and BE3_11 = 1
BO3_03 = 1	sex = 2

다음으로 각 생성된 룰에 대한 평가를 하기 위해서 정확도(Accuracy, ACC), 민감도(Sensitivity, SENS), 특이성(Specificity, SPEC), 양성 예측도(Positive predictive value, PPV), 음성 예측도(negative predictive value, NPV)를 측정한다[12]. 각 평가를 위한 식은 식1과 같으며, 오차 행렬은 <표 6>과 같다. 그리고 운동 및 식사량

감소에 대한 결과는 <표 7>과 같다.

$$ACC = \frac{TP + TN}{TP + FN + TN + FP} * 100 \quad (eq. 1)$$

$$SENS = \frac{TP}{TP + FN} * 100$$

$$SPEC = \frac{TN}{TN + FP} * 100$$

$$PPV = \frac{TP}{TP + FP} * 100$$

$$NPV = \frac{TN}{TN + FN} * 100$$

<Table 6> Confusion matrix

		Predicted Outcome	
		Positive	Negative
Actual Outcome	Positive	TP (True Positive)	FN (False Negative)
	Negative	FP (False Positive)	TN (True Negative)

<Table 7> Experimental result(Testing set 30%)

Exercise (BO3_01)		
	Multivariate analysis Based Decision Tree(C5.0)	Multivariate analysis Based Association Rule(GRI)
ACC	73.569	52.234
SENS	98.538	62.448
SPEC	11.913	26.736
PPV	73.420	68.030
NPV	76.744	22.190
Decrease of meal (BO3_03)		
	Multivariate analysis Based Decision Tree(C5.0)	Multivariate analysis Based Association Rule(GRI)
ACC	59.146	57.651
SENS	87.193	63.823
SPEC	20.531	49.180
PPV	60.169	63.283
NPV	53.797	49.763

실험을 위한 모델은 총 2가지로 다변량 분석 기반의 C5.0의사결정 트리 알고리즘(Multivariate analysis Based Decision Tree(C5.0))과 다변량 분석 기반의 GRI 연관성 규칙 알고리즘(Multivariate analysis Based Association Rule(GRI))이다. 운동 추천에 대한 정확도는 각각 73.569%, 52.234%로 의사결정 트리 알고리즘이 더 유의한 것으로 나타나있으며, 평균 62.902%로서, 두가지 룰을 사용하면, 운동 요인 추천에 대해 60%정도의 만족

도를 줄 수 있다.

식사 감소량 추천에 대한 정확도는 각각 59.146%, 57.651%이며, 의사결정 트리 알고리즘이 더욱 유의한 것으로 나타났다. 두가지 룰을 사용한다면 평균 58.389%로서, 약 60%정도의 식사량 감소 요인에 대해 만족도를 줄 수 있다. 또한, 민감도와 양성예측도는 는 연관성 규칙 알고리즘이 더욱 성능이 좋은 것으로 나타난다. 따라서 두 가지 모델을 통해 추출된 룰을 모두 이용하면 더욱 좋은 모델이 나올 수 있다.

제안하는 모델은 모두 약 60%정도의 정확도를 보이고 있으며, 민감도 약 78%이상, 양성예측도 약 65%이상으로 추천 요인에 따르는 의사결정 모델에 서비스 제공에 유의한 것으로 보인다. 실제 서비스 제공할 경우, 운동 및 식사 서비스 두 가지를 제공할 경우, 정확도가 100%가 넘기 때문에 사용자가 만족하는 서비스를 제공받을 수 있다. 따라서 상황인지 기반 건강서비스 에서 입력되는 상황 데이터에 따라서 환자에게 선호하는 추천 요인을 제공할 수 있다.

## 5. 결론

상황인지 기반 유헬스 서비스는 환자의 상황에 따라 건강 콘텐츠를 추천할 수 있으며, 선호하는 서비스를 제공하는 맞춤형 모델이 주목을 받고 있다. 본 논문에서는 건강 서비스를 제공하는 추천 요인에 대해 다변량 분석 기법을 통해 유의한 상황 데이터를 판별하였으며, 이를 통해서 추천 의사결정 모델을 생성하였다. 추천 의사결정 모델은 판별된 상황데이터를 바탕으로 의사결정 트리의 C5.0과 연관성 규칙의 GRI를 통해서 룰베이스를 생성하였다.

실험 결과, 상황 데이터에 대한 판별을 통해 추천 요인에 유의한 상황 데이터 10개를 추출 할 수 있었으며, 26개의 추천 요인에 대한 의사결정 모델을 생성하였으며, 정확도가 약 60%이상으로 나타나고 있다. 제안하는 방법은 상황인지 기반 유헬스 서비스를 제공받는 환자에게 추천 의사결정 모델을 통해서 질 높은 건강 서비스를 제공할 수 있을 것으로 기대된다.

## ACKNOWLEDGMENTS

This study was supported by a grant of the Korean Health Technology R&D Project, Ministry of Health & Welfare, Republic of Korea.(A112020).

## REFERENCES

- [1] Dong-Kyun Park, Jong-Hun Kim, Jae-Kwon Kim, Eun-Young Jung and Young-Ho Lee, U-health Service Model for Managing Health of Chronic Patients in Multi-platform Environment. The Journal of Korea Contents Association, Vol. 11, No. 8, pp. 23-32, 2011.
- [2] Jae-Kwon Kim, Jong-Hun Kim, Eun-Young Jung, Dong-Kyun Park and Young-Ho Lee, U-Health Platform based Health Management Service Model using Context Information. The Journal of Digital Policy and Management, Vol. 10, No. 8, pp. 185-192, 2012.
- [3] Byung-Mun Lee, Jae-Kwon Kim, Jong-Hun Kim, Young-Ho Lee, and Un-Gu Kang, A Customized Exercise Service Model based on the Context-Awareness in u-Health Service. Journal of Korean Institute of Information Technology, Vol. 9, No. 2, pp. 141-152, 2011.
- [4] Chi-Chun L., Chi-Hua C., Ding-Yuan C. and Hsu-Yang K., Ubiquitous Healthcare Service System with Context-awareness Capability: Design and Implementation. Expert Systems with Applications, Vol. 38, No. 4, pp. 4416-4436, 2010.
- [5] J. Ben Schafer, Dan F., Jon H. and Shilad S., Collaborative filtering recommender systems. Lecture Notes In Computer Science, Vol. 4321, pp. 291-324, 2007.
- [6] The Fifth Korea National Health and Nutrition Examination Survey (KNHANES V-2), 2011, Korea Centers for Disease Control and Prevention
- [7] A. K. Dey, G. D. Abowd, and D. Salber, A Conceptual Framework and a Toolkit for Supporting the Rapid Prototyping of Context-Aware Applications. Journal of Human Computer

- Interaction, Vol. 16, No. 2, pp. 97-166, 2001.
- [8] Jae-Kwon Kim, Young-Ho Lee, Jong-Hun Kim, Dong-Kyun Park and Un-Gu Kang, Semantics Environment for U-health Service driven Naive Bayesian Filtering for Personalized Service Recommendation Method in Digital TV. Journal of The Korea Society of Computer and Information, Vo. 12, No. 8, pp. 81-90, 2012.
- [9] Pang Ning Tan., Michel S. and Vinpin K. Introduction to Data Mining. Addison Wesley, 2007.
- [10] Anooj PK, Clinical decision support system: risk level prediction of heart disease using decision tree fuzzy rules. International Journal of IJRCS, Vol. 3, No. 3, pp. 1659 - 1667, 2012.
- [11] Smyth, P. & Goodman, R.M., Rule induction using information theory. AAAI Press, 1991.
- [12] Son CS, Kim YN, Kim HS, Park HS and Kim MS, Decision-making model for early diagnosis of congestive heart failure using rough set and decision tree approaches. Journal of Biomedical Informatics, Vol. 45, No. 5, pp. 999-1008, 2012.

#### 김재권(Kim, Jae Kwon)



- 2011년 2월 : 가천의과대 정보처리과(학사)
- 2011년 2월 ~ 현재 : 인하대학교 컴퓨터정보과 석사과정
- 관심분야 : 인공지능, 데이터마이닝, 분산처리, 모델링&시뮬레이션
- Email : [jaekwonkorea@naver.com](mailto:jaekwonkorea@naver.com)

#### 이영호(Lee, Young Ho)



- 1996년 2월 : 한국외국어대학교 응용전산학과(석사)
- 2005년 2월 : 아주대학교의료정보학과(박사)
- 1999년 2월 ~ 2002년 9월 : IBM Korea BI&CRM EM
- 2010년 2월 ~ 현재 : 가천대학교 전산정보센터장
- 2007년 2월 ~ 현재 : ISO/TC215 전문의원
- 2002년 10월 ~ 현재 : 가천대학교 IT대학 교수
- 관심분야 : 메디컬 인포메틱스, 유헬스케어
- Email : [lyh@gachon.ac.kr](mailto:lyh@gachon.ac.kr)