

빅 데이터 기반 만성질환 관리 시스템

강희범 · 이종원 · 김경환 · 김창수 · 정희경*

배재대학교

A web-based Obesity Management system using Body variations

Hee-Beom Kang · Jong-Won Lee · Kyung-Hwan Kim · Chang-Su Kim · Hoe-Kyung Jung*

Paichai University

E-mail : {rkdgml90, starjwon, shwan10}, {ddoja, hkjung}@pcu.ac.kr

요 약

오늘날 만성질환자에게 데이터를 제공해주고 관리하는 시스템의 필요성이 대두되고 있다. 그러나 대부분의 질병관리 시스템의 경우 사용자에게 광범위한 데이터를 제공하거나, 중요한 키워드 및 데이터를 제공해주지 않는 문제점이 존재하였다.

본 논문에서는 질병에 대한 데이터에 대해 R프로그래밍을 통해 분석하여 해당 질병에 가장 관련이 높은 키워드를 사용자에게 추천해준다. 이를 통해 사용자가 자신의 질병을 관리할 시 중요한 키워드들을 효율적으로 관리할 수 있는 빅 데이터 기반 만성질환 관리 시스템을 연구하였다. 제안하는 시스템을 활용하여 사용자는 불필요한 데이터나 키워드를 제외하고 필요로 하는 데이터와 키워드를 볼 수 있을 것으로 사료된다.

ABSTRACT

Today, need for a development system that provides the data to a chronic disease, and management has emerged. However, for most of the disease management system provides a wide range of data to the user and problem does not provide for important keyword or data existed.

In this paper, analyzing the data for a disease through the R Programing it makes like the most relevant keyword in the illness to the user. This study was a system in which only the important parts when the user to manage their disease can be efficiently managed. By utilizing the proposed system to the user it is considered to be Except for unnecessary data or keyword and to be able to see the data and the keyword in need.

키워드

Big data, Healthcare, Recommendation, Scraping

I. 서 론

최근 정보통신기술의 발달 및 서비스의 보급으로 환경이 급격히 변함에 따라 삶의 질이 향상되고 있다. 이로 인해 노령화된 인구의 증가로 만성질환이 차지하는 비율의 가속화가 증가하고 있고, 이를 예방하고 관리하기 위한 시스템 개발이 활발히 이루어지고 있다[1,2]. 하지만 현재 질병관리 시스템의 경우에는 사용자의 편의에 맞추어 관리하는 것이 가장 효율적이지만 대부분의 시스템은 모든 질병을 한 번에 같이 관리하는 것이 대부분

이었다[3,4].

이를 해결하기 위해 본 논문에서는 Node.js를 통해 웹 스크래핑을 하여 사용자의 질병에 대한 정보를 수집하고, R을 통한 분석을 통해 키워드만 추출한 뒤, 가장 많이 사용되어진 키워드를 저장하였다. 그리고 키워드들을 추출하여 질병을 가진 사용자들의 성향을 체크할 수 있도록 자동으로 체크박스를 생성되게 하였고, 체크된 값을 Count하여 그래프로 시각화하였다. 그리고 혈당, 혈압 등 단순한 데이터만 사용자에게 보여주지 않고 체중이나 BMI 지수 등을 함께 보여줌으로써 사용자가 여러 부분을 같이 관리할 수 있도록 한

빅 데이터 기반 질병 키워드 추천 시스템을 연구 하였다.

II. 시스템 설계

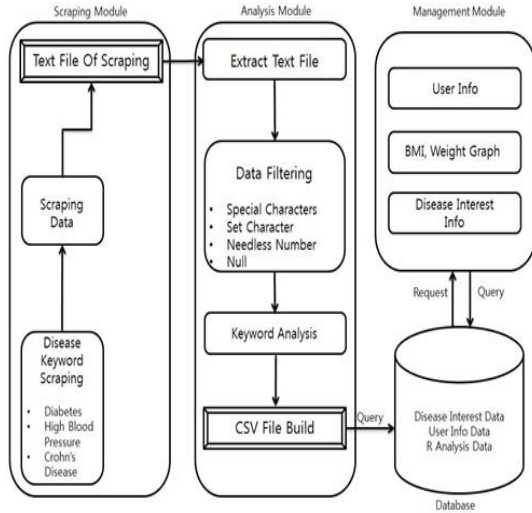


그림 1. 시스템 구조도

그림 1은 빅 데이터 기반 질병 키워드 추천 시스템의 전체 구조도이다. 모듈별로 Scraping Module과 Analysis Module, Management Module로 나누어진다. Scraping Module에서는 Node.js를 사용하여 질병에 대한 데이터를 수집하여 Text 파일로 저장하게 된다. 질병은 만성질환으로써 현재 대두가 되고 있는 당뇨병(Diabetes), 고혈압(High Blood Pressure), 크론병(Crohn's Disease)을 선택하였다. Analysis Module에서는 R Studio를 사용하여 Scraping Module에서 저장한 Text 파일에서 키워드를 추출하는 역할을 한다. 우선 Text 파일을 추출해 R Studio의 변수에 저장한다. 그리고 저장된 데이터에서 의미 있는 키워드를 추출하기 위해서 필터링을 하게 되는데 단계로는 특수문자나 순번을 나타내는 숫자를 삭제하고, 설정해둔 조사나 맺음말 등 필요 없는 단어를 삭제하게 된다. 마지막으로 공백이 된 부분과 기존에 존재했던 공백을 삭제하고, 분석을 시작한다. 분석에는 필터링 된 키워드를 가지고 Count를 하게 되고, Count에 따른 정렬을 한다. 그리고 분석된 키워드를 CSV 파일 형식으로 저장한다. Management Module은 Web에서 질병 관리에 대한 데이터를 제공하고, 분석된 데이터를 시각화하여 보여주는 것이다. 관리 데이터는 사용자의 이름, 나이, 질병, 같은 질병을 가지고 있는 사용자의 명수를 나타내 보여준다.

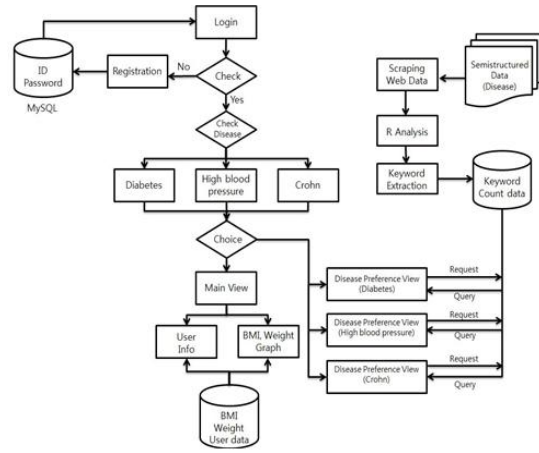


그림 2. 시스템 흐름도

그림 2는 전체 시스템의 흐름도이다. 처음 사용자가 로그인을 하는 것으로부터 시작한다. 로그인을 해서 저장된 회원가입 DB와 일치하면 다음으로 넘어가게 되고 아니면 회원가입으로 이동하여 가입 후 다시 로그인해야한다. 로그인이 되면 사용자의 질병을 체크하게 되는데 당뇨병, 고혈압, 크론병 중 어디에 속해있는지 확인한다. 확인 후 두 개의 버튼이 있는데 첫 번째는 Main View에는 사용자의 이름과 나이, 질병, 같은 질병을 가지고 있는 사용자들의 수를 볼 수 있고, 사용자가 신장과 체중을 입력할 수 있는 텍스트박스과 그에 대한 최근 값을 DB에서 불러와 그래프와 표로 나타낸 것을 볼 수 있다. 질병 관리를 하는데 체중과 BMI 수치를 함께 나타낸 것은 질병은 비만과 관련이 높아서 같이 관리하는 것이 효율적이기 때문이다. 두 번째는 질병에 대한 키워드의 성향을 시각화한 것을 볼 수 있다. 로그인 때 체크한 질병에 따라 당뇨병은 당뇨병 키워드 성향 분석, 고혈압은 고혈압 키워드 성향 분석, 크론병은 크론병 키워드 성향 분석 페이지로 자동으로 변경된다. 시각화하기 위해 제공되는 키워드는 비정형 데이터를 Node.js를 이용한 웹 스크래핑을 이용하여 수집하고 R을 통해 키워드를 분석하고 추출하여 저장된 데이터를 사용한다.

III. 고찰 및 결론

기존의 질병과 관련된 시스템은 질병관련 데이터를 나열하여 보여주는 서비스를 제공하고 있어 필요한 정보를 찾기 힘들었다. 또한 같은 질병을 가진 타 사용자와의 데이터 공유가 되지 않았고, 병원을 가야만 자신과 비슷한 질병을 가진 타 사용자의 질병에 대한 관심분야 데이터를 얻는 것이 가능했다.

이를 해결하기 위해 본 논문에서는 질병 키워드를 추천하는 시스템을 연구하였다. 본 시스템은

웹의 당뇨병, 고혈압, 크론병의 데이터를 수집하고 빈도수와 빈도수 비율을 사용하여 분류 및 추출하여 데이터베이스 테이블에 적재하였다. 적재된 키워드를 자동적으로 체크박스로 생성되게 하고 이를 통해 사용자가 중요 키워드 중 관심 있는 키워드를 선택하여 습득할 수 있다. 또한 사용자는 키워드 데이터를 보고 자신이 관심 있는 키워드와 타 사용자들의 관심 키워드 그래프를 비교하여 추천받을 수 있다. 이로 인해 질병에 대한 정보를 사용자가 습득할 시 복잡한 과정을 거치지 않고 웹에 나타난 질병정보만을 보고 질병에 대한 중요한 관심분야를 얻을 수 있을 것으로 사료된다.

ACKNOWLEDGMENTS

This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education(No. 2014R1A1A2059842)

참고문헌

- [1] H. I. Jung, J. G. Yang, J. I. Woo, B. M. Lee, J. S. O, K. Y. Chung, Y. H. Lee, "Evolutionary rule decision using similarity based associative chronic disease patients," *Cluster Computing*, vol. 18, no. 1, pp. 279-291, 2015.5
- [2] Eugenio Santoro, Gianluca Castelnuovo, Italo Zoppis, Geancarlo Mauri, Francesco Sicurello, "Social media and mobile applications in chronic disease prevention and management," *Frontiers in Psychology*, vol. 6, 2015.5
- [3] G. T. Ryu, H Choi, "Implementation of U-Healthcare System for Chronic Disease Management," *Electronics and Information Engineers*, vol. 51, no. 1, 2014.6
- [4] G. J. Kim, J. S. Han, "Chronic Disease Management using Smart Mobile Device," *The Society of Digital Policy & Management*, vol. 14, no. 4, pp. 335-342, 2014.3