

<http://dx.doi.org/10.17703/JCCT.2023.9.3.731>

JCCT 2023-5-86

음성인식과 자연어 처리 딥러닝을 통한 전자의무기록 자동 생성 시스템

Automatic Electronic Medical Record Generation System using Speech Recognition and Natural Language Processing Deep Learning

손현곤*, 류기환**

Hyeon-kon Son*, Gi-hwan Ryu**

요약 최근 의료 현장은 전자의무기록, 전자건강기록 등의 의료 기록을 전산화하여 저장하고 관리하는 시스템이 의무적으로 적용되거나 전체 의료 현장에 보급되어 환자 개개인의 과거 의료 기록을 추가적인 의료 행위에 활용하고 있다. 그러나 일반적인 의료 문진 및 상담 간 발생하는 의료진과 환자 간의 대화는 별도로 기록되거나 저장되지 않고 있어 추가적인 환자의 주요 정보는 효율적으로 활용되지 못하고 있다. 이에 따라, 의료 문진 현장에서 발생하는 의료진과 환자와의 대화를 저장하고 이를 텍스트 데이터로 변환하여 주요한 문진 내용만 자동으로 추출, 요약하여 정보화하는 음성인식과 자연어 처리 딥러닝을 통한 의료 상담 요약문을 자동으로 생성하는 전자의무기록 시스템을 제안한다. 본 시스템은 의료 종사자와 환자의 의료 상담 내용의 인식 과정을 거쳐서 텍스트 정보를 획득한다. 이렇게 획득된 텍스트를 복수의 문장으로 구분하고, 생성된 문장에 포함된 복수 키워드의 중요도를 산출한다. 산출된 중요도를 기반으로 복수의 문장에 순위를 매기고, 순위를 기반으로 문장들을 요약하여 최종 전자 의무기록 데이터를 생성한다. 제안하는 시스템 성능은 정량적 분석을 통하여 우수함을 확인한다.

주요어 : 상담 요약, 전자의무기록, 음성인식, 자연어 처리

Abstract Recently, the medical field has been applying mandatory Electronic Medical Records (EMRs) and Electronic Health Records (EHRs) systems that computerize and manage medical records, and distributing them throughout the entire medical industry to utilize patients' past medical records for additional medical procedures. However, the conversations between medical professionals and patients that occur during general medical consultations and counseling sessions are not separately recorded or stored, so additional important patient information cannot be efficiently utilized. Therefore, we propose an electronic medical record system that uses speech recognition and natural language processing deep learning to store conversations between medical professionals and patients in text form, automatically extracts and summarizes important medical consultation information, and generates electronic medical records. The system acquires text information through the recognition process of medical professionals and patients' medical consultation content. The acquired text is then divided into multiple sentences, and the importance of multiple keywords included in the generated sentences is calculated. Based on the calculated importance, the system ranks multiple sentences and summarizes them to create the final electronic medical record data. The proposed system's performance is verified to be excellent through quantitative analysis.

Key words : Consultation Summary, Electronic Medical Records, Speech Recognition, Natural Language Processing

*정회원, 광운대학교 대학원 실감융합콘텐츠학과 박사과정
(제1저자)

**정회원, 광운대학교 스마트융합대학원 교수 (교신저자)
접수일: 2023년 3월 31일, 수정완료일: 2023년 4월 18일
게재확정일: 2023년 5월 1일

Received: March 31, 2023 / Revised: April 18, 2023

Accepted: May 1, 2023

**Corresponding Author: allryu@kw.ac.kr

Dept. of Professor, Department of Tourism Industry, Graduate
school of smart convergence, Kwangwoon University, Korea

I. 서 론

의료 산업은 인간의 건강과 질병 치료에 관련된 분야에서 매우 중요한 역할을 담당한다. 병원에서는 여러 의료 행위를 수행하고 그에 따른 증거와 정보로써 의무 기록을 생산한다. 이러한 의료 행위는 의료인이 의학적 근거 하에 내리는 지속적인 의사결정과 그 적용 과정이라 할 수 있으며 이 과정은 의무 기록으로 남겨진다. 다른 산업 분야와 마찬가지로 의료산업의 정보화 추세에 따라 의료 기관의 의무로 EMR의 필요성이 대두되고 있다.

EMR(Electronic Medical Record)은 전자의무기록 시스템으로, 환자의 의료 정보를 전자적으로 기록하고 저장하는 시스템이다. EMR의 장점으로는 정보 공유의 효율성 증대, 진단 및 치료의 정확성 확보, 의료 직원의 업무 효율성 향상, 비용 절감 및 보안성 증대 효과 등이 있다. 4차 산업사회 시대에 다른 산업 분야와 마찬가지로 의료 산업의 정보화 추세에 따라 의무적으로 EMR의 필요성이 부각되고 있다.

현재 종합병원은 전문 전사자를 통해 판독의가 영상을 보면서 음성으로 기록한 내용을 전문 전사자를 통해 텍스트로 변환하고 있으나 이들의 이직률이 높아서 인력 부족 현상이 심화되고 있다. 특히 EMR의 근간이 되는 진료 영상을 판독하는 전문의의 부족 현상이 심각해지고 있다. 이로 인하여 환자들의 영상 판독을 위한 대기 시간도 급증하고 있는 상황으로 일부 병원에선 판독 업무를 외주로 돌리고 있으나 그 비용이 매우 높아 비효율적이다. 또한, EMR의 복잡성이 갈수록 증가하고 있는 추세이다. 이를 극복하기 위한 영상 판독과 진단 보조 서비스에 음성인식 기술을 적용하고 자연어 처리 알고리즘을 통한 의료 상담 요약문 생성 시스템이 요구되고 있다[1].

II. 이론적 배경

1. EMR 시스템 요구 기술들

EMR을 구현하기 위해 다양한 기술들이 사용된다[2]. 클라우드 컴퓨팅 기술은 EMR 구현에서 가장 중요한 요소 중 하나로 데이터 저장 및 처리를 위한 적절한 인프라를 보유하는 것이다. 이를 위해 클라우드 기술을 사용하는 것이 효과적일 수 있다. 예를 들어, AWS (Amazon Web Services)의 EMR 서비스는 클라우드

기반의 분산 데이터 처리 플랫폼으로 대규모 데이터 집합을 처리하고 분석하기 위한 다양한 기술들을 포함하고 있다[3].

빅데이터 기술은 EMR은 대규모 데이터를 다루므로 빅데이터 기술이 필요하다. 빅데이터 기술로는 Apache Hadoop, Apache Spark, 인공지능 등이 있다[4]. 데이터 처리 기술은 EMR에는 데이터를 저장하고 검색하기 위한 데이터베이스가 필요하다. 이를 위해 NoSQL 데이터베이스, 예를 들어 MongoDB, Cassandra, DynamoDB 등이 사용될 수 있다[5]. 데이터 시각화 도구는 EMR에 저장된 데이터를 시각화하고 분석하기 위해, 데이터 시각화 도구를 사용할 수 있다. 이를 위해 Tableau, Power BI 등이 사용될 수 있다[6]. 보안 기술로는 EMR에는 민감한 환자 정보가 포함되므로, 보안 기술이 필요하다. 이를 위해 데이터 암호화, 액세스 제어, 보안 로그 등의 기술이 사용될 수 있다[7][8]. 마지막으로, 인공지능 기술은 EMR에서는 예측 모델링, 자연어 처리 등의 인공지능 기술을 사용하여 진단, 치료 등의 의료 서비스를 개선할 수 있다. 이를 위해 딥러닝, 강화학습 등의 기술이 사용될 수 있다[9][10].

2. 음성 처리 기술

음성 처리 기술은 인간의 음성을 기계가 인식하고 분석하는 기술이다. 여러 응용 분야에서 활용되고 있다. 의료 분야에서 음성 인식 및 처리는 EMR, 음성 기반 수술 로봇, 중증 장애인을 위한 환경 제어 분야, 그리고 언어 장애 치료 등에서 적용되는 기술이다. 특히, EMR 시스템에서는 음성처리 기술은 의사 혹은 간호사와 환자 사이 대화를 전용 마이크나 헤드셋을 통하여 대화를 인식하고 처리하여 텍스트를 생성하도록 돕는 것이다.

음성 처리 기술 분야는 음성인식 (Speech Recognition), 음성 합성 (Speech Synthesis), 음성 분석 (Speech Analysis), 음성 기반 검색 (Voice-based Search), 음성 감정 인식 (Speech Emotion Recognition), 음성 인식 기반 대화형 인터페이스 (Voice-based Interactive Interface) 등이 있다. EMR 시스템 구현에서 음성 인식 기술은 최근에 중요성이 점점 커지고 있다.

음성 인식은 음성을 텍스트로 변환하는 기술이다. 음성 인식은 음성 인식 엔진을 통해 구현할 수 있다. 음성 인식은 자연어 처리, 기계 학습 및 신호 처리 기술을 활용하여 구현된다.

3. 자연어 처리 기술

자연어 처리(Natural Language Processing, NLP)는 컴퓨터와 인간 언어 사이의 상호 작용, 특히 컴퓨터가 대량의 자연어 데이터를 처리하고 분석하도록 프로그래밍하는 방법과 관련된 언어학, 컴퓨터 과학 및 인공지능의 하나의 분야이다. NLP의 궁극적인 목표는 문서 내 언어의 문맥적 뉘앙스를 포함하여 문서의 내용을 '이해'할 수 있는 컴퓨팅 작업이다. NLP 기술은 문서에 포함된 정보와 통찰을 정확하게 추출할 수 있을 뿐만 아니라 문서 자체를 분류하고 정리할 수 있다. NLP의 주요 과제는 음성 인식, 자연어 이해 및 자연어 생성 등이 있다.

초기에, 많은 NLP 시스템들은 문법을 쓰거나 정지를 위한 경험적 규칙(Rules, Knowledges)을 데이터베이스화한 사전 검색과 결합한 일련의 규칙들의 수동 코딩과 같은 상징적인 방법들에 따라 설계되었다. 기계학습 알고리즘을 기반으로 하는 최근의 시스템은 수작업으로 만든 규칙보다 많은 장점이 있다.

4. 레코드 생성 기술

레코드 생성 기술은 데이터 저장 및 관리를 위한 방법 중 하나이며 데이터베이스, 전자문서, 음성 녹음, 이미지 파일 등 다양한 형태의 데이터를 저장하는 데 사용된다. 최신 텍스트 생성 기술 중 하나는 GPT-3 (Generative Pre-trained Transformer 3)이다. GPT-3는 OpenAI에서 개발한 인공지능 언어 모델로, 높은 자연어 이해 능력을 갖추고 있어 대화형 대답 생성, 문장 완성, 번역, 요약 등의 작업을 수행할 수 있다. GPT-3는 175억 개의 매개 변수를 가지고 있으며, 이는 이전 모델들보다 대폭 증가한 수치이다. 다른 최신 텍스트 생성 기술로는 GPT-3와 유사한 구조를 갖는 GPT-Neo는 오픈 소스 기반으로 개발되어 사용자가 모델을 수정하고 훈련하는 장점이 있다. 다른 기술 중 하나는 GShard이다. GShard는 구글에서 개발한 다중 GPU 클러스터에서 효율적인 모델 학습을 가능하게 하는 기술로, 대규모 텍스트 생성 작업을 빠르고 효율적으로 처리할 수 있다. 이 외에도 BERT (Bidirectional Encoder Representations from Transformers), T5 (Text-to-Text Transfer Transformer) 등 다양한 텍스트 생성 기술들이 개발되고 있으며, 이러한 기술들은 텍스트 생성 분야에서 높은 성능을 보여주고 있다.

5. 메디컬 데이터 분석

메디컬 데이터 분석은 의료 분야에서 수집된 데이터를 분석하여 의학적인 식견에 도달하는 과정이다. 메디컬 데이터 분석은 건강한 인구에서부터 질병을 가진 환자까지 다양한 환경에서 수집된 데이터를 포함한다. 예를 들어, 환자의 진단서, 의료 기기에서 수집된 생체 신호, 환자의 건강 기록, 의료 직원의 일일 업무 기록 등이 해당된다. 메디컬 데이터 분석을 통해 의료 전문가들은 환자의 건강 상태를 파악하고, 치료 계획을 세우고 예방 전략을 개발하는 데에 활용할 수 있다. 또한, 의료 데이터 분석은 의료 시스템을 개선하고 비용을 절감하는 데에도 사용될 수 있다. 메디컬 데이터 분석은 다양한 기술과 도구를 사용하여 수행된다. 예를 들어, 데이터 마이닝, 기계학습, 통계 분석 등이 사용될 수 있다.

III. 연구내용 및 방법

1. 시스템 구성

음성 인식과 자연어 처리 딥러닝을 이용하여 의료 상담 요약문과 EMR을 생성하는 시스템은 입력부, 음성 인식부, 텍스트 분석부, 저장부, 조회부 다섯 부분으로 구성되었다.

입력부는 의료 상담이 녹취된 오디오 신호를 입력받기 위한 구성이다. 오디오 신호를 시스템에 업로드하기 위한 음성 업로드 모듈을 포함하며, 마이크를 통해 입력된 의료 상담을 기반으로 음성을 업로드한다.

음성 인식부는 의료 상담이 녹취된 오디오 신호를 인식하여 텍스트로 변환한다. 음성 인식부는 음향 모델 기반 인식 모듈, 음절 추출 모듈, 언어 모델 기반 인식 모듈, 텍스트 생성 모듈 등을 포함한다. 음향 모델 기반 인식 모듈은, 음향 모델을 통해 오디오 신호를 분석하여 하나의 음소 내지는 특정 단위에 해당하는 소리를 추출하기 위한 모듈이다. 음절 추출 모듈은 음향 모델 기반 인식 모듈을 통해 추출된 하나의 음소를 이용하여 음절을 추출한다. 언어 모델 기반 인식 모듈은 언어 모델을 통해 음소 또는 음절을 확률적으로 조합하여 하나의 단어 내지는 문장을 생성하기 위한 모듈이다[11]. 텍스트 생성 모듈은 상술한 모듈들을 통해 수행된 분석 결과에 따라 오디오 신호에 매칭되는 전체 텍스트를 획득한다.

제안된 시스템의 흐름과 기능적 구성을 도식화하기 위한 블록도는 그림 1과 같다.

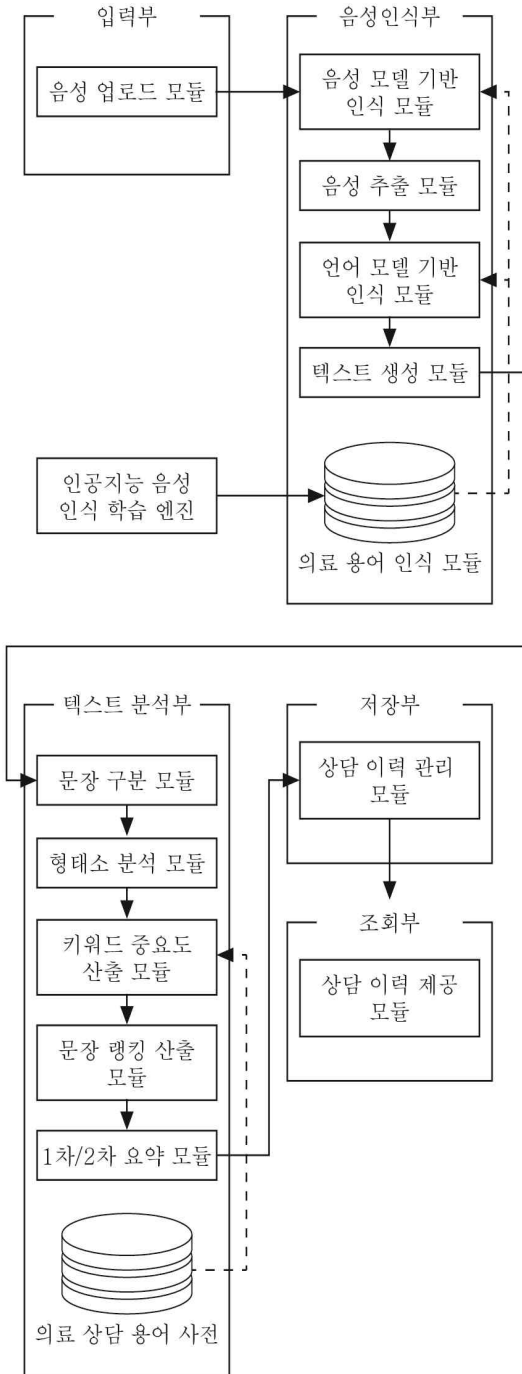


그림 1. 의료 상담 요약문과 전자의무기록 생성 시스템 블록도
 Figure 1. Block diagram for generating medical service summaries and EMRs system

텍스트 분석부는 인식된 텍스트를 분석하여 요약하기 위한 구성이다. 문장 구분 모듈, 형태소 분석 모듈, 키워드 중요도 산출 모듈, 문장 랭킹 산출 모듈, 1차/2차 요약 모듈 등을 포함한다. 문장 구분 모듈은 텍스트에 포함된 문장 부호 내지는 어미를 식별하여 텍스트를 복수의 문장으로 구분한다. 형태소 분석 모듈은 구분된 복수의 문장 각각에 포함된 키워드들을 형태소 단위로 추출한다. 키워드 중요도 산출 모듈은 추출된 각 키워드의 중요도를 산출한다. 이때, 키워드 중요도 산출 모듈은 각 키워드의 출현 빈도(ex. 다양한 의료 상담 내용)를 통해 각 키워드의 중요도를 산출한다. 또한, 키워드 중요도 산출 모듈은 의료 상담 용어 사전에 포함된 단어와 매칭되는 키워드에 대해서는 중요도에 가중치를 부여할 수도 있다. 문장 랭킹 산출 모듈은 키워드 각각의 중요도를 이용하여 각 문장에 순위를 매길 수 있다. 각 문장에 포함된 키워드들의 중요도를 합산하여 각 문장의 스코어를 산출한다. 그리고, 문장 랭킹 산출 모듈은 점수가 높을수록 순위가 높아지도록 각 문장의 순위를 매길 수 있다. 1차/2차 요약 모듈은 상술한 복수의 문장 중 일부를 이용하여 요약을 수행하기 위한 모듈이다. 복수의 문장 중 기설정된 요약물에 대응되는 수만큼의 문장들을 선택하며, 선택된 문장들 내에서 주격/목적격 조사를 제외한 조사를 제외한다.

저장부는 요약된 텍스트(요약문)를 저장 및 관리하기 위한 모듈이다. 상담 이력 관리 모듈을 포함한다. 상담 이력 관리 모듈은 특정 환자에 대해 수행된 복수의 의료 상담을 기반으로 각각 획득된 복수의 요약문을 환자에 대한 상담 이력으로 저장한다. 또한, 상담 이력 관리 모듈은 새롭게 수행된 의료 상담에 따라 획득된 요약문을 추가하여 환자의 상담 이력을 업데이트한다. 상담 이력 관리 모듈은 의료 상담 전체가 담긴 원문 텍스트 및 요약문을 각각 상담 이력으로 등록할 수도 있다. 이 경우, 상담 이력 관리 모듈은 원문 텍스트가 등록된 시점으로부터 제1 기간이 지난 시점에 원문 텍스트를 삭제하고, 요약문이 등록된 시점으로부터 제2 기간이 지난 시점에 요약문을 삭제한다. 여기서, 제2 기간은 제1 기간보다 길 수 있다. 즉, 원문 텍스트가 아닌 요약문이 장기 보관됨으로써, 서버의 메모리 용량이 효율적으로 사용된다. 또는, 상담 이력 관리 모듈은 요약문은 시간이 지나도 삭제하지 않되, 등록된 원문 텍스트만 설정된 기간 이후에 삭제할 수도 있다.

조희부는 환자의 상담 이력을 조회하며 상담 이력 제공 모듈로 구성된다. 환자 각각의 상담 이력을 제공하며, 요약문을 포함하여 환자의 의료 상담 이력을 제공하여 의료 상담을 수행할 수 있게 서비스를 제공한다.

2. 의료 상담 요약문과 전자의무기록 생성

음성인식과 자연어 처리 딥러닝을 통한 전자의무기록 자동 생성 시스템으로 환자의 이름이 입력되면 해당 환자에 대한 상담 이력에 포함된 하나 이상의 요약문인 상담 내용을 제공하며, 상담 이력 제공 모듈은 입력된 키워드에 따라 요약문을 검색하여 제공할 수도 있다.

예를 들어, 식사 빈도, 흡연 빈도와 같은 키워드에 따라 검색이 수행되는 경우, 상담 이력 제공 모듈은 해당 키워드와 관련된 생활 정보가 담긴 요약문을 제공한다. 이러한 요약문은, 향후 수행되는 해당 환자의 의료 상담 내지는 의료 분석에 유의미하게 활용될 수 있다.

요약된 의료 상담의 최종 요약문을 이용하여, 환자의 생활 방식, 병력 등을 분석한다. 그리고, 환자의 생활 방식 및 병력을 이용하여 환자에 대하여 예측되는 질병을 판단하고, 판단된 질병에 대한 정보(질병 종류, 예방 수칙, 진단 주기 등)를 제공할 수도 있다. 이를 위해, 생활 방식 및 과거 병력에 대한 정보가 입력되면 향후 예측되는 질병에 대한 정보를 출력하도록 훈련된 인공지능 모델을 이용한다.

또한, 요약된 의료 상담의 최종 요약문을 이용하여, 의료 상담을 수행하기 위한 인공지능 모델을 훈련한다. 즉, 실제 의사와 환자 간의 대화를 요약하여 요약문을 획득한 뒤, 획득된 요약문을 의료 상담용 인공지능 모델의 훈련 데이터로 이용한다.

인공지능 모델은, 상담의 어조나 말투 또는 상황 등과 무관하게 의료 상담 내 객관적 정보를 포함하는 주요 문장만을 기반으로 빠르게 훈련하며, 결과적으로는, 의료 상담을 빠르게 수행하여 객관적인 의료 정보만을 기계적으로 제공하도록 훈련된다. 이러한 훈련 방식은, 실제 의사처럼 인간적 면모를 보이거나 종합적인 외부 상황을 고려하지는 못하지만, 객관적 또는 기계적 진단만을 수행하는 인공지능 모델의 역할을 명확히 하며, 사진 상담, 간이 진료 등에 해당하는 간편 의료 서비스 등의 다양한 의료 상담과 진료 분야의 인공지능 모델 훈련에 적합하다.

IV. 실험 및 결과

한국건설생활환경시험연구원(KCL)을 통해 시스템의 정량적인 시험과 측정을 진행하였다. 음성 채널 처리는 주변음, 잡음 등의 음향 전처리와 멀티채널의 스트리밍 기술을 사용하였으며, 대량의 의무기록 데이터를 훈련한 음성 인식 엔진은 음향 모델, 언어 모델로 의료 전문 음성 인식 모델을 사용하였다. 음성 인식을 통해 변환된 의료 텍스트를 자연어 처리(NLP)를 통해서 식별하며 EMR 생성 엔진으로 처리된 텍스트는 오류 자동 수정과 휴먼 보정 프로세스를 거쳐 완성된 형태의 전자의무기록으로 생성한다. 키워드 추출, 주제 분류, 텍스트 요약, 의도 및 탐지 등의 텍스트 분석을 하였으며 풀 텍스트 검색엔진을 통해 대량의 텍스트에 손쉽게 접근하도록 지원하였으며 대시보드를 통해 EMR 데이터를 관리하는 분석 시각화를 위한 웹 서비스를 진행하였다. 표 1은 계량적 목표 및 실험 결과값을 나타낸다.

표 1. 주요 성능 지표의 실험 및 결과

Table 1. Experiments and results of key performance indicators

실험	주요 성능 지표	결과
1	의료 음성인식 정확도	96.98%
2	실시간 자동보정 속도	1ms
3	분석 서비스 종류	5종
4	분석 지원 진료 과목	3개
5	음성 입력 기기	2종

음성인식 정확도는 소화기내과, 복부내과, 흉부외과 영상판독 녹취파일은 10건씩 무작위 추출하여 해당 기록에 대해 96.98% 정확도를 달성하였으며, 실시간 보정 속도는 0.5s 이하로 기록문에 대한 영문/숫자/특수문자 변환 및 오타자 교정이 필요한 120개 문장을 측정하여 문장당 보정 속도 0.001s를 달성하였다. 분석 서비스 종류는 5종 이상으로 3개 진료과의 2개월분 데이터를 대상으로 요약, 병증 분류, 환부 분류, 위험요소 판단, 키워드 분석 등의 5종의 분석 서비스를 제공하며, 소화기내과, 흉부외과, 복부내과의 3개 진료과목을 지원할 수 있도록 하였다. 음성 입력 기기는 2종 이상으로 영상판독에서 의사들이 주로 쓰는 전용 마이크와 전사자들이 주로 쓰는 헤드셋 등 2종의 장치가 작동되도록 개발하였다.

V. 결론

본 논문은 의료 문진과 상담 내용을 자동으로 추출, 요약하여 지식화하는 음성인식과 자연어 처리 딥러닝을 통한 EMR 시스템을 제안하였다. 제안된 시스템은 의료 종사자와 환자 간의 대화를 전기 오디오 신호로 변환하여 인공지능 자연어 처리 딥러닝 기술을 이용해 의료 상담 요약문과 전자의무기록을 자동으로 생성한다. 이를 통해 환자 개인의 과거 의료 기록을 효율적으로 활용할 수 있으며 의료 현장에서의 업무 효율성과 품질을 향상할 수 있다. 시스템의 정량적 분석 결과 제안된 시스템은 5가지 측면에서 우수한 성능을 보였다.

이러한 결과를 토대로 본 논문은 의료 분야에서의 음성인식 기술과 자연어 처리 딥러닝 기술의 활용 가능성을 제시하였으며, 이 연구가 실제 의료 현장에서 적용되어 의료 서비스 품질 향상에 이바지하게 되기를 기대한다.

References

- [1] Wang, Samuel J., et al., "A cost-benefit analysis of electronic medical records in primary care.", *The American journal of medicine*, Vol. 114. No. 5, pp. 397-403, April 2003. DOI: [https://doi.org/10.1016/S0002-9343\(03\)00057-3](https://doi.org/10.1016/S0002-9343(03)00057-3)
- [2] Noraziani, K., et al., "An overview of electronic medical record implementation in healthcare system: Lesson to learn.", *World Applied Sciences Journal*, Vol. 25, No. 2, pp. 323-332, September 2013. DOI: <https://doi.org/10.5829/idosi.wasj.2013.25.02.2537>
- [3] Li, Zhuo-Rong, et al., "A secure electronic medical record sharing mechanism in the cloud computing platform.", 2011 IEEE 15th international symposium on consumer electronics (ISCE). IEEE, pp. 98-103, August 2011. DOI: <https://doi.org/10.1109/ISCE.2011.5973792>
- [4] SARANGI, Susanta; SAHIDULLAH, Md; SAHA, Goutam. Optimization of data-driven filterbank for automatic speaker verification. *Digital Signal Processing*, Vol. 104, September 2020:102795. DOI: <https://doi.org/10.1016/j.dsp.2020.102795>
- [5] Shao, Shih Chieh, et al., "The Chang Gung Research Database—a multi institutional electronic medical records database for real world epidemiological studies in Taiwan.", *Pharmacoepidemiology and drug safety* Vol. 28, No. 5, pp. 593-600, January 2019. DOI: <https://doi.org/10.1002/pds.4713>
- [6] Huang, Edward W., Sheng Wang, and ChengXiang Zhai, "VisAGE: integrating external knowledge into electronic medical record visualization.", *PACIFIC SYMPOSIUM ON BIOCOMPUTING 2018: Proceedings of the Pacific Symposium*, pp. 278-589, January 2018. DOI: https://doi.org/10.1142/9789813235533_0053
- [7] Wu, Sihua, and Jiang Du, "Electronic medical record security sharing model based on blockchain.", *Proceedings of the 3rd International Conference on Cryptography, Security and Privacy*, pp. 13-17, January 2019. DOI: <https://doi.org/10.1145/3309074.3309079>
- [8] Adamu, Jibril, Raseeda Hamzah, and Marshima Mohd Rosli, "Security issues and framework of electronic medical record: A review.", *Bulletin of Electrical Engineering and Informatics*, Vol. 9, No. 2, pp. 565-572, April 2020. DOI: <https://doi.org/10.11591/eei.v9i2.2064>
- [9] Shinozaki, Ayaka, "Electronic medical records and machine learning in approaches to drug development.", *Artificial intelligence in Oncology drug discovery and development*. IntechOpen, 2020. DOI: <https://doi.org/10.5772/intechopen.92613>
- [10] Shickel, Benjamin, et al., "Deep EHR: a survey of recent advances in deep learning techniques for electronic health record (EHR) analysis.", *IEEE journal of biomedical and health informatics*, Vol. 22, No. 5, pp. 1589-1604, September 2018. DOI: <https://doi.org/10.1109/JBHI.2017.2767063>
- [11] Kihun Nam, "A Study on Processing of Speech Recognition Korean Words", *The Journal of the Convergence on Culture Technology*, Vol. 5, No. 4, pp. 407-412, November 2019. DOI: <https://doi.org/10.17703/JCCT.2019.5.4.407>