헬스 빅데이터 플랫폼에서 이기종 라이프로그 마이닝 모델

강지수¹, 정경용^{2*}
¹경기대학교 컴퓨터과학과 학생, ²경기대학교 컴퓨터공학부 교수

Heterogeneous Lifelog Mining Model in Health Big-data Platform

JI-Soo Kang¹, Kyungyong Chung^{2*}

¹Data Mining Lab., Department of Computer Science, Kyonggi University

²Division of Computer Science and Engineering, Kyonggi University

요 약 본 논문에서는 헬스 빅데이터 플랫폼에서 이기종 라이프로그 마이닝 모델을 제안한다. 이는 사용자의 라이프 로그를 실시간으로 수집하고 헬스케어 서비스를 제공하기 위한 온톨로지 기반의 마이닝 모델이다. 제안하는 방법은 이기종 라이프로그 데이터를 분산처리하고, 클라우드 컴퓨팅 환경에서 실시간으로 처리한다. 이를 이기종 온톨로지를 기반으로 구성한 환경에 적합하도록 상위 온톨로지 방식으로 지식베이스를 재구성한다. 재구성한 지식베이스는 Jena 4.0 추론엔진을 이용해 추론 규칙들을 생성하고, 규칙 기반 추론 방법으로 실시간 헬스 서비스를 제공한다. 라이프로그 마이닝을 숨겨진 관계에 대한 분석과 시계열적 생체신호에 대한 예측모델을 구성한다. 이는 관계나 추론규칙에서 포함되지 않은 음의 상관관계나양의 상관관계를 탐색하여 사용자의 생체신호에 대한 변화를 감지하고 예방 의료 서비스를 현실화하는 실시간 헬스케어서비스가 가능하다. 성능 평가는 제안한 이기종 라이프로그 마이닝 모델 방법이 정확도에서 0.734, 재현율에서 0.752로 다른모델에 비해 우수하게 나타난다.

주제어 : 온톨로지, 라이프로그, 데이터 마이닝, 헬스케어, 잠재적 건강위험, 클라우드 컴퓨팅

Abstract In this paper, we propose heterogeneous lifelog mining model in health big-data platform. It is an ontology-based mining model for collecting user's lifelog in real-time and providing healthcare services. The proposed method distributes heterogeneous lifelog data and processes it in real time in a cloud computing environment. The knowledge base is reconstructed by an upper ontology method suitable for the environment constructed based on the heterogeneous ontology. The restructured knowledge base generates inference rules using Jena 4.0 inference engines, and provides real-time healthcare services by rule-based inference methods. Lifelog mining constructs an analysis of hidden relationships and a predictive model for time-series bio-signal. This enables real-time healthcare services that realize preventive health services to detect changes in the users' bio-signal by exploring negative or positive correlations that are not included in the relationships or inference rules. The performance evaluation shows that the proposed heterogeneous lifelog mining model method is superior to other models with an accuracy of 0.734, a precision of 0.752.

Key Words: Ontology, Lifelog, Data Mining, Healthcare, Potential Health-risk, Cloud Computing

^{*}This work was supported by the Korea Foundation for the Advancement of Science & Creativity (KOFAC), and funded by the Korean Government(MOE).

^{*}Corresponding Author: Kyungyong Chung (dragonhci@gmail.com)

1. 서론

의료산업에서 융합기술에 대한 연구가 계속됨에 따라 사람의 기대수명이 증가하고 있다[1]. 하지만 환경오염과 같은 다양한 외부요인으로 인해 암 발병률 및 사망률도 같이 증가한다. 따라서 기대수명 대비 만성질환 환자의 수가 늘어난다. 이에 우리는 장수와 질병의 공존이 필수 불가결 한 유병장수시대에서 생활하게 된다. 유병장수시 대에서는 지속적인 건강관리와 증진이 헬스케어의 중요 한 요소이다. 실제로 다수의 만성질환자는 나이가 많으 며 암환자는 지속적인 항암치료나 대형 수술로 인해 생 존율 및 면역력이 저하된다[2]. 따라서 완치판정 이후에 도 지속적인 건강관리가 필요하다. 때문에 만성질환자의 질병에 대한 지속적인 모니터링 서비스를 목표로 하며 잠재적 건강이상의 변화를 감지하고 질병을 예방 할 수 있는 건강예방 인공지능 기술이 필요하다[1,2]. 예방 의료 는 헬스 빅데이터 플랫폼을 이용하여 잠재적 건강위험을 분석하고 그 결과로 질병의 초기 발병 발견을 중요시하 며 예방에 초점을 둔다.

최근 헬스케어산업이 각광받으며 다양한 라이프로그를 수집할 수 있는 사물인터넷 및 웨어러블 디바이스와 개인 건강 기기가 연구, 개발되고 있다[3]. 이는 의료관리부문의 취약점을 보완하고 지속가능한 헬스케어를 제공한다. 건강 디바이스가 보편화됨에 따라 이기종 라이프로그를 통해 건강이상 징후를 탐지하기 용이해졌다[1-4]. 하지만 이를 실시간으로 분석하고 체계적으로 관리할 수 있는 서비스 모델에 대한 연구는 미약하다. 따라서 본 논문에서는 헬스 빅데이터 플랫폼에서 이기종 라이프로그마이닝 모델을 제안한다.

2. 관련연구

2.1 빅데이터 전처리

빅데이터는 환경에 따라 다양한 형태로 수집이 되어 활용 가능한 형태로 변환하는 전처리 과정이 필요하다. 이는 수집한 데이터를 활용하기 유용한 형태로 정제하는 과정이다. 전처리 과정을 거치지 않은 원시 데이터는 필 요한 속성이 없는 불완전 문제, 데이터 내에 의미가 맞지 않는 값이 있거나 범위 밖의 이상치가 있는 노이즈 문제, 그리고 여러 데이터를 통합할 때 발생할 수 있는 비일관 성 문제가 나타난다. 그림 1은 빅데이터의 전처리 프로세 스를 나타낸다. 빅데이터 전처리 프로세스는 데이터 클리닝, 데이터 통합, 데이터 변환, 데이터 축소, 데이터 이 산화의 총 다섯 가지 단계로 구분이 가능하대4.51.

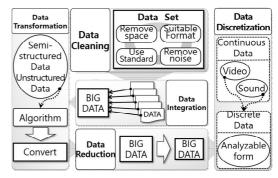


Fig. 1. Big Data Preprocessing Process

데이터 클리닝은 데이터의 공백을 채우고 잡음은 제 거하며 모순된 데이터를 해결하는 과정이다. 데이터 집 합을 쉽게 사용할 수 있도록 만들며 불일치나 오류를 수 정하고 추론엔진이 읽을 수 없는 요소를 제거한다. 데이 터 통합은 데이터가 다양한 형태로 곳곳에 산재해 있어 부분적으로만 효용가치가 있고 전사적으로 즉각적인 활 용이 가능한 데이터가 되지 못하고 있는 것들을 가치 있 고 재사용 가능한 정보 데이터로 가공하는 과정이다. 데 이터 변환은 반정형, 비정형 데이터를 추론엔진이 이해 할 수 있도록 변환하는 과정이다. 마이닝 알고리즘의 효 율성을 극대화시키기 위해 데이터를 알고리즘에 따라 변 환한다. 데이터 축소는 빅데이터의 원형을 유지하면서 크기는 더 작지만, 분석 결과는 동일한 데이터 집합의 표 현을 만드는 과정이다. 편집, 정렬, 통합 등을 이용하여 데이터 중복을 없애는 등 가공되지 않은 원시 데이터를 보다 요약된 형태로 변경한다. 마지막으로 데이터 이산 화는 시간과 공간에서 발생하는 연속적인 값 또는 속성 을 분석이 가능한 이산 형태로 표현하는 방법이다[4,5].

2.2 빅데이터 처리 및 분석

빅데이터 처리기술은 방대한 양의 데이터와 데이터 생성 속도, 데이터 종류의 다양성을 통합적으로 고려해야 한다[6]. 예로 빅데이터 일괄 처리기술, 빅데이터 실시간 처리기술, 빅데이터 처리 프로그래밍 지원 기술이 있으며 하둡, R, NoSQL 등의 기술이 대표적으로 사용된다. 빅데이터 일괄 처리기술은 빅데이터를 수집하여 여러 서버로 분산하고 병렬 처리한 후, 이를 다시 수집하여 결과

를 도출하는 기술이다. 여기에는 구글 맵-리듀스, 하둡 맵-리듀스, 마이크로소프트 드라이애드 등의 기술이 있 다[7]. 빅데이터 실시간 처리기술은 실시간으로 들어오는 빅데이터를 처리하기 위한 기술로 분산 환경에서 스트리 밍 데이터를 분석할 수 있는 트위터의 스톰이 있다. 또한 센서 네트워크에서 스트림 처리 기술로 강화된 고성능 컴퓨팅을 지원하는 IBM의 인포스피어 스트림이 있다[8]. 그리고 빅데이터 처리가 가능하고 모든 유형의 분산 및 병렬 데이터 처리가 가능한 프로그래밍 언어가 지속적으 로 개발되고 있다.

빅데이터 분석 기술에 대한 연구로는 실시간 비정형 빅데이터를 분석하기 위한 프레임워크를 개발이 진행중 에 있다. 이는 비정형 데이터의 전처리부와 엔진, 관리 프 로세스로 구성되었으며 실시간 분석을 관리하기 위한 사 용자 인터페이스를 제공한다[9]. 또한 하둡 기반의 빅데 이터 시스템을 이용하여 라이프로그 분석을 통한 상황인 식 서비스를 제공하는 플랫폼이 있다. 이는 사용자의 일 상생활을 기록한 라이프로그로부터 의미있는 상황인식 서비스를 제공해 주는 기술이다[10].

3. 헬스 빅데이터 플랫폼에서 이기종 라이프로그 마이닝 모델

3.1 이기종 라이프로그 기반 분산 파일처리

이기종으로 구성된 라이프로그 데이터는 수집방법이 나 경로에 따라 정형, 비정형, 반정형으로 분류를 한다. 비정형과 반정형 데이터는 특징 추출과 전처리 작업을 통해 정형 데이터로 변환한다[4]. 본 논문에서는 이기종 라이프로그 마이닝 모델을 개발하기 위해 라이프로그. EMR, PHR, 상황정보 등으로 데이터 모델을 구성한다. 이는 실시간으로 추출한 사용자의 정보에 대한 피드백을 통해 개인화된 헬스 서비스를 제공한다[1-3]. 실시간 데 이터 처리에 대한 하드웨어, 소프트웨어적 부하를 줄이 기 위해 하둡 분산파일 시스템의 클러스터 컴퓨팅을 활 용하며, 클라우드 환경에서 빠르게 연결되는 고속 전용 회선과 P2P 데이터 파이프라인을 구성한다[11]. 또한 사 전에 정의된 라이프케어 모델에 따른 일괄처리 시스템 형태의 아키텍처 설계와 별도의 데이터 분석 모델을 구 성하여 공통 데이터 모델[12]을 적용한다. 이를 통해 이 기종 라이프로그를 효율적으로 통합 및 관리한다. 따라 서 데이터 관리의 시간적, 비용적, 보안 및 무결성을 향상 시키는 모델이 구성된다. 또한 이는 병렬적 분산처리 플 랫폼이므로 분석이 필요하여 처리시간이 많이 소요되는 이기종 빅데이터 처리에 효율적이다. 그리고 데이터의 암호화 처리와 접근 제한으로 인해 보안이 뛰어나 헬스 플랫폼에서 의료 데이터 처리에 적합하다.

3.2 이기종 온톨로지 기반 지식베이스 구성

전처리 과정을 거쳐 구성한 이기종 라이프로그 데이 터에 온톨로지 추론엔진을 이용하여 규칙을 발견한다. 발견된 추론 규칙으로 지식베이스를 생성한다. 또한 변 화하는 상황에 따라 추가되는 추론규칙으로 지식베이스 를 확장한다. 온톨로지 기반 지식베이스는 데이터와 속 성들 간의 관계가 명시되어 있으며, 관계에 대한 추론을 통해 지식을 생성한다. 그림 2는 이기종 라이프로그 온톨 로지를 나타낸다. 이는 기존에 존재하는 헬스 온톨로지 [13]에 실시간 서비스와 서비스 확장을 위하여 상위 온톨 로지 방식으로 재구성한다. 상위 온톨로지 모델링은 추 론 장점을 극대화하여 지식표현 방법의 개발 및 획득, 추 론 융합 모델을 구축하고 지식 획득 및 지식 정제를 통해 능동적 지식을 생성한다.

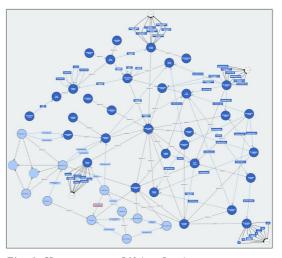


Fig. 2. Heterogeneous Lifelog Ontology

라이프로그는 시간의 흐름에 따라 위치, 행동 상태, 건 강 상태 등의 상황정보가 끊임없이 변화한다[1-3]. 이러 한 특성으로 인해 F-logic과 SWRL 추론방식을 이용한 다. 이를 통한 규칙기반 추론으로 시계열 데이터 변화에 따라 엠비언트 상황에 맞추어 지식베이스를 확장한다 [14]. 실시간 건강관리 서비스를 위한 라이프로그 온톨로 지 추론 모델을 개발하여 지식베이스 기반 적응적인 의사결정 및 지식베이스 확장, 지식 표현 및 획득, 추론을 진행한다. 그림 3은 지식베이스에서 온톨로지 추론 과정을 나타낸다. 이는 사용자가 당뇨환자로 가정된 추론규칙이다. 또한 추론된 RULE은 IF-THEN 형식을 따라 작성되었으며 경우에 따라 입력 변수를 논리곱 한다.

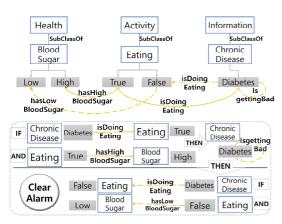


Fig. 3. Ontology Inference Process in Knowledge Base

<IF> 만약, 만성질환으로 당뇨를 앓고 있는 사용자가무언가를 먹는 행동을 하며 <AND> 먹는 행동을 하는 사용자의 혈당이 높아지는 것이 논리곱으로 입력되어진다. <THEN> 그렇다면 사용자가 앓고 있는 만성질환인당뇨가 더 악화될 수 있다는 잠재적 건강 위험을 예측할수 있다. 따라서 이때 사용자에게 혈당이 높아졌다는 경고와 인슐린 투여 같은 적절한 조치를 안내하는 문구를 전달한다. <THEN> 이런 상황에서 새로운 상황정보가들어오게 된다고 가정하여 <IF> 당뇨환자인 사용자가먹는 행동을 멈추었을 때, 일정시간이 지난 후 <AND>혈당도 같이 낮아졌음이 입력된다면 이는 높은 혈당으로인한 잠재적 건강 위험이 사라진다는 결과이다. 따라서사용자에게 안내되고 있던 높은 혈당 알람을 중지한다.이와 같이 사용자의 잠재적 건강 위험에 대한 지식베이스를 구축한다.

3.3 빅데이터 기반 이기종 라이프로그 마이닝

이기종 라이프로그 온톨로지는 다양한 클래스와 속성의 관계를 포함하고 있으며, 라이프로그 중 시계열 생체 신호는 서로 상호작용하여 변화한다. 라이프로그 마이닝 은 시간의 흐름에 따라 변화하는 생체신호에 대한 분석 을 통해 변화를 예측하고, 예측결과를 이용하여 건강위험을 감지한다. 시계열 생체신호가 발생 시간 순으로 측정된다. 이를 구성된 트랙잭션에서 의미적 관계에 대한분석으로 온톨로지에서 나타나지 않은 숨겨진 관계를 탐색하고, 생체신호의 변화를 예측한다. 생체신호는 사용자의 주변의 엠비언트 환경이나 사용자의 활동에 따라 수시로 변화한다. 따라서 라이프로그 마이닝을 통해 사용자의 건강에 이상 징후가 나타날 경우 사전에 감지하여건강위험을 예방한다.

Table 1. Relation of association rules analyzed through life log mining

No.	Biological Signal Correlation	
1	{ultraviolet} => {body temperature}	positive
2	{blood sugar} => {weight}	positive
3	{temperature} => {activity}	negative
4	{activity} => {bmi}	negative
5	{age} => {heart beat}	negative
6	{step count} => {sleep time}	positive
7	{ultraviolet} => {wearing mask}	positive
8	{temperature} => {eczema}	negative
9	{humidity} => {steps}	negative

표 1은 라이프로그 마이닝을 통해 분석한 연관규칙들 의 관계이다. 이는 총 372개의 연관규칙 중에서 빈도가 높은 규칙과 특이한 규칙을 나타낸다. 라이프로그 마이 닝은 데이터간의 빈발항목집합과 후보항목집합의 반복 적인 프로세스로 연관관계를 발견하고 지지도, 신뢰도, 향상도를 이용하여 분석하는 Apriori 알고리즘[4,5]을 사 용한다. 또한 관계 속성은 음의 상관관계와 양의 상관관 계로 나타낸다. 음의 상관관계는 두 속성 a와 b의 상관계 수가 음수일 때, 속성 a가 평균보다 높으면 속성 b는 평 균보다 작고, 속성 a가 평균보다 높으면 속성 b는 평균보 다 낮을 가능성이 높음을 나타낸다. 양의 상관관계는 두 속성 a와 b가 양수일 때, 속성 a가 평균보다 높으면 속성 b역시 평균보다 높을 가능성이 있음을 나타낸다. 표 1에 서 {ultraviolet} => {body temperature}, positive는 자외 선과 체온이 양의 상관관계임을 나타낸다. {temperature} => {activity}, negative는 온도와 활동이 음의 상관관계 임을 나타낸다.

4. 모델 및 성능평가

4.1 이기종 라이프로그 마이닝 모델

이기종 라이프로그 마이닝 모델은 온톨로지 추론규칙과 시계열 생체신호를 이용하여 사용자의 건강에 대한 이상 징후를 예측하고, 이에 대한 행동지침을 제공한다. 웨어러블 엠비언트 센서를 통해 사용자의 라이프로그를 실시간으로 수집하고 클라우드 환경에서 처리한다. 분산클라우드 컴퓨팅에서 사용자의 현재 라이프로그와 관련이 있는 추론규칙을 기반으로 행동요령이나 건강관련 지식을 제공한다. 사용자는 일상생활에서 행동 예측에 대한 의사결정으로 제공받은 정보를 반영하여 건강에 대한위험을 회피하거나 예방할 수 있다. 또한 라이프로그 마이닝을 통해 시계열적 특성을 나타내는 생체신호에 대한변화를 예측하고, 이에 대한 예측 결과가 부정적일 경우사용자에게 경고와 대처방안을 제공한다. 그림 4는 이기종 라이프로그 마이닝 모델 구성도를 나타낸다.

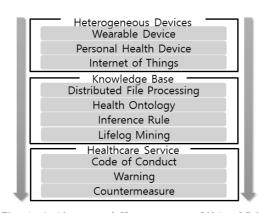


Fig. 4. Architecture of Heterogenonous Lifelog Mining Model

4.2 성능 평가

제안하는 모델에 대한 성능을 평가하기 위해 정확도와 재현율을 평가한다. 제안하는 라이프로그 마이닝 모델(LMM: Lifelog Mining Model)과 단순 온톨로지 모델(SOM:Simple Ontology Model), 시계열 예측 모델(TPM:Time Series Prediction Model)을 비교한다. 성능평가는 사용자의 건강상태에 대한 상황을 내부, 외부, 서비스로 약 500개로 정의한다. 이는 건강보험심사평가원[15]으로부터 제공받은 데이터를 기반으로 구성한다. 각모델을 이용한 사용자 건강 상황을 예측하고 이에 대한

정확도와 재현율을 평가한다. 정확도는 전체 예측 결과 중 예측한 상황과 실제 상황이 같은 비율을 나타낸다. 재현율은 모델이 긍정적으로 예측한 사용자 건강상황이 실제로 긍정적인 경우의 비율을 나타낸다. 그림 5는 정확도와 재현율 평가 결과를 나타낸다. 평가 결과에서 제안하는 이기종 라이프로그 마이닝 모델 방법이 정확도에서 0.734, 재현율에서 0.752로 다른 모델에 비해 우수하게 나타난다. 단순 온톨로지 모델은 실시간으로 변화하는 생체신호의 변화에 대한 고려를 하지 않아 정확도가 낮아지고, 시계열 예측 모델의 경우 사용자의 상황정보를 고려하지 않아 재현율이 낮게 나타난다.

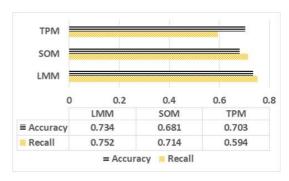


Fig. 5. Performance Result of Accuracy and Recall

5. 결론

본 논문에서는 헬스 빅데이터 플랫폼에서 이기종 라 이프로그 마이닝 모델을 제안하였다. 분산처리와 클라우 드 컴퓨팅 환경에서 데이터 파이프라인을 구성을 통해 실시간으로 라이프로그 데이터의 수집과 서비스가 가능 하도록 한다. 확장성을 높이기 위해 이기종 환경에 적합 한 상위 온톨로지 방식으로 온톨로지를 재구성하였다. 추론엔진을 이용해 추론 규칙들을 생성하고, 규칙 기반 추론 방법으로 실시간 의료 서비스를 제공한다. 예방 의 료를 위해 관계나 추론규칙에서 포함되지 않은 음의 상 관관계나 양의 상관관계를 라이프로그 마이닝을 통해 탐 색하였다. 이는 사용자의 생체신호에 대한 변화를 감지 하고 예방 의료 서비스를 현실화할 수 있는 적응적인 건 강관리 기술이다. 제안하는 모델에 대한 성능 평가를 진 행한 결과가 정확도는 0.734, 재현율은 0.752로 향상된 것 으로 나타났다. 이를 통해 기존의 질병 발생 후 치료 프 로세스의 의료 서비스에서 탈피하여 질병 예방 관리 프

로세스의 건강관리 서비스가 가능해진다. 라이프로그 온 톨로지를 통한 건강관리는 지속적인 만성질환 관리와 잠 재적 건강 위험을 감지할 수 있다.

REFERENCES

- J. K. Kim, J. H. Kim, D. K. Park & Y. H. Lee. (2012).
 U-Health Platform based Health Management Service Model using Context Information. *Journal of Digital Convergence*, 10(8), 185-192.
- [2] H. Jung & K. Chung. (2016). Life Style Improvement Mobile Service for High Risk Chronic Disease based on PHR Platform. Cluster Computing, 19(2), 967–977.
- [3] K. Chung, Y. Na & J. Lee. (2013). Interactive Design Recommendation using Sensor based Smart Wear and Weather WebBot. *Wireless Personal Communications*, 73(2), 243–256.
- [4] J. C. Kim & K. Chung. (2018). Mining Health-Risk Factors using PHR Similarity in a Hybrid P2P Network. Peer-to-Peer Networking and Applications, 11(6), 1278-1287.
- [5] J. S. Lee. (2014). A Study on the Data Mining Preprocessing Tool For Efficient Database Marketing. Journal of Digital Convergence, 12(11), 257–264.
- [6] C. W. Song, H. Jung & K. Chung. (2018). Development of a Medical Big-Data Mining Process using Topic Modeling. Cluster Computing, DOI 10.1007/s10586-017-0942-0.
- [7] S. J. Choi, J. W. Park, J. B. Kim & J. H. Choi. (2014). A Quality Evaluation Model for Distributed Processing Systems of Big Data. *Journal of Digital Contents Society*, 15(4), 533–545.
- [8] S. Sakr. (2013). An Introduction to InfoSphere Streams and InfoSphere BigInsights, IBM, https://www.ibm.com/developerworks/library/bd-stream sintro/.
- [9] J. Kim, N. Kim, H. Yu, S. Yun & B. Lee. (2013). Framework for Unstructured Big Data Analysis. Information Communication Technology Platform, 1(1), 1–4.
- [10] H. Yoo & K. Chung. (2018). Mining-based Lifecare Recommendation using Peer-to-Peer Dataset and Adaptive Decision Feedback. Peer-to-Peer Networking and Applications, 11(6), 1309–1320.
- [11] K. Chung & R. C. Park. (2016). P2P Cloud Network Services for IoT based Disaster Situations Information.

- Peer-to-Peer Networking and Applications, 9(3), 566-577.
- [12] H. Jung & K. Chung. (2016). Knowledge-based dietary nutrition recommendation for obese management. Information Technology and Management, 17(1), 29-42.
- [13] J. H. Kim, J. W. Kim, D. S. Lee & K. Y. Chung. (2014). Ontology Driven Interactive Healthcare with Wearable Sensors. *Multimedia Tools and Applications*, 71(2), 827–841.
- [14] S. K. Kim & K. H. Ahn. (2008). Web Ontology Modeling Based on Description Logic and SWRL. Journal of the Korean Society for Information Management, 25(1), 149–171.
- [15] Health Insurance Review and Assessment Service(HIRA). http://opendata.hira.or.kr/.

강 지 수(Kang, Ji Soo)

[학생회원]



- · 2016년 2월 ~ 현재 : 경기대학교 컴퓨터과학과 학부생
- · 2017년 2월 ~ 현재 : 경기대학교 컴퓨터과학과 데이터마이닝 연구 실 연구원
- · 관심분야 : 데이터 마이닝, 헬스케

어, 빅데이터 분석, 인공지능, 신경망, 추천 시스템

· E-Mail: kangis920@gmail.com

정 경 용(Chung, Kyungyong)

[종신회원]



- · 2000년 2월 : 인하대학교 전자계 산공학과 (공학사)
- · 2002년 2월 : 인하대학교 전자계 산공학과 (공학석사)
- · 2005년 8월 : 인하대학교 컴퓨터 정보공학부 (공학박사)
- · 2006년 3월 ~ 2017년 2월 : 상지대학교 컴퓨터정보공 학부 교수
- 2017년 3월 ~ 현재 : 경기대학교 컴퓨터공학부 교수
- · 관심분야: 데이터 마이닝, 헬스케어, 빅데이터, 지능시 스템, 인공지능, HCI, 정보검색, 추천 시스템
- · E-Mail: dragonhci@gmail.com