
신경망을 이용한 생활습관성 질환 시스템 설계

이영호*, 정경용**, 강운구***

Design of Life Habits Disease System using Neural Network

Young-Ho Lee*, Kyung-Yong Chung**, Un-Gu Kang***

요 약 현대사회는 IT융합기술의 발달로 정보의 양이 급속도로 늘어나고 있으며, 이로 인하여 많은 데이터 속에 원하는 정보를 용이하게 획득하거나 검색하는 기술도 발전되고 있다. 이에 따라 다양하고 많은 건강정보 제공 사이트가 개발되어 운영되고 있지만, 웹서비스에 기반한 정보 제공의 한계와 개인화의 부족으로 사용자의 건강관리와 증진에 효과적이지 못한 결과를 보이고 있다. 건강정보 지원 서비스는 생체정보를 획득하고, 획득된 데이터를 다시 컴퓨터에 입력하여 기존 네트워크 기반을 통하여 전송하는 형태로 개발되고 있기 때문에 불편함은 물론 비효율적이다. 본 논문에서는 기존의 의료 데이터와 Framingham 위험인자(FRS)를 활용, 신경망을 이용한 생활습관성 질환 시스템 설계를 제안한다. 제안하는 시스템을 통하여 생활습관성 질환 환자의 고통호소를 의사가 신속하게 파악할 수 있도록 기초자료와 가이드라인을 제공하게 되고, 따라서 환자의 안위 증진이 향상되게 된다.

주제어 : 헬스케어, 생활습관성 질환, 신경망, 진단

Abstract In modern society, the amount of information has been significantly increased according to the development of IT convergence technology. That leads to develop information obtaining and searching technologies from lots of data. However, they don't seem to have been actually effective for users' health care and promotion not only for limitations of Web-based information and but for lack of personalization. Health information support services have been generally developed in a format of inputting data on bio-information acquired into the computer for the existing network-based transmission. In this paper, a life habits disease system that uses the existing medical data and Framingham risk factor(FRS) performs neural network is proposed. Based on the proposed system, it is possible to provide the fundamental data and guideline to doctors for recognizing the life habits disease diagnosis of patients and that represents increases in the welfare of patients.

Key Words : Healthcare, Life Habits Disease, Neural Network, Diagnosis

1. 서 론

산업화에 따라 급속한 경제성장을 이루고 인구의 도시집중이 심화되고 식생활을 포함한 생활양식의 변화, 의료 기술의 발달 등으로 평균 수명이 연장되어 인구 고령화의 속도가 빨라짐에 따라 만성질환의 증가 속도 또

한 빨라지고 있다. 또한 이러한 사회 및 경제적 여건의 변화에 따라 질병의 양상도 크게 변화하여 감염성 질환이 주된 사망원인이었던 과거와는 달리 순환기계 질환, 암, 당뇨, 고혈압과 같은 비감염성 만성질환이 사망원인이 되고 있다. 이러한 만성질환에 의한 사망이 높게 나타난 현상의 원인은 현대인의 잘못된 생활습관에서 비롯되

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

*가천대학교 IT대학 교수(제1저자)

**상지대학교 컴퓨터정보공학부 교수

***가천대학교 IT대학 교수(교신저자)

논문접수: 2012년 5월 9일, 1차 수정을 거쳐, 심사완료: 2012년 5월 23일

었기 때문이다. 만성질환은 개인의 생활습관과 환경적 요인이 복합적으로 작용하여 장기간에 누적된 결과로 볼 수 있다[1][2][3]. 본 논문에서는 순환기계 질환, 압, 당뇨, 고혈압과 같은 비감염성 만성질환 발생을 사전에 예방하고 최소화할 위해 지속적으로 측정된 생체신호(콜레스테롤, 맥박, 흡연정도, HDL 수치 등)를 패턴화하고 향후 특정 시점에서의 건강 상태를 예측하는 방법론을 개발하고자 한다. 이는 개인별로 다른 생체신호 데이터를 이용하여 자신의 건강상태를 진단하는 것이 아니라 개인 맞춤형 건강관리 시스템을 이용하여 평소 자신의 생활패턴을 바탕으로 건강관리를 수행하게 된다. 제안하는 시스템은 단순한 컴퓨터 시스템이 아니라 주치의가 개인 환자의 모니터링 된 데이터를 지속적으로 관리해주고 그에 대한 결과를 3단계(건강, 심려, 위험)분류하여 평가하게 된다. 주치의의 평가된 데이터는 신경망을 이용한 생활습관성 질환 시스템의 결과에 대한 데이터로 규정되며 이를 통해 생활습관성 질환 시스템은 자율적으로 학습하며 환자 개인의 패턴을 생성하게 된다.

2. 관련연구

2.1 유 헬스 서비스

유 헬스(U-Health)는 IT융합기술을 활용하여 예방, 진단, 치료 및 사후 관리 등의 의료 서비스를 이용하는 것을 말한다. 종래의 원격보건의료, e-Health보다 넓은 개념의 서비스로 유·무선 정보통신 인프라와 디바이스를 이용하여 제공하는 모든 보건의료 서비스를 포괄한다. 대표적인 유 헬스 서비스에는 모바일 건강관리, 의료정보 온라인 제공, 질병 모니터링이 있다[4].

본 연구에서 사용하는 데이터의 형태가 시계열 자료임을 고려하여 분석하는 방법론으로 시계열 분석을 활용하였다. 시계열 분석은 과거를 이해하고 미래를 예측하기 위해 기관측 된 현상을 설명할 수 있는 법칙을 찾는 것으로써 시계열 생성 과정에 대한 모형을 과거 자료에 내재된 규칙성에 근거한 추정 가능하며 시스템의 미래 행위는 과거와 크게 다르지 않다는 가정 하에서 주어진 자료의 특정 파악, 모델링 및 예측의 세 단계를 거쳐서 이루어진다[5]. 시계열 자료의 대표적인 분석 방법은 지수 회귀법, 이동 평균법, 박스-젠킨스(Box-Jenkins), 신경망 기법이 있다. 기존 연구에서는 시계열 분석에 있어

신경망의 유용성을 검증하여 신경망의 예측성과를 객관적으로 평가하고자 다층 퍼셉트론 모형을 많이 사용하였으며 자기회귀-이동평균 모형과 비교평가 한 연구가 진행되어 왔다[6][7]. 다른 연구에서는 기내식 수요 예측에 효율적인 예측 기법을 도입하여 낭비 요소를 제거함으로써 원가 절감뿐만 아니라 향상된 서비스를 제공할 수 있도록 하기 위해 시계열 분석과 신경망을 활용하여 수요 예측을 위한 연구를 진행되었다[8].

2.2 임상 의사 결정 지원 시스템

보건·의료분야에서 환자 진료시에 의사는 진단과 치료에 관한 수많은 의사결정을 내려야 한다. 차세대 IT융합기술을 이용한 임상지침 기반의 임상 의사결정 지원 시스템(Clinical Decision Support System; CDSS)은 의료진과 자동화된 시스템 간의 다이얼로그를 생성하고 상대적 중요도를 통해 의료진으로 하여금 환자 상황에 적합한 임상 정보와 일반적 의학 지식에 더 잘 접근할 수 있도록 하며 진단에 대한 확률이나 치료 대한 정보를 컴퓨터가 직접·간접적으로 지원할 수 있는 시스템을 의미한다[8]. 지식 베이스를 지능 시스템에서 해석할 수 있는 형태로 표현하는 것과 이를 자동화할 수 있는 IT융합기술이 필수적이다[9]. 임상 의사결정 지원 시스템의 활용도를 높이기 위한 단말기와 무선 통신을 이용하여 임상 환경에서 현장 진료를 지원하기 위한 지능 시스템 개발에 대한 연구가 활발히 진행되고 있다[10][15].

임상 현장에서는 관상동맥질환의 예방과 평가를 위한 도구로서 Framingham 연구에 기초한 위험 점수도구(risk scoring tool)를 활용하고 있다[11]. 대규모 전향 코호트(prospective cohort) 연구를 통해서 심혈관질환(CHD)의 5년에서 10년 후의 발병까지의 환자 위험률을 평가한 것이다. 이는 단일한 위험인자를 평가하는 것보다 다중적인 위험인자를 평가하기 때문에 정확하고 가장 효과적인 심혈관 질환을 예측하고 치료를 결정하는데 지침이 된다고 할 수 있다[12]. 일반적으로, Framingham 공식은 30세 이하, 또는 65세 이상에서보다 중년의 백인과 흑인 남녀에서 위험척도를 정확하게 예측할 수 있다. Framingham 공식에 사용되는 위험인자는 나이, 저밀도 콜레스테롤(LDL-C) 또는 총콜레스테롤(TC), 고밀도 콜레스테롤(HDL-C), 혈압(SBP& DBP), 당뇨와 흡연이다. 이러한 위험 인자들에 대한 위험 정도에 따라 점수를 부여하여 그 총합을 이용하여 10년 후 심혈관 질환 발병 위

협성을 나타낸다. 여기서 혈압과 콜레스테롤 기준은 JNC-C(Joint National Committee)와 NCEP ATP II (National Cholesterol Education Program, Adult Treatment Panel)에 의한 것이다[13]. 40세 이상 중년 남성의 규칙적인 운동 습관이 관상 동맥 질환을 예측할 수 있는 Framingham 위험인자와 대사 증후군 결정요인에 미치는 영향을 알아보는 연구에서 대사증후군과 Framingham 위험인자의 관계와 유산소 능력과의 관계를 알아보기 위해 상관 분석을 실시한 결과 VO₂peak와 는 음의 상관을 나타내었고, 대사증후군 결정 요인과는 양의 상관을 나타내었다[14]. 기존에 연구에서는 Framingham 위험인자를 비교 연구의 도구로 사용하거나 심혈관 질환의 환자 위험률을 평가하는 도구로 사용 하였다. 또한, Framingham 위험인자는 일반화된 척도로 이용하기 때문에 환자 개인에 맞춘 위험률을 평가하는데 한계가 있다.

따라서 본 연구에서는 유헬스 환경에서의 의사의 진 단에 지원 가능한 임상 의사 결정 지원을 목표로 시계열 자료를 자가 학습을 통해 분석하고 예측하며 결과를 Framingham 위험인자에 적용함으로써 개인 패턴에 맞 춘 개인화된 생활습관성 질환 진단 서비스를 제공하는 시스템을 설계한다.

3. 신경망을 이용한 생활습관성 질환 시스템 설계

3.1 신경망을 이용한 질환 진단 모형

신경망의 학습을 통한 환자의 상태를 예측값을 Framingham 위험인자에 적용하여 환자의 심혈관 질환 의 위험률의 패턴을 평가하였다. Framingham 위험인자 의 입력 요소를 [표 1]에 제시한다.

[표 1]과 같은 수치 척도를 이용하여 나이, HDL, Systolic BP, 콜레스테롤, 흡연 유무의 6가지 요인 각각 에 점수를 찾아 그것을 합산하여 환자의 10년 이내에 심 혈관 질환이 발병할 절대적 위험률을 계산하였다. 진단 모델에서 각각 요인의 점수를 평가한다. 평가 점수를 예 측 모델에서 입력층의 입력 노드를 통하여 입력받는다.

〈표 1〉 Framingham 위험인자의 입력 요소

Systolic BP(mmHg)	IF Untreated	IF Treated
<120	0	0
120-129	0	1
130-139	1	2
140-159	1	2
≥160	2	3

Total Cholesterol	point				
	age 20-39	age 40-49	age 50-59	age 60-69	age 70-79
<160	0	0	0	0	0
160-199	4	3	2	1	0
200-239	7	5	3	1	0
240-279	9	6	4	2	1
≥280	11	8	5	3	1

	point				
	age 20-39	age 40-49	age 50-59	age 60-69	age 70-79
Nonsmoker	0	0	0	0	0
smoker	8	5	3	1	1

Point Total	10-Year Risk %
<0	<1
0	1
1	1
2	1
3	1
4	1
5	2
6	2
7	3
8	4
9	5
10	6
11	8
12	10
13	12
14	16
15	20
16	25
≥17	≥30

Age	Points
20-34	-9
35-39	-4
40-44	0
45-49	3
50-54	6
55-59	6
60-64	10
65-69	11
70-74	12
75	13

HDL (mg/dL)	Points
≥160	-1
50-59	0
40-49	1
<40	2

다층퍼셉트론 신경망은 입력층과, 은닉층을 가지고 있 으며, 뉴런들 사이의 연결 링크는 각각의 가중치를 갖는 다. 다층 퍼셉트론은 가중내의 뉴런들 사이의 연결과 출 력층에서 입력층으로의 직접적인 연결은 존재하지 않는 전방향 신경망이다. 대부분의 다층 퍼셉트론의 학습은 역전파 알고리즘을 사용하여 수행할 수 있다. 이는 주어진 입력에 대해 원하는 출력 결과를 학습시키고자 할 때 사용하며, 출력층의 각 뉴런에서 발생하는 출력 오차를 각 층의 역으로 전파시켜 나가면서 연결 링크의 가중치 수정을 통해 오차를 최소화시킨다. 전방향 신경망은 뉴

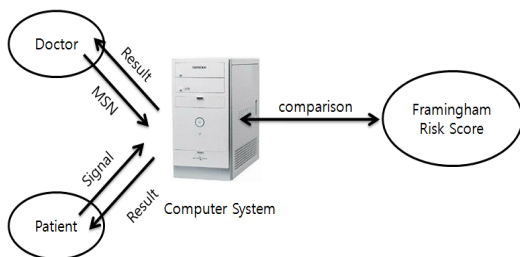
런이 정렬된 층에 배치되어 구성된다. 3층으로 구성되는 신경망을 $I*H*O$ 으로 표시한다. I는 입력층에 속한 뉴런의 수, H는 중간층에 속한 뉴런의 수, O는 출력층에 속한 뉴런의 수이다. 입력층에 속한 뉴런은 입력을 중간층에 제공한다. 중간층과 출력층에 속한 뉴런은 하위층에 속한 뉴런의 출력 값과 연결 강도의 합을 계산 후 전달 함수를 통하여 출력 값을 계산한다.

3.2 생활습관성 질환 진단 시스템

3.2.1 대상자 설정 및 주요데이터 설정

본 연구는 40세 이상 50세 이하 남자를 대상으로 하였다. 대상자 설정 및 주요데이터 설정할 때 특이사항에 대한 질병의 유무를 기준으로 정상과 비정상 군집을 분류하였다. 환자의 성별, 나이, 생체신호(수축기 혈압, 이완기혈압, 총콜레스테롤, BMI 등), 흡연유무를 주요데이터로 설정하였다. 생활습관성 질환 진단 시스템의 시나리오 오는 [그림 1]과 같이 3단계로 구성되어 있으며 단계별 내용은 다음과 같다.

개인의 연령, HDL, 맥박, 흡연유무, 콜레스테롤 수치와 같은 자신의 생체신호를 매일 자가 측정한다. 측정된 기록 데이터는 컴퓨터를 통해 하루 단위로 모니터링 되고 데이터베이스에 저장된다. 측정 데이터의 기록 결과는 의사의 모바일에 MSN으로 전송되고 의사는 MSN로 전송된 환자의 예측 결과를 통해 3단계(건강, 심려, 위험)로 분류 평가한다. 평가 결과는 생활습관성 질환 시스템의 자가 학습을 위한 데이터로 활용되며 생체신호는 학습을 통해 패턴화된다. 이를 바탕으로 특정 시점의 예측 결과를 확인한다. 특정 시점을 생체신호를 Framingham 위험인자의 데이터로 입력하여 위험률을 확인하고 시점의 위험률을 확인하여 환자의 상태를 평가한다.



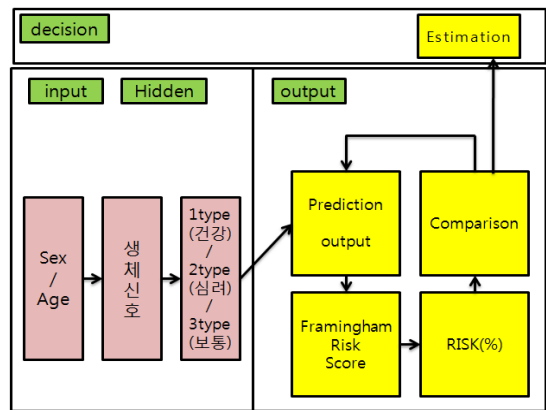
[그림 1] 생활습관성 질환 진단 시스템

Framingham 위험인자는 현재의 건강 상태를 통해서 향후 10년 후 위험률에 대한 수치를 확인해 주는 것이다.

따라서 현재 측정된 데이터 및 미래에 대한 생체신호 데이터를 예측하여 환자가 살아가는 건강 상태가 10년 후에 어떤 영향을 미칠 것인지를 평가해주게 되며 백분율로 표현되는 것을 활용하는 개인화된 예측반영 진단 규칙이 되는 것이다.

3.2.2 제안하는 시스템의 동작 과정

제안하는 신경망을 이용한 생활습관성 질환 진단 시스템의 동작 과정은 다음과 같다. [그림 2]는 제안하는 생활습관성 질환 진단 시스템의 동작 과정을 나타낸다.



[그림 2] 생활습관성 질환 진단 시스템의 동작 과정

[단계 1] 개인의 성별, 연령, 흡연유무를 입력하고 생체신호(수축기 혈압, 이완기혈압, 총콜레스테롤, BMI 등)를 측정한다. 측정된 데이터를 입력층의 4개의 입력노드를 통하여 입력받는다.

[단계 2] 입력된 데이터를 건강, 심려, 위험)로 3가지로 분류한다.

[단계 3] 1개의 은닉층의 8개의 노드들과 입력층의 노드들은 병렬적으로 연결되어 있는 형태를 지닌다. 점수화된 각각의 생체신호에 대한 가중치 적용하여 환자의 상태를 패턴화한다.

[단계 4] 학습을 통해 패턴화 된 예측값의 특정 X_t 시점을 Framingham 위험인자 위험률을 평가한다.

[단계 5] 학습을 통해 패턴화 된 예측의 특정 X_{t+1} 시점을 Framingham 위험인자 위험률을 평가한다.

- [단계 6] X_t 와 X_{t+1} 의 시점간 위험률을 비교한다.
- [단계 7] 위험률을 평가를 위해 [단계 4]부터 반복한다.
- [단계 8] 평가된 위험률을 바탕으로 환자의 패턴을 정의하고 건강 상태를 결정한다.

4. 결론

본 논문에서는 생활습관성 질환으로 인한 돌연사를 예방하고 개인 맞춤형 건강관리를 위한 생활습관성 질환 진단 시스템을 신경망 모델을 이용하여 설계하였다. 제안하는 방법은 진단과 예측을 통한 지능 시스템으로 단순한 컴퓨터에만 의존한 것이 아닌 주치의에 의한 3단계 (건강, 심려, 위험) 평가를 바탕으로 자율적으로 학습하여 환자 개인의 패턴을 생성함으로써 환자 건강 상태를 예측하여 개인에 맞춘 건강관리를 제공받게 된다. 제안하는 생활습관성 질환 진단 시스템을 통한 환자의 생체 신호 예측값을 Framingham 위험인자에 적용하여 10년 후의 질병 위험률을 확인하고 환자의 건강 패턴을 생성하여 환자의 건강 상태를 평가하여 관리한다. 차세대 IT 융합기술을 이용한 지능 시스템으로써 사람의 경험을 신경망을 통해 학습해서 가중치를 가지고 진단하며 예측 결과를 가지고 미래시점을 예측하는 임상 의사 결정 지원 시스템이 된다.

본 연구는 다음과 같은 한계점을 가지고 있다. 환자 패턴에 맞춘 진단으로 신경망으로 측정된 수학적 값만으로는 환자의 상태를 정확히 알 수 없다. 또한 의료 전문가가 쉽게 이해할 수 있는 근거를 뒷받침할 수 있는 시스템의 연구가 진행되어야 한다. 신경망의 예측 결과는 미래 시점의 예측 결과이기 때문에 미래에 적용 가능하나, 현재 시점에는 적용 불가능하다.

참 고 문 헌

- [1] 이상영, 만성질환 관리를 위한 지속적 건강관리체계 구축, 보건복지포럼 통권, 제87호, pp.72-81, 2004.
- [2] 허미경, 만성질환예방을 위한 보건교육제도의 개선방안, 한국보건사회연구원, 보건복지포럼 통권, 제21호, pp.25-36, 1998.
- [3] 김진태, 김명성, 이진우, 네트워크 기반의 u-Healthcare

- 서비스 추진동향, ETRI, 주간기술동향통권 제1321호, 2007.
- [4] 김성욱, 이성호, 유헤스(u-Healthcare)시대의 도래, 삼성경제연구소, 제602호, 2007.
- [5] 지원철, "신경망을 이용한 시계열의 분해분석", 홍익대학교 정보산업공학과, 대한산업공학회지, 제 25권, 1호, pp.111-124, 1999.
- [6] 윤여창, 허문열, "신경망이론에 의한 시계열자료의 분석", 한국통계학회논문집, 제4권, 제1호, pp.91-99, 1997.
- [7] 서창적, 광수환, "시계열분석법과 인공신경망을 이용한 기내식 수요예측", 한국생산관리학회, 제9권, 제2호, pp.67-95, 1998.
- [8] 박종선 외 9명, *의료정보학 개론*, 보문각, 2008.
- [9] 조인숙, 김지현, 이은정, 김소영, 이재호, 김윤, "고혈압관리 의사결정지원시스템 요구사항 규명을 위한 접근", 대한의료정보학회지, 제13권, 제3호, 2007.
- [10] 장희정, 김교창, 이강선, 진종훈, 최진욱, "무선 인터넷 서비스를 이용한 의사결정 지원 시스템의 설계 및 구현", 한국인터넷정보학회 2002 추계학술발표대회 논문집, 제3권, 제2호, 2002.
- [11] Sheridan S, Poignon M, and Mulrow C, "Framingham-based Tools to Calculate the Global Risk of Coronary Heart Disease: a Systematic Review of Tools for Clinicians, J. of Gen. Intern. Med., Vol. 18, No. 12, pp.1039-1052, 2003.
- [12] Murray CJ, Lauer JA, Hutubessy RC, Niessen L, Tomijima N, Rodgers A, Lawes CM, Evans DB, "Effectiveness and Costs of Interventions to Lower Systolic Blood Pressure and Cholesterol: a Global and Regional Analysis on Reduction of Cardiovascular-Disease Risk", Evidence and Information for Policy, World Health Organization, 27, Geneva, Switzerland.
- [13] Wilson PWF, D'Aostino RB, Levy D, Belanger AM, Silbershatz H, Kannel WB, "Prediction of Coronary Heart Disease Using Risk Factor Categories", Circulation, Vol. 97, pp.1837-1847, 1998.
- [14] 주기찬, 이한준, "40세 이상 중년 남성의 규칙적인 운동습관과 Ramingham Risk Score 및 대사증후군 결정요인간의 관계", 한국운동생리학회, 운동과학, 제 15권, 제3호, 2006.

[15] 이강대, 이영호, 강운구, 정경용, “병원내 신속 대응을 위한 RFID 기반의 의약품 물류 시스템”, 한국콘텐츠학회논문지, 제8권, 제10호, pp.1-10, 2008.

이 영 호



- 1996년 한국외국어대학교 응용전산학과(석사)
- 2005년 아주대학교 의료정보학과(박사)
- 1999년 ~ 2002년 IBM Korea BI&CRM EM
- 2007년 ~ 현재 ISO/TC215전문위원
- 2002년 ~ 현재 가천대학교 IT학과 교수
- 관심분야 : 메디컬인포매틱스, 유헬스케어
- E-mail : lyh@gachon.ac.kr

정 경 용



- 2000년 인하대학교 전자계산공학과(학사)
- 2002년 인하대학교 컴퓨터정보공학과(석사)
- 2005년 인하대학교 컴퓨터정보공학과(박사)
- 2006년 ~ 현재 상지대학교 컴퓨터정보공학부 교수
- 관심분야 : 지능시스템, 데이터마이닝, HCI
- E-mail : dragonhci@hanmail.net

강 운 구



- 2001년 인하대학교 전자계산공학과(박사)
- 2006년 뉴미디어연구소장
- 2007년~현재 가천대학교 u-헬스케어연구소장
- 1994년~현재 가천대학교 IT대학 교수
- 현재 IT융합산업정책위원
- 관심분야 : 의료정보, 유헬스케어, SW공학
- E-mail : ugkang@gachon.ac.kr