



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2022년09월19일
(11) 등록번호 10-2444460
(24) 등록일자 2022년09월14일

(51) 국제특허분류(Int. Cl.)

G16H 10/20 (2018.01) G06F 16/33 (2019.01)
G06F 16/332 (2019.01) G06F 40/284 (2020.01)
G16H 50/20 (2018.01) G16H 50/30 (2018.01)
G16H 50/50 (2018.01) G16H 80/00 (2018.01)

(52) CPC특허분류

G16H 10/20 (2021.08)
G06F 16/3329 (2019.01)

(21) 출원번호 10-2021-0114558

(22) 출원일자 2021년08월30일

심사청구일자 2021년08월30일

(56) 선행기술조사문헌

KR101785780 B1*

(뒷면에 계속)

(73) 특허권자

주식회사 초메디신

대전광역시 유성구 대학로 99, 408호 대전팀스타운(궁동, 충남대학교)

세종대학교 산학협력단

서울특별시 광진구 능동로 209 (군자동, 세종대학교)

(72) 발명자

조상연

대전광역시 유성구 전민로 71, 112동 906호(전민동, 삼성푸른아파트)

이승원

서울특별시 광진구 능동로23길 11-6, 304호(군자동)

(뒷면에 계속)

(74) 대리인

특허법인오암

전체 청구항 수 : 총 1 항

심사관 : 김채연

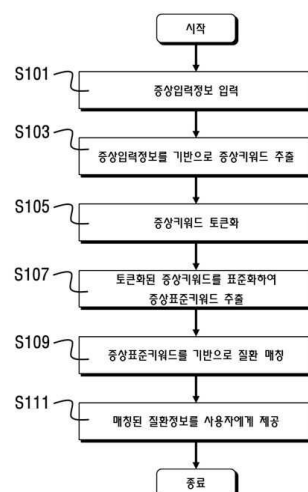
(54) 발명의 명칭 인공지능 기반의 증상 및 질환 매칭을 위한 챗봇 서비스 제공방법

(57) 요약

본 발명은 인공지능 기반의 증상 및 질환 매칭을 위한 챗봇 서비스 제공방법에 관한 것이다.

보다 상세하게는 본 발명의 일 실시예에 따른 장치에 의하여 수행되는 인공지능 기반의 증상 및 질환 매칭을 위한 챗봇 서비스 제공방법에 있어서, 사용자로부터 자연어 기반의 증상입력정보를 입력받는 단계, 상기 증상입력정보를 기반으로 증상키워드를 추출하는 단계, 상기 증상키워드를 토큰화하는 단계, 토큰화된 상기 증상키워드를 표준화하여 증상표준키워드를 추출하는 단계, 상기 증상표준키워드를 기반으로 질환과 매칭하는 단계 및 매칭된 질환정보를 사용자에게 제공하는 단계를 포함할 수 있다.

대표도 - 도1



(52) CPC특허분류

G06F 16/3334 (2019.01)

G06F 40/284 (2020.01)

G16H 50/20 (2018.01)

G16H 50/30 (2018.01)

G16H 50/50 (2018.01)

G16H 80/00 (2021.08)

H04L 51/02 (2022.05)

(72) 발명자

고현용

대전광역시 서구 월평동로 45, 108동 802호(월평동, 진달래아파트)

한해성

서울특별시 노원구 공릉로59길 25, 702동 104호(하계동, 청솔아파트)

김은선

서울특별시 서대문구 북아현로14길 9-6(북아현동)

김은

경기도 부천시 신흥로275번길 42, 213동 2202호(약대동, 약대두산위브 트레지움)

이현진

서울특별시 강남구 삼성로 150, 108동 1403호(대치동, 한보미도맨션)

정정민

경기도 용인시 기흥구 예현로 15, 103동 1102호(서천동, 서그내마을에스케이아파트)

이재훈

대구광역시 수성구 들안로 360, 104동 2504호(수성동4가, 수성 태영 데시앙)

(56) 선행기술조사문헌

KR1020210057308 A*

KR102221126 B1*

KR1020190074130 A

KR102269425 B1

KR102246827 B1

KR1020160145978 A

KR1020200117118 A

*는 심사관에 의하여 인용된 문헌

명세서

청구범위

청구항 1

장치에 의하여 수행되는 인공지능 기반의 증상 및 질환 매칭을 위한 챗봇 서비스 제공방법에 있어서,
 사용자로부터 자연어 기반의 증상입력정보를 입력받는 단계;
 상기 증상입력정보를 기반으로 증상키워드를 추출하는 단계;
 상기 증상키워드를 토큰화하는 단계;
 토큰화된 상기 증상키워드를 표준화하여 증상표준키워드를 추출하는 단계;
 상기 증상표준키워드를 기반으로 질환과 매칭하는 단계; 및
 매칭된 질환정보를 사용자에게 제공하는 단계;를 포함하며,
 상기 토큰화하는 단계는,
 Soynlp, koNLPy 또는 토큰라이저 중 적어도 하나 이상을 이용하여 토큰화하고,
 상기 증상표준키워드 중에서 질환 매칭을 위한 필수항목의 존재여부를 확인하는 단계; 및
 상기 증상표준키워드 중에서 상기 필수항목이 존재하지 않은 경우, 기설정된 추가문진매뉴얼을 상기 사용자에게 제공하는 단계;를 더 포함하며,
 질환 매칭을 위하여 제1 기계학습모델을 통하여 상기 증상표준키워드마다 상기 증상표준키워드에 대응하는 가중치를 출력하는 단계;를 더 포함하고,
 상기 매칭하는 단계는,
 상기 증상표준키워드 및 상기 가중치를 입력으로 매칭되는 질환 및 질환가능성을 출력하는 제2 기계학습모델을 통하여 질환과 매칭하며,
 상기 제1 기계학습모델 또는 제2 기계학습모델을 학습시키기 위해 데이터를 전처리하되, 데이터 검토 및 정리와, 질환 관련 작업과, 데이터 전처리를 수행하며,
 상기 데이터 검토 및 정리는, 오타자 수정과, 의사가 환자를 직접 진찰하여 획득한 정보들을 제외하는 진단 결과 값 제외와, 의학 전문용어들을 보편적 단어로 수정하는 전문용어 수정과, 영어-한국어 번역을 수행하고,
 상기 질환 관련 작업은, 질환명을 한국어로 번역하는 질환 번역과, 둘이상의 질환들을 동일하게 표기하거나, 하나의 상위 질환으로 통합하는 질환 통합을 수행하며,
 상기 데이터 전처리는, 결측값이 과반수 이상인 행을 제거하는 결측값 통일 및 과반수 이상 데이터 제거와, 질병명 번역 및 통합과, 체질량 지수(BMI)를 계산하여 비만도를 정의하는 체중 상태, 나이 그룹화 및 성별 표현 정리와, 환자의 발현 증상인 positive value만 추출하는 Negative value 제거와, 수치화된 통증정도를 텍스트로 변환하는 NRS 점수 텍스트화와, 질병 예측에 사용할 수 있는 열들을 모두 연결하는 한 문장으로 만들기와, 맞춤법 교정과, 유저의 입력을 고려한 용어 정리와, 형태소 분석과, 불용어 및 불필요한 기호 제거를 순차 수행하는 인공지능 기반의 증상 및 질환 매칭을 위한 챗봇 서비스 제공방법.

청구항 2

삭제

청구항 3

삭제

청구항 4

삭제

청구항 5

삭제

발명의 설명

기술 분야

[0001] 본 발명은 인공지능 기반의 증상 및 질환 매칭을 위한 챗봇 서비스 제공방법에 관한 것이다. 보다 상세하게는 사용자로부터 증상에 대한 정보를 입력받아 인공지능 기반으로 질환 매칭을 수행하되, 챗봇 기반으로 정보를 제공하는 기술에 관한 것이다.

배경 기술

[0002] 본 명세서에서 달리 표시되지 않는 한, 이 섹션에 설명되는 내용들은 이 출원의 청구항들에 대한 종래 기술이 아니며, 이 섹션에 포함된다고 하여 종래 기술이라고 인정되는 것은 아니다.

[0003] 대부분 일반 환자들이 찾는 병원은 일차 의료기관이다. 일차 의료기관의 보편적인 진료시간은 평일 오전 9시~오후 6시이다. 하지만 모든 환자들이 이 시간대에 병원을 방문할 수 있는 것은 아니다. 또한 점심시간은 최소 1시간 휴진이며 토요일은 오전 진료만 보는 곳이 많다. 만약 공휴일에 증상이 발생한다면 제시간에 병원에 방문하는 것이 더욱 어려워진다.

[0004] 2020년 경증환자 대형병원 진료비 부담이 높아졌고, 공공의료와 일차 의료기관 부족으로 의료비가 증가하였다. 지난 20년간의 연도별 건강보험 진료비 추이를 살펴보면 해가 갈수록 진료비가 증가하는 경향을 알 수 있다. 뿐만 아니라 65세 이상 노인 진료비 또한 꾸준히 증가해 진료비 부담으로 인한 환자들의 병원 방문이 어려운 경우를 보여준다.

[0005] 이처럼 병원 방문 시간과 맞지 않거나 진료비 부담, 코로나로 인한 병원 방문 걱정, 경미한 증상이라는 생각 등 여러 요인으로 인해 환자가 제때 병원에 가지 못하는 경우가 많다. 필요할 때 병원에 가지 못하면 시간이 지나면서 증세가 심각해지고 조기 발견이 필요한 질병들이 악화될 위험이 있다. 따라서 와이파이 통신을 이용해 스마트폰으로 시간과 장소의 제약 없이 진단 받을 수 있게 된다면 환자들의 편의가 증가하고 악화 위험을 줄일 수 있을 것이다.

[0006] 추가적으로, 환자는 병원을 방문할 때 자신의 주요 증상에 따라 해당 증상을 다루는 과에 찾아가서 진료를 받는다. 이 과정에서 자신이 실제로 앓고 있는 병과 맞지 않는 과를 찾아가서 정확하지 못한 진단을 받거나, 여러 과를 옮겨 다니며 진료과정이 번거로운 경우가 발생할 수도 있다. 그러나 환자가 의사와 대면하지 않고 간단하게 진단을 받을 수 있다면, 본인의 질병 분류 범위를 알게 되어 다음 병원 방문에 도움이 될 것이다.

[0007] 또한, 인포데믹(Infodemic)에 대한 문제도 있다. 인포데믹은 정보(information)와 전염병(epidemic)의 합성어로, 잘못된 정보가 미디어·인터넷 등의 매체를 통해 급속하게 퍼져나가는 것이 전염병과 유사하다는 데서 생겨난 용어이다. 최근 코로나 19로 인해 많은 사람이 알게 된 '인포데믹'이라는 용어는 2003년 5월 미국 전략분석기관 인텔리브리지(Intellibridge)의 창립자 데이비드 로스코프(David Rothkopf)가 워싱턴포스트에 기고한 글에서 처음으로 '정보전염병(인포데믹)'이라는 용어를 사용하면서 알려졌다.

[0008] 특히, 전문성이 필요한 의학정보의 인포데믹은 심각하다. 한국에서 가장 유명한 포털사이트인 네이버에만 하더라도 출처가 불분명한 잘못된 의학정보들이 널리 퍼져 있고, 가짜 정보 및 잘못된 의학정보로 인한 사례, 각종 SNS에서 검증되지 않은 정보가 다수의 '좋아요'를 받으며 사실로 인식되는 경우도 어렵지 않게 찾아볼 수 있다.

[0009] 정보화 시대에는 잘못된 정보도 대량으로 생산되기 때문에 전문적인 지식 없이는 잘못된 정보를 의심없이 사실로 받아들이는 경우가 많아 '인포데믹'이 가속화되고 있다. 환자가 잘못된 정보를 믿고 의사에게 진단을 받으러 오면 오히려 의사를 불신하고 여러가지 검사를 요구하여 정확한 진단을 방해하고 의료자원의 낭비를 초래하는 문제 또한 발생한다.

[0010] 환자가 잘못된 의료정보를 접하게 되는 가장 큰 이유는 병에 대한 두려움 때문이다. 병에 대한 두려움이 환자의

조바심으로 이어져 검색으로 획득한 불분명한 출처의 정보를 사실로 받아들이는 것이다. 부정확한 정보확산을 예방하기 위해서는 신뢰성 있는 정보를 환자가 손쉽게 제공받는 환경이 마련되어야 한다. 따라서 환자에게 전문성 있는 진단 결과와 질환 정보를 제공하는 시스템이 필요해지고 있다.

선행기술문헌

특허문헌

[0011] (특허문헌 0001) 등록특허공보 제10-2262278호

발명의 내용

해결하려는 과제

[0012] 개시된 실시 예는 사용자의 증상정보를 기반으로 질환을 진단하는 챗봇 서비스를 제공하는 것이다.

[0013] 또한, 개시된 실시 예는 사용자의 증상정보를 자연어로 입력받고, 이를 표준화하여 인공지능기반으로 매칭되는 질환을 도출하는 것이다.

[0014] 본 발명에서 이루고자 하는 기술적 과제들은 이상에서 언급한 기술적 과제들로 제한되지 않으며, 언급하지 않은 또 다른 기술적 과제들은 아래의 기재로부터 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에게 명확하게 이해될 수 있을 것이다.

과제의 해결 수단

[0015] 상술한 목적을 달성하기 위하여, 본 발명의 일 실시예에 따른 장치에 의하여 수행되는 인공지능 기반의 증상 및 질환 매칭을 위한 챗봇 서비스 제공방법에 있어서, 사용자로부터 자연어 기반의 증상입력정보를 입력받는 단계, 상기 증상입력정보를 기반으로 증상키워드를 추출하는 단계, 상기 증상키워드를 토큰화하는 단계, 토큰화된 상기 증상키워드를 표준화하여 증상표준키워드를 추출하는 단계, 상기 증상표준키워드를 기반으로 질환과 매칭하는 단계 및 매칭된 질환정보를 사용자에게 제공하는 단계를 포함할 수 있다.

[0016] 이 때, 상기 토큰화하는 단계는, Soynlp, koNLPy 또는 토큰나이저 중 적어도 하나 이상을 이용하여 토큰화할 수 있다.

[0017] 이 때, 본 발명의 일 실시예에 따른 인공지능 기반의 증상 및 질환 매칭을 위한 챗봇 서비스 제공방법은 상기 증상표준키워드 중에서 질환 매칭을 위한 필수항목의 존재여부를 확인할 수 있다.

[0018] 이 때, 본 발명의 일 실시예에 따른 인공지능 기반의 증상 및 질환 매칭을 위한 챗봇 서비스 제공방법은 상기 증상표준키워드 중에서 상기 필수항목이 존재하지 않은 경우, 기설정된 추가문진매뉴얼을 상기 사용자에게 제공할 수 있다.

[0019] 이 때, 본 발명의 일 실시예에 따른 인공지능 기반의 증상 및 질환 매칭을 위한 챗봇 서비스 제공방법은 질환 매칭을 위하여 제1 기계학습모델을 통하여 상기 증상표준키워드마다 상기 증상표준키워드에 대응하는 가중치를 출력할 수 있다.

[0020] 이 때, 상기 매칭하는 단계는, 상기 증상표준키워드 및 상기 가중치를 입력으로 매칭되는 질환 및 질환가능성을 출력하는 제2 기계학습모델을 통하여 질환과 매칭할 수 있다.

발명의 효과

[0021] 개시된 실시 예에 따르면, 사용자의 증상정보를 기반으로 질환을 진단하는 챗봇 서비스를 제공할 수 있다.

[0022] 또한, 개시된 실시 예에 따르면, 사용자의 증상정보를 자연어로 입력받고, 이를 표준화하여 인공지능기반으로 매칭되는 질환을 도출할 수 있다.

[0023] 본 발명에서 얻을 수 있는 효과는 이상에서 언급한 효과들로 제한되지 않으며, 언급하지 않은 또 다른 효과들은 아래의 기재로부터 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에게 명확하게 이해될 수 있을 것이다.

도면의 간단한 설명

- [0024] 본 발명의 특정한 바람직한 실시예들의 상기에서 설명한 바와 같은 또한 다른 측면들과, 특징들 및 이득들은 첨부 도면들과 함께 처리되는 하기의 설명으로부터 보다 명백하게 될 것이다.
- 도 1은 본 발명의 일 실시예에 따른 인공지능 기반의 증상 및 질환 매칭을 위한 챗봇 서비스 제공방법의 흐름도이다.
- 도 2는 기계학습모델을 학습하기 위한 기초데이터의 예시를 나타내는 도면이다.
- 도 3은 증상입력정보에 대한 예시를 나타내는 도면이다.
- 도 4는 데이터의 검토 및 정리를 나타내는 도면이다.
- 도 5는 본 발명의 일 실시예에 따라 증상과 질환을 매칭하는 기본 모델을 설명하기 위한 도면이다.
- 도 6은 본 발명의 일 실시예에 따라 증상과 질환을 매칭하는 최종 모델을 설명하기 위한 도면이다.
- 도 7은 본 발명의 일 실시예에 따른 챗봇의 동작을 나타내는 흐름도이다.
- 도 8은 본 발명의 일 실시예에 따른 챗봇의 실시예를 나타내는 도면이다.
- 도 9는 본 발명의 일 실시예에 따른 컴퓨터 시스템을 나타낸 도면이다.
- 상기 도면들을 통해, 유사 참조 번호들은 동일한 혹은 유사한 엘리먼트들과, 특징들 및 구조들을 도시하기 위해 사용된다는 것에 유의해야만 한다.

발명을 실시하기 위한 구체적인 내용

- [0025] 본 발명의 이점 및 특징, 그리고 그것들을 달성하는 방법은 첨부되는 도면과 함께 상세하게 후술되어 있는 실시예들을 참조하면 명확해질 것이다. 그러나 본 발명은 이하에서 개시되는 실시예들에 한정되는 것이 아니라 서로 다른 다양한 형태로 구현될 수 있으며, 단지 본 실시예들은 본 발명의 개시가 완전하도록 하고, 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에게 발명의 범주를 완전하게 알려주기 위해 제공되는 것이며, 본 발명은 청구항의 범주에 의해 정의될 뿐이다. 명세서 전체에 걸쳐 동일 도면부호는 동일 구성 요소를 지칭한다.
- [0026] 본 발명의 실시예들을 설명함에 있어서 공지 기능 또는 구성에 대한 구체적인 설명이 본 발명의 요지를 불필요하게 흐릴 수 있다고 판단되는 경우에는 그 상세한 설명을 생략할 것이다. 그리고 후술되는 용어들을 본 발명의 실시예에서의 기능을 고려하여 정의된 용어들로서 이는 사용자, 운용자의 의도 또는 관례 등에 따라 달라질 수 있다. 그러므로 그 정의는 본 명세서 전반에 걸친 내용을 토대로 내려져야 할 것이다.
- [0027] 본 발명의 이점 및 특징, 그리고 그것들을 달성하는 방법은 첨부되는 도면과 함께 상세하게 후술되어 있는 실시예들을 참조하면 명확해질 것이다. 그러나 본 발명은 이하에서 개시되는 실시예들에 한정되는 것이 아니라 서로 다른 다양한 형태로 구현될 수 있으며, 단지 본 실시예들은 본 발명의 개시가 완전하도록 하고, 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에게 발명의 범주를 완전하게 알려주기 위해 제공되는 것이며, 본 발명은 청구항의 범주에 의해 정의될 뿐이다. 명세서 전체에 걸쳐 동일 참조 부호는 동일 구성 요소를 지칭한다.
- [0028] 이때, 처리 흐름도 도면들의 각 블록과 흐름도 도면들의 조합들은 컴퓨터 프로그램 인스트럭션들에 의해 수행될 수 있음을 이해할 수 있을 것이다. 이들 컴퓨터 프로그램 인스트럭션들은 범용 컴퓨터, 특수용 컴퓨터 또는 기타 프로그램 가능한 데이터 프로세싱 장비의 프로세서에 탑재될 수 있으므로, 컴퓨터 또는 기타 프로그램 가능한 데이터 프로세싱 장비의 프로세서를 통해 수행되는 그 인스트럭션들이 흐름도 블록(들)에서 설명된 기능들을 수행하는 수단을 생성하게 된다. 이들 컴퓨터 프로그램 인스트럭션들은 특정 방식으로 기능을 구현하기 위해 컴퓨터 또는 기타 프로그램 가능한 데이터 프로세싱 장비를 지향할 수 있는 컴퓨터 이용 가능 또는 컴퓨터 판독 가능 메모리에 저장되는 것도 가능하므로, 그 컴퓨터 이용가능 또는 컴퓨터 판독 가능 메모리에 저장된 인스트럭션들은 흐름도 블록(들)에서 설명된 기능을 수행하는 인스트럭션 수단을 내포하는 제조 품목을 생산하는 것도 가능하다. 컴퓨터 프로그램 인스트럭션들은 컴퓨터 또는 기타 프로그램 가능한 데이터 프로세싱 장비 상에 탑재되는 것도 가능하므로, 컴퓨터 또는 기타 프로그램 가능한 데이터 프로세싱 장비 상에서 일련의 동작 단계들이 수행되어 컴퓨터로 실행되는 프로세스를 생성해서 컴퓨터 또는 기타 프로그램 가능한 데이터 프로세싱 장비를 수행하는 인스트럭션들은 흐름도 블록(들)에서 설명된 기능들을 실행하기 위한 단계들을 제공하는 것도 가능하다.

다.

- [0029] 또한, 각 블록은 특정된 논리적 기능(들)을 실행하기 위한 하나 이상의 실행 가능한 인스트럭션들을 포함하는 모듈, 세그먼트 또는 코드의 일부를 나타낼 수 있다. 또, 몇 가지 대체 실행 예들에서는 블록들에서 언급된 기능들이 순서를 벗어나서 발생하는 것도 가능함을 주목해야 한다. 예컨대, 잇달아 도시되어 있는 두 개의 블록들은 사실 실질적으로 동시에 수행되는 것도 가능하고 또는 그 블록들이 때때로 해당하는 기능에 따라 역순으로 수행되는 것도 가능하다.
- [0030] 이 때, 본 실시 예에서 사용되는 '~부'라는 용어는 소프트웨어 또는 FPGA(field-Programmable Gate Array) 또는 ASIC(Application Specific Integrated Circuit)과 같은 하드웨어 구성요소를 의미하며, '~부'는 어떤 역할들을 수행한다. 그렇지만 '~부'는 소프트웨어 또는 하드웨어에 한정되는 의미는 아니다. '~부'는 어드레싱할 수 있는 저장 매체에 있도록 구성될 수도 있고 하나 또는 그 이상의 프로세서들을 재생시키도록 구성될 수도 있다. 따라서, 일 예로서 '~부'는 소프트웨어 구성요소들, 객체지향 소프트웨어 구성요소들, 클래스 구성요소들 및 태스크 구성요소들과 같은 구성요소들과, 프로세스들, 함수들, 속성들, 프로시저들, 서브루틴들, 프로그램 코드의 세그먼트들, 드라이버들, 펌웨어, 마이크로코드, 회로, 데이터, 데이터베이스, 데이터 구조들, 테이블들, 어레이들, 및 변수들을 포함한다. 구성요소들과 '~부'들 안에서 제공되는 기능은 더 작은 수의 구성요소들 및 '~부'들로 결합되거나 추가적인 구성요소들과 '~부'들로 더 분리될 수 있다. 뿐만 아니라, 구성요소들 및 '~부'들은 디바이스 또는 보안 멀티미디어카드 내의 하나 또는 그 이상의 CPU들을 재생시키도록 구현될 수도 있다.
- [0031] 본 발명의 실시예들을 구체적으로 설명함에 있어서, 특정 시스템의 예를 주된 대상으로 할 것이지만, 본 명세서에서 청구하고자 하는 주요한 요지는 유사한 기술적 배경을 가지는 여타의 통신 시스템 및 서비스에도 본 명세서에 개시된 범위를 크게 벗어나지 아니하는 범위에서 적용 가능하며, 이는 당해 기술분야에서 숙련된 기술적 지식을 가진 자의 판단으로 가능할 것이다.
- [0033] 이하에서, 첨부된 도면을 참조하여 본 발명의 일 실시예에 따른 인공지능 기반의 증상 및 질환 매칭을 위한 챗봇 서비스 제공방법에 대해 상세하게 설명한다.
- [0034] 도 1은 본 발명의 일 실시예에 따른 인공지능 기반의 증상 및 질환 매칭을 위한 챗봇 서비스 제공방법의 흐름도이다.
- [0035] 도 1을 참조하면, 본 발명의 일 실시예에 따라 장치에 의하여 수행되는 인공지능 기반의 증상 및 질환 매칭을 위한 챗봇 서비스 제공방법은, 사용자로부터 자연어 기반의 증상입력정보를 입력받을 수 있다(S101).
- [0036] 이 때, 상기 증상입력정보는 사용자의 성별, 나이, 키, 몸무게, 병력 등에 관한 정보 및 현재 증상에 관한 정보를 포함할 수 있다. 현재 증상에 관한 정보는 자연어를 기반으로 입력받을 수 있다. 예를 들어, '지금 속이 매우 더부룩해요' 또는 '설사가 3시간째 나오고 있어요' 등으로 입력받을 수 있다.
- [0037] 또한, 본 발명의 일 실시예에 따른 인공지능 기반의 증상 및 질환 매칭을 위한 챗봇 서비스 제공방법은 상기 증상입력정보를 기반으로 증상키워드를 추출할 수 있다(S103).
- [0038] 예를 들면, '제가 오늘 아까부터 배가 아파서요 밥을 못먹었어요' 를 입력받은 경우, 질환과 관련된 키워드만을 '오늘, 배가, 아프다 등' 으로 추출할 수 있다.
- [0039] 또한, 본 발명의 일 실시예에 따른 인공지능 기반의 증상 및 질환 매칭을 위한 챗봇 서비스 제공방법은 상기 증상키워드를 토큰화할 수 있다(S105).
- [0040] 이 때, 상기 S105 단계는 Soynlp, koNLpy 또는 토큰나이저 중 적어도 하나 이상을 이용하여 토큰화할 수 있다. soynlp는 한국어 처리를 위한 파이썬 패키지 중 하나다. koNLpy에서 제공하는 형태소분석기는 형태소 기반으로 문서를 토큰화할 수 있는 기능을 제공하지만 새롭게 만들어진 미등록 단어들은 인식이 잘 되지 않는 단점을 가지고 있다. 이를 해결하기 위해서는 사용자 사전에 단어를 등록하는 절차를 거쳐야 한다. soynlp는 이러한 과정을 돕기 위해 사용자 사전과 형태소분석 없이 cohesion 기반으로 토큰화를 할 수 있는 기능을 제공한다. koNLpy는 파이썬으로 한국어 정보처리를 할 수 있게 하는 패키지이다. 꼬꼬마, 한나눔, MeCab-ko 등 국내외에서 개발된 여러 형태소 분석기를 포함하고, 자연어처리에 필요한 각종 사전, 말뭉치, 도구 및 다양한 튜토리얼을 포함하여 누구나 손쉽게 한국어 분석을 할 수 있도록 만들었다. 토큰나이저 또한, Soynlp와 koNLpy와 마찬가지로 자연어에서 단어 또는 서브워드 단위로 쪼개어 사전에 등록된 표준단어 등으로 변환해주는 라이브러리를 의미한다.
- [0041] 또한, 본 발명의 일 실시예에 따른 인공지능 기반의 증상 및 질환 매칭을 위한 챗봇 서비스 제공방법은 토큰화

된 상기 증상키워드를 표준화하여 증상표준키워드를 추출할 수 있다(S107).

- [0042] 예를 들면, '배가 아픈데/아파서/아파다/완전아프네/아프네 등' 으로 추출된 증상키워드를 '아프다' 라는 표준화된 키워드로 변환하여 증상표준키워드로 추출할 수 있다.
- [0043] 또한, 본 발명의 일 실시예에 따른 인공지능 기반의 증상 및 질환 매칭을 위한 챗봇 서비스 제공방법은 상기 증상표준키워드를 기반으로 질환과 매칭할 수 있다(S109).
- [0044] 또한, 본 발명의 일 실시예에 따른 인공지능 기반의 증상 및 질환 매칭을 위한 챗봇 서비스 제공방법은 상기 증상표준키워드 중에서 질환 매칭을 위한 필수항목의 존재여부를 확인할 수 있다.
- [0045] 사용자가 자신의 상태에 대하여 자연어로 입력하다보면 질환과 매칭하기에 필요한 최소한의 정보를 놓치는 경우가 생길 수 있다. 이 때, 입력된 정보만을 가지고 질환을 매칭하는 것은 매칭정확도를 저하시키고, 불확실한 진단으로 인해 약물 오용 등의 문제가 발생할 수 있어, 질환 매칭을 위한 필수항목의 존재여부를 확인하는 것이 바람직하다.
- [0046] 따라서, 필수항목의 존재여부를 확인해본 결과 상기 증상표준키워드 중에서 상기 필수항목이 존재하지 않은 경우, 기설정된 추가문진메뉴얼을 상기 사용자에게 제공할 수 있다.
- [0047] 이는 사용자에게 질환 매칭을 위해서 필요한 정보가 무엇인지 안내를 함과 동시에, 정확한 매칭을 위하여 필요한 정보를 보다 구체적으로 사용자에게 입력받기 위함이다. 이는 챗봇을 기반으로 입력받을 수 있다.
- [0048] 이 때, 상기 추가문진메뉴얼은 'OLD CoEx CAFE 약사과의과' 와 환자의 기본 정보(Sex/Age/Height/Weight), 환자의 주요 호소 증상(Chief complaint)을 포함할 수 있다. 이와 관련해서는 후술하도록 한다.
- [0049] 또한, 본 발명의 일 실시예에 따른 인공지능 기반의 증상 및 질환 매칭을 위한 챗봇 서비스 제공방법은 질환 매칭을 위하여 제1 기계학습모델을 통하여 상기 증상표준키워드마다 상기 증상표준키워드에 대응하는 가중치를 출력할 수 있다.
- [0050] 이는 증상표준키워드마다 질환과의 상관관계가 높은 키워드가 있고, 상관관계가 낮은 키워드가 있으므로 이에 대한 차등을 주기 위함이다. 예를 들어, '오늘', '복통', '3시간' 이란 증상표준키워드가 추출되었다면, '복통' 이란 키워드가 질환 매칭에 가장 적합한 키워드일 수 있다. 따라서, '복통' 이란 키워드가 가중치가 높도록 상기 제1 기계학습모델을 학습시킬 수 있다.
- [0051] 또한, 본 발명의 일 실시예에 따른 인공지능 기반의 증상 및 질환 매칭을 위한 챗봇 서비스 제공방법은 상기 매칭하는 단계에서 상기 증상표준키워드 및 상기 가중치를 입력으로 매칭되는 질환 및 질환가능성을 출력하는 제2 기계학습모델을 통하여 질환과 매칭할 수 있다.
- [0052] 이 때, 본 발명의 일 실시예에 따른 인공지능 기반의 증상 및 질환 매칭을 위한 챗봇 서비스 제공방법을 실행하는 장치는 상기 제1 기계학습모델 및 상기 제2 기계학습모델을 생성하고 학습하는 인공지능모듈을 포함할 수 있다.
- [0053] 이 때, 인공지능모듈은 딥러닝을 통하여 상기 함수에서의 복수 개의 입력들의 가중치(weight)를 학습을 통하여 산출할 수 있다. 또한, 이러한 학습을 위하여 활용되는 인공지능망 모델로는 RNN(Recurrent Neural Network), DNN(Deep Neural Network) 및 DRNN(Dynamic Recurrent Neural Network) 등 다양한 모델들을 활용할 수 있을 것이다.
- [0054] 여기서 RNN은 현재의 데이터와 과거의 데이터를 동시에 고려하는 딥러닝 기법으로서, 순환 신경망(RNN)은 인공 신경망을 구성하는 유닛 사이의 연결이 방향성 사이클(directed cycle)을 구성하는 신경망을 나타낸다. 나아가, 순환 신경망(RNN)을 구성할 수 있는 구조에는 다양한 방식이 사용될 수 있는데, 예컨대, 완전순환망(Fully Recurrent Network), 홉필드망(Hopfield Network), 엘만망(Elman Network), ESN(Echo state network), LSTM(Long short term memory network), 양방향(Bi-directional) RNN, CTRNN(Continuous-time RNN), 계층적 RNN, 2차 RNN 등이 대표적인 예이다. 또한, 순환 신경망(RNN)을 학습시키기 위한 방법으로서, 경사 하강법, Hessian Free Optimization, Global Optimization Method 등의 방식이 사용될 수 있다.
- [0055] 또한, 본 발명의 일 실시예에 따른 인공지능 기반의 증상 및 질환 매칭을 위한 챗봇 서비스 제공방법은 매칭된 질환정보를 사용자에게 제공할 수 있다(S111).
- [0057] 이하에서는 본 발명의 일 실시예에 따른 인공지능 기반의 증상 및 질환 매칭을 위한 챗봇 서비스의 구현을 위한

여 기계학습모델을 학습시키기 위한 기초데이터를 추출하여 모델을 생성하는 과정을 설명하도록 한다.

- [0059] 도 2는 기계학습모델을 학습하기 위한 기초데이터의 예시를 나타내는 도면이다. 도 2를 참조하면, 인공지능모델을 통해 기계학습모델을 생성하기 위한 의료데이터를 시중에 존재하는 ‘guide to REAL CPX’, ‘한권으로 끝내는 CPX 개정 4판’ 등 총 5권의 서적에 적힌 질병예시, 증례들을 바탕으로 수집할 수 있다.
- [0060] 다만, 시중에 존재하는 데이터는 각 질병 당 케이스 2~3개뿐으로 기계학습모델을 학습하기에 부족하였다. 이에 실제 문진데이터를 통해 기초데이터를 보충할 수 있었다.
- [0062] 도 3은 증상입력정보에 대한 예시를 나타내는 도면이다.
- [0063] 도 3을 참조하면, 문진 데이터의 각 컬럼은 ‘OLD CoEx CAFE 약사가외과’와 환자의 기본 정보(Sex/Age/Height/Weight), 환자의 주요 호소 증상(Chief complaint)으로 구성될 수 있다.
- [0064] 또한, 데이터 중 ‘level’ 컬럼은 질환의 분류 과정을 의미한다. ‘level1’은 계통(기관계), ‘level2’는 주요 증상, ‘level3’는 기관, ‘level4’는 조직, 마지막으로 ‘level5/diagnosis’는 질환명에 해당할 수 있다. ‘level2’, ‘level5/diagnosis’와 다른 level 컬럼들은 분류를 위한 참고 용도로써, 예측에 직접적으로 사용되지는 않을 수 있다. 질환 매칭 모델 구축에는 전처리 과정을 거친 약 220가지의 질환이 사용되었는데, 이 질환들은 respiratory, circulator 등 6가지 계통과 acute abdominal pain, dyspepsia 등 27가지의 주요 증상으로 분류할 수 있다.
- [0066] 도 4는 데이터의 검토 및 정리를 나타내는 도면이다.
- [0067] 도 4를 참조하면, 기계학습모델을 학습시키기 위한 데이터의 포맷은 서로 상이하여 이를 전처리하는 과정이 필요하다. 데이터의 전처리 과정으로써 아래와 같은 과정을 수행할 수 있다.
- [0069] - 데이터 검토 및 정리
- [0070] (1) 오타자 수정
- [0071] 대부분의 데이터값이 문자열(텍스트) 자료형인데다가 입력이 수기로 이루어졌기 때문에, 틀린 맞춤법이나 잘못된 기입된 데이터값을 수정하는 작업을 수행할 수 있다.
- [0072] (2) 진단 결과 값 제외
- [0073] raw 데이터 값들 중에는 신체 검사 등 의사가 환자를 직접 진찰하여 획득한 정보들도 존재한다. 본 발명의 일 실시예에 따른 기계학습모델은 챗봇을 기반으로 환자에게 묻는 형태로 정보를 얻기 때문에 진찰을 통한 정보는 필요하지 않았고, 이에 따라 해당 데이터 값을 제외하는 작업을 거칠 수 있다.
- [0074] e.g. 겨드랑이 림프절 촉진, 좌상복부 지라 촉진
- [0075] (3) 전문용어 수정
- [0076] 데이터가 실제 문진 과정에서 수집되었기 때문에, 일상생활에서 흔히 쓰이지 않는 의학 전문용어들을 상당수 있다. 사용자들이 텍스트를 입력할 때 이와 같은 단어들을 사용하지 않으므로, 이를 일반 사용자들이 보편적으로 떠올릴 수 있는 단어로 바꾸는 작업을 수행할 수 있다.
- [0077] e.g. 회발월경 → 월경/생리 주기가 길다, 소양증 → 가려움증
- [0078] (4) 영어-한국어 번역
- [0079] (2)번과 같은 맥락으로 영어로 기입된 용어들을 한국어로 번역하는 작업을 거쳤다.
- [0080] e.g. NSAIDs → 소염진통제, DM → 당뇨
- [0081] 검토한 사항들은 raw 데이터의 직접 수정 없이 별도의 엑셀 파일(용어 정리.xlsx)을 생성해 도 4와 같이 정리할 수 있다.
- [0083] - 질환 관련 작업
- [0084] 타겟값인 질환이 예측에 더 효과적으로 사용될 수 있도록 다음과 같은 작업을 수행할 수 있다.
- [0085] (1) 질환 번역
- [0086] raw 데이터의 질환명은 영어로 표기되어 있을 수 있다. 예측 질환명이 영어로 나타난다면 직관적인 이해가 불가

한 경우가 많을 것이고 사용자가 질환을 직접 검색해 볼 때에도 불편함이 있을 수 있다고 판단하여 이를 한국어로 번역하여 통일 시킬 수 있다.

[0087] e.g. Hyperthyroidism → 갑상선기능항진증

[0088] (2) 질환 통합

[0089] 각 질환당 케이스의 수는 일부 질환을 제외하고 다섯 개로, 이는 예측에 충분하지 않은 수의 데이터이다. 또한 동일한 질환이지만 다르게 표기되거나 매우 세분화된 질환들이 존재하여 오히려 사용자에게 혼동을 줄 가능성도 있었다. 따라서 둘 이상의 질환들을 동일하게 표기, 하나의 상위 질환으로 합하는 작업을 수행할 수 있다.

[0090] e.g. 불안정 협심증(unstable angina) + 안정 협심증(stable angina) → 협심증(angina)

[0091] 위식도 역류 질환(Gastroesophageal reflux disease), 역류식도염(Reflux esophagitis) → 역류성 식도염

[0092] 결과적으로 한 질환에 대해 보다 다양하고 많은 데이터의 획득과 함께, 이 과정이 예측 정확도의 향상 및 사용자 친화적인 진단 수행에 기여할 수 있다.

[0094] - 전처리 진행 과정

[0095] 위 작업들을 바탕으로 파이썬을 이용하여 수행한 전처리 작업들은 다음과 같은 순서로 진행할 수 있다.

[0096] (1) 결측값 통일 및 과반수 이상 데이터 제거

[0097] 데이터에는 '해당 사항 없음', '(-)', '-' 등과 같은 결측값이 있다. 서로 다른 결측값 표현들을 빈칸(' ')으로 통일하였다. 또한 데이터의 질을 높이기 위해 결측값이 과반수 이상인 행들은 제거하였다.

[0098] (2) 질병명 번역 및 통합

[0099] 위에서 언급한 질환 관련 작업으로 질병명 번역 및 통합을 진행할 수 있다.

[0100] (3) 체중 상태, 나이 그룹화, 성별 표현 정리

[0101] 체중 상태: 환자들의 키와 몸무게 데이터를 이용하여 체질량 지수(BMI)를 계산한 후 비만도를 정의하였다. 통영시 보건소의 자가진단 BMI에 의한 비만도 계산법을 참고하였으며, 아래 표 1과 같이 비만도를 정의하였다.

표 1

계산식	신체질량지수(BMI) = $\frac{\text{체중}(kg)}{[\text{신장}(m)]^2}$	
판정기준	저체중	20 미만
	정상	20 - 24
	과체중	25 - 29
	비만	30 이상

[0103] 나이 그룹화: 나이는 0s, 10s, 20s, 30s, ..., 90s로 그룹화를 진행하였다. 그 이외의 나이표현들 (e.g 생후 ~주, ~개월 등)은 정규표현식을 이용하여 생후 몇 주 인지 계산한 뒤 '신생아', '영아', '유아' 로 그룹화할 수 있다.

[0104] 성별 표현 정리: 데이터의 기관계 마다 성별 표현방법이 다르기에 (e.g. F, M, 여성, 남성, 남자, 여자) 모두 '남자' 혹은 '여자' 로 통일할 수 있다.

[0105] (4) Negative value 제거

[0106] 데이터에는 환자의 발현된 증상인 positive value뿐만 아니라 발현되지 않은 negative value도 포함되어 있다. 유저가 발현되는 증상 위주로 챗봇에 적는다는 가정아래 negative value는 제거하고 positive value만 남기도록 remove_negative_values 함수를 만들어 적용할 수 있다. 정규표현식을 이용하여 괄호 안 텍스트들을 제거한 뒤, negative value인 '-' 에 해당하는 값들은 제거할 수 있다.

[0107] 함수 적용 전: '[진진] 발열 / 식은땀 / 체중 감소 / 오한/ 피로감(-/-/-/-), [소화기] 구역 / 구토 / 식욕 부진 / 변비 / 설사(-/#, 2회 흰 물만 나옴 /-/-/-), 속쓰림, 혈변(-), [응급] 어지러움 / 갈증 / 소변량 감소 / 호흡곤란(##/##/-)'

- [0108] 함수 적용 후: ', 구토, 속쓰림, [응급] 어지러움 같은 소변량 감소 '
- [0109] 추가적으로, 데이터가 대부분은 동일한 형식으로 작성되었지만 일부는 조금 다른 형식으로 작성되어있어 remove_negative_values 함수가 제대로 적용되지 않는 경우가 종종 있었다. 이런 경우 수기로 데이터를 수정하는 과정을 거쳐 최대한 오류없이 positive value를 추출할 수 있다.
- [0110] (5) NRS 점수 텍스트화
- [0111] NRS (Numerical Rating Scale)은 통증 평가 척도 중 하나인 수치평가척도이다. 현재의 주관적 통증 정도를 0~10의 숫자로 객관화해 말하라고 환자에게 물었을 때 그 값들의 분포를 나타낸 것이다. 통증정도를 숫자에서 텍스트로 변환하기 위해 NRS 7~10은 '심함', NRS 4~6은 '중간', NRS 3이하는 '약함'으로 대체하였다.
- [0112] (6) 한 문장으로 만들기
- [0113] 데이터에서 질병예측에 사용할 수 있는 열들을 모두 연결해 한 문장으로 표현하는 'All'이라는 새로운 열을 만들 수 있다.
- [0114] 사용된 열들: Chief complaint, Age, Sex, Onset, Location, Duration, Course, Experience, Character, Associated Sx., Factor, Event, 약물 투약력, 사회력, 가족력, 외상력, 과거력, 여성력, Obesity
- [0115] (7) 맞춤법 교정
- [0116] 데이터 검토 중 오타자 수정에서 작업한 파일을 반영한 수기 교정을 진행하였다. 그 후, 추가 맞춤법 교정을 위해 네이버 맞춤법 검사기 등 상용검사기를 이용한 파이썬용 한글 맞춤법 검사 라이브러리인 py-hanspell을 적용하였다.
- [0117] (8) 유저의 입력을 고려한 용어 정리
- [0118] 일반인 유저들을 고려한 용어 정리는 위에서 언급한 데이터 검토 중 전문 용어 수정 및 영어-한국어 번역 파일을 반영한 replace_diffi_to_easy 함수와 replace_eng_to_kor 함수를 이용하여 정리할 수 있다.
- [0119] (9) 형태소 분석
- [0120] 한국어 정보처리를 위한 파이썬 패키지인 KoNLPy를 이용하여 형태소 분석을 통한 동사 및 형용사의 어간 추출을 진행할 수 있다. (erase_stopwords 함수에 포함)
- [0121] 적용 전: '머리가 아프고 어지러워요.'
- [0122] 적용 후: '머리' '가' '아프다' '어지럽다' ' ' '.'
- [0123] (10) 불용어, 불필요한 기호 제거
- [0124] 내용에서 '의', '가', '은', '는' 등 큰 의미가 없는 불용어 목록을 정의한 후 불용어를 제거하였다. 또한 (), !, @, # 등의 불필요한 기호들도 있어서 정규표현식을 활용하여 이를 제거할 수 있다.
- [0126] 또한, 사용자로부터 증상입력정보를 입력받기 위해서는 챗봇을 통해 사용자가 입력할 정보를 안내하는 것이 중요하다. 따라서, 사전에 문진항목을 생성하여 이를 챗봇 형식으로 사용자에게 먼저 제공할 수 있다.
- [0127] 의사들의 문진과정은 주로 'OLD CoEx CAFE 약사가외과'라는 매뉴얼을 따른다. 이 매뉴얼의 의미는 다음과 같다.
- [0128] O) Onset ? 증상이 발생한 시점
- [0129] L) Location ? 증상의 위치
- [0130] D) Duration - 증상의 지속 시간(1일 당 반복 횟수)
- [0131] Co) Course ? 증상의 정도(점점 심해짐, 유지 등)
- [0132] Ex) Experience ? 비슷한 증상의 경험 여부
- [0133] C) Character ? 증상의 양상 및 특징(통증의 양상, 색, 질감 등)
- [0134] A) Associated Sx. - 동반 증상

- [0135] F) Factor ? 유발요인(악화요인, 완화요인)
- [0136] E) Exam ? 건강검진 이력
- [0137] 약) 약물복용력 ? 복용하고 있거나 복용했던 약 이력
- [0138] 사) 사회력 ? 술/담배/커피 유무, 직업 등
- [0139] 가) 가족력 ? 주요증상에 관련하여 가족이 앓은 병
- [0140] 외) 외과(수술력, 외상력) ? 다쳐서 외상을 입거나 수술을 한 이력
- [0141] 과) 과거력 ? 과거 앓았던 질병 이력
- [0142] 추가적으로 사용자로부터 정확한 정보를 원활히 입력 받기 위해 입력 예시를 제시하고 답변은 되도록이면 선택지에서 고를 수 있게 할 수 있다. 각 주요증상 별로 나타나는 질병들의 관련증상들과 악화/완화요인들을 조사하고 데이터에 기입된 증상들을 합하여 질문목록의 선택지로 반영할 수 있다.
- [0144] - 문진 항목 목록
- [0145] 아래의 목록은 위의 과정을 거쳐 만들어진 공통질문과 각 주요증상 27개에 맞는 질문목록 및 선택지를 선정한 것이다. 또한 아래 목록에서는 생략하였지만, 모든 주요증상 질문의 마지막에는 ‘다른 특이사항 여부’를 입력받음으로써 사용자가 추가로 언급하고 싶은 부분을 보완할 수 있다.
- [0147] (0) 공통 질문
- [0148] * C.C: 주된 증상이 무엇인가요?
- [0149] * 공통사항: 기본정보를 입력해주세요.
- [0150] - 성별/나이(연령대)
- [0151] - 키/몸무게 → BMI 측정
- [0152] * Onset: 언제부터 증상이 시작되었나요? (1일 전/2일 전/1달 전 등)
- [0153] * Location: 어느 부위가 불편한가요? (전체/위/아래/옆구리 등)
- [0154] (1) 급성 복통
- [0155] Duration: 증상이 한번 생기면 얼마나 나타나나요?(하루 종일/1시간 지속됨/간헐적/가끔씩)
- [0156] Character: 증상의 양상이 어떻게 되나요?
- [0157] - 통증의 퍼짐/속이 쓰림/더부룩함/가슴이 답답한 느낌
- [0158] Factor: 어떤 경우에 증상 이 더 심해지거나 완화되나요(복수선택 가능)? 해당 사항이 있으면 써주세요.
- [0159] - 자세에 따른 증상 변화(돌아누우면 완화, 반듯이 누우면 악화 등)
- [0160] - 식사 전후 증상 변화(식사 후 악화 등)
- [0161] - 음주 후 증상 변화
- [0162] - 운동 후 증상 변화
- [0163] - 해당사항 없음
- [0164] Associated Sx.: 아래와 같은 증상들 중 나타나는 증상을 모두 입력해주세요. 이외에 동반된 증상이 있다면 자유롭게 입력해주세요.
- [0165] - 식욕부진/ 구토/ 메스꺼움/ 설사/ 변비/ 적색변/ 흑색변/ 가스배출감소/ 체중감소/ 혈뇨/ 가슴 작열감/ 스트레스/ 소변감소/ 토혈/ 복부단단함
- [0166] Patient history: 해당되는 내용이 있으면 써 주시고 없으면 그대로 비워주세요.
- [0167] - 약 복용력: 복용하고 있는 약물이 있다면 입력해주세요. (진통소염제)

- [0168] - 사회력: 음주/흡연을 얼마나 하시나요?(음주 ~회/주, 흡연 ~갑/일 ~년)
- [0169] - 수술력: 복부 수술 경력이 있으면 입력해주세요
- [0170] - 과거력: 앓고 있거나 앓았던 질병이 있으면 작성해주세요
- [0171] - 여성력: 임신 가능성이 있으면 체크해주세요(임신가능성)
- [0172]
- [0173] (2) 소화불량/만성 복통
- [0174] Duration: 증상이 한번 생기면 얼마나 나타나나요?(하루 종일/1시간 지속됨/간헐적/가끔씩)
- [0175] Character: 증상의 양상이 어떻게 되나요?
- [0176] - 통증의 퍼짐/속이 쓰림/더부룩함/가슴이 답답한 느낌
- [0177] Factor: 어떤 경우에 증상 이 더 심해지거나 완화되나요(복수선택 가능)? 해당 사항이 있으면 써주세요.
- [0178] - 자세에 따른 증상 변화(돌아누우면 완화, 반듯이 누우면 악화 등), 식사 전후 증상 변화(식사 후 악화 등), 음주 후 증상 변화, 운동 후 증상 변화, 해당사항 없음
- [0179] Associated Sx.: 아래와 같은 증상들 중 나타나는 증상을 모두 입력해주세요. 이외에 동반된 증상이 있다면 자유롭게 입력해주세요.
- [0180] - 속쓰림/ 설사/ 변비/ 적색변/ 흑색변/ 체중감소/ 가슴 작열감/ 스트레스/ 신물올라옴 등
- [0181] Patient history: 해당되는 내용이 있으면 써 주시고 없으면 그대로 비워주세요.
- [0182] - 약 복용력: 증상을 완화하기 위해 복용한 약물이 있으면 입력해주세요 (진통소염제 등)
- [0183] - 사회력: 음주/흡연을 얼마나 하시나요?(음주 ~회/주, 흡연 ~갑/일 ~년)
- [0184] - 가족력: 가족 중에 복부질환을 앓으셨던 분이 계신가요?(아버지, 대장암)
- [0185] - 외상력(수술력): 복부 수술 경력이 있으면 입력해주세요
- [0186] - 과거력: 앓고 있거나 앓았던 질병이 있으면 입력해주세요
- [0187]
- [0188] (3) 토혈
- [0189] Character: 증상의 양상이 어떻게 되나요? 다음의 사항을 입력 및 선택해주세요.
- [0190] - 피의 색(선홍색, 검붉거나 새까만 색) / 토의 양(소주잔, 종이컵 ~컵)
- [0191] - 기침 도중에 피가 남 / 코나 입에서 피가 남 / 토에 음식물 동반함 / 덩어리진 피 등
- [0192] Associated Sx.: 아래와 같은 증상들 중 나타나는 증상을 모두 입력해주세요. 이외에 동반된 증상이 있다면 자유롭게 입력해주세요.
- [0193] - 어지러움 / 숨이 참 / 배꼽 주변 통증 / 명치 부위 통증 / 속쓰림 / 혈변(붉은 색) / 가슴 통증 / 기침 / 피가 나면 잘 멎지 않음 / 코피가 자주 남 / 명이 잘 들 / 체중변화 / 소화불량 / 구역감 / 황달 / 발열 / 오한 / 신티름, 신물 / 변비 / 복통 / 구토 / 음식 삼키기 힘들 / 만성피로 / 식욕떨어짐 / 복부에 물 찬 느낌 / 검은색 변/복부단단함 / 체중증가/ 소화불량
- [0194] Event/과거력: 다음 중 해당사항이 있다면 모두 입력해주세요. 이외에 투병 중이거나 진단받았던 질환이 있다면 자유롭게 입력해주세요.
- [0195] - 위, 식도 내시경 검사 결과 / 과거 위염, 위궤양, 십이지장궤양, 식도염, 정맥류 진단 / 간 질환(바이러스성 만성 간염, 간경화, 간암 등) / 과거 복부 수술
- [0196] Patient history: 해당되는 내용이 있으면 써 주시고 없으면 그대로 비워주세요.
- [0197] - 약물 복용력: 다음 중 복용중이거나 오랜 시간동안 드신 약이 있으면 입력해주세요(진통소염제(이부프로펜,

아스피린 등) / 항응고제 / 항혈소판제)

- 사회력: 음주/흡연을 얼마나 하시나요? (음주 ~회/주, 흡연 ~갑/일 ~년)

(4) 피부발진

Character: 증상의 양상이 어떻게 되나요? 해당되는 사항을 선택해주세요.

- 발진 부위가 커지거나 퍼짐 / 증상이 나타났다가 사라짐 / 발진 부위에 가려움이나 통증

Event/Factor/과거력: 다음 중 해당되는 사항이 있다면 모두 입력해주세요.

- 새 옷, 장신구 착용 / 새로운 비누, 세제, 로션, 화장품 이용 / 음주 / 최근 처음 먹어본 음식 / 알레르기 / 심한 운동 / 뜨거운 목욕탕, 사우나 이용 / 벌레 물림 / 과거 알레르기 비염, 아토피 피부염 투병 / 햇빛 노출에 민감(노출되면 가렵거나 붉게 변함) / 굶으면 더욱 가려움 / 밤에 가려움 / 피곤하면 심해짐 / 외부 활동 / 단체생활 / 성 접촉

Associated Sx. 아래와 같은 증상들 중 나타나는 증상을 모두 입력해주세요.

이외에 동반된 증상이 있다면 자유롭게 입력해주세요.

- 발열 / 오한 / 관절 통증 / 피로 / 구강 궤양 / 성기 궤양

Patient history: 해당되는 내용이 있으면 써 주시고 없으면 그대로 비워주세요.

- 가족력: 혹시 가족 중에 비슷한 증상을 가진 분이 있나요?(아버지: 발열 등)

(5) 다뇨증

Duration: 소변을 하루에 몇 회 가량 보나요? 몇 시간 간격으로 소변을 보나요?

Character: 증상의 양상이 어떻게 되나요? 다음의 사항을 입력 및 선택해주세요.

- 한 번 소변을 볼 때의 양(종이컵 ~컵) / 소변의 색(붉은색, 콜라색 등) / 소변의 농도

- 1회 소변량 증가함 / 소변에서 냄새남

Associated Sx.: 아래와 같은 증상들 중 나타나는 증상을 모두 입력해주세요. 이외에 동반된 증상이 있다면 자유롭게 입력해주세요.

- 자주 목이 마름 / 물을 많이 마심 / 소변을 보기 위해 잠에서 깬 / 체중 감소 / 두통 / 욕지기(메스꺼움) / 시야장애(시야 축소, 물체가 두 개로 보임, 시야 흐림 등) / 전신의 쇠약 / 무력감 / 피로 / 피부 가려움 / 소변 볼 때 아랫배 통증 / 발열

Event/과거력: 다음 중 해당되는 사항이 있다면 모두 입력해주세요. 이외에 투병 중이거나 진단받았던 질환이 있다면 자유롭게 입력해주세요.

- 과거 고혈압, 당뇨, 만성콩팥질환, 뇌종양 진단 / 머리 다침 / 뇌수술

Patient history: 해당되는 내용이 있으면 써 주시고 없으면 그대로 비워주세요.

- 약물 복용력: 현재 복용중인 약, 한약, 건강보조식품, 다이어트 식품이 있으면 입력해주세요

- 가족력: 가족 중에 당뇨를 진단 받으신 분이 있나요?(아버지: 당뇨 등)

- (남성) 성욕 감퇴 또는 발기 부전 증상이 있으면 입력해주세요(발기 부전)

- (여성) 혹시 생리주기의 변화는 없나요? 규칙적인가요?

(6) 씹노

Duration: 소변을 하루에 몇 회 가량 보나요? 몇 시간 간격으로 소변을 보나요?

- [0228] Character: 증상의 양상이 어떻게 되나요? 다음의 사항을 입력 및 선택해주세요.
- [0229] - 한 번 소변을 볼때의 양(종이컵 ~컵) / 소변의 색(붉은색,콜라색 등) / 소변의 농도
- [0230] - 1회 소변량 감소함 / 소변에 거품 있음
- [0231] Associated Sx.: 아래와 같은 증상들 중 나타나는 증상을 모두 입력해주세요. 이외에 동반된 증상이 있다면 자유롭게 입력해주세요.
- [0232] - 체중 증가 / 설사 / 구토 / 메스꺼움 / 소변 볼 때 아랫배 통증 / 발열 / 오한 / 소변이 마려운데 배출이 힘들 / 심한 운동 / 근육통 / 옆구리 통증 / 숨이 참 / 흑색변(새까만 색) / 혈변(붉은 색) / 어지러움 / 피로 / 얼굴이나 다리 등 몸이 부음 / 두근거림 / 인후통 / 두통 / 눈통증 / 시야장애(시야 축소, 물체가 두 개로 보임, 시야 흐림 등) / 잔뇨감
- [0233] Event/과거력: 다음 중 해당되는 사항이 있다면 모두 입력해주세요. 이외에 투병 중이거나 진단받았던 질환이 있다면 자유롭게 입력해주세요.
- [0234] - 소변 이상 또는 신장 질환 / 고혈압 / 당뇨 / 고지혈증 / 심부전 / 신장 이식
- [0235] Patient history: 해당되는 내용이 있으면 써 주시고 없으면 그대로 비워주세요.
- [0236] - 약물 복용력: 복용 중인 약이 있으면 입력해주세요(소염제, 고혈압약, 항생제, 항암제 등)
- [0237]
- [0238] (7) 붉은색 소변
- [0239] Character: 증상의 양상이 어떻게 되나요? 다음의 사항을 입력 및 선택해주세요.
- [0240] - 소변의 색(선홍색, 김붉은색 등)
- [0241] - 혈뇨가 소변 보는 내내 나옴 / 혈뇨가 소변 보기 시작 후 초반에만 나옴 / 소변에 거품 있음 / 소변에 핏덩이 동반됨
- [0242] Associated Sx.: 아래와 같은 증상들 중 나타나는 증상을 모두 입력해주세요. 이외에 동반된 증상이 있다면 자유롭게 입력해주세요.
- [0243] - 소변 볼 때 아랫배 통증 / 잔뇨감 / 피로 / 발열 / 오한 / 얼굴이나 다리 등 몸이 부음 / 멍이 잘 듦 / 잇몸 피 / 옆구리 통증 / 체중 감소 / 체중 증가 / 인후통 / 소변 줄기 가늘어짐 / 호흡곤란, 호흡불편 / 구토 / 근육통
- [0244] Factor/사회력: 다음 중 해당되는 사항이 있다면 모두 입력해주세요.
- [0245] - 성관계 / 심한 운동 / 감기 / 사우나 / 다이어트
- [0246] Event/외상력: 옆구리나 배를 다친 적이 있나요?
- [0247] Patient history: 해당되는 내용이 있으면 써 주시고 없으면 그대로 비워주세요.
- [0248] - 약물 복용력: 최근에 새로 복용한 약이 있나요?
- [0249] (리팜핀(결핵약), 비타민, 아스피린, 와파린(항응고제), 진통제, 고혈압약 등)
- [0250] - 가족력: 가족 중에 혈뇨가 있거나 콩팥 질환을 가지고 있는 분이 있나요?
- [0251] (아버지: 요로결석 등)
- [0252] - 과거력: 다음 중 해당되는 사항이 있다면 모두 입력해주세요. 이외에 투병 중이거나 진단받았던 질환이 있다면 자유롭게 입력해주세요.
- [0253] (혈관염 / 루푸스신염 / 출혈성 질환 / 항암치료 / 고혈압 / 결핵)
- [0254] - (여성) 마지막 생리가 언제였나요? 지금 현재 생리 기간은 아닌가요?
- [0256] (8) 배뇨 이상
- [0257] Character: (소변에 관한 것들 다 넣음)

- [0258] - 하루에 소변을 보는 횟수가 어떻게 되나요? (하루 7번 등)
- [0259] - 해당되는 증상이 있으면 작성해주세요.
- [0260] 잔뇨감/혈뇨/거품뇨/탁한 소변/ 배뇨통/ 요실금/ 빈뇨/ 야뇨/ 절박뇨/ 요정체/ 소변 줄기 가늘어짐/소변 줄기 끊김/요도 분비물 등
- [0261] (빈뇨: 소변을 자주 봄, 야뇨: 밤 중에 자다가도 일어나서 소변을 봄, 절박뇨: 소변이 마렵기 시작하면 참기 힘들, 요 정체: 소변을 보려고 하면 소변이 잘 안 나와서 시간이 걸림)
- [0262] Associated Sx: 함께 동반된 증상이 있다면 작성해주세요.
- [0263] - 발열/오한/옆구리 통증/체중감소/시아장애/근육통 등
- [0264] Factor: 어떤 경우에 증상이 완화되거나 더 심해지나요?
- [0265] - 날씨에 따른 증상 변화, 자세에 따른 증상 변화, 기침, 웃음에 따른 증상 변화 등
- [0266] ex) 날씨가 추워지면 증상이 심해짐 등
- [0267] Patient history: 해당되는 내용이 있으면 써 주시고 없으면 그대로 비워주세요.
- [0268] - 복용력: 현재 복용중인 약이 있다면 적어주세요.
- [0269] - 과거력: 다음과 같은 질환을 진단받거나 치료받은 적이 있다면 적어주세요.
- [0270] 척수손상, 추간판탈출, 뇌졸중, 치매, 파킨슨병 등 신경계 질환/ 전립선 또는 요로 계통 질환/당뇨 등
- [0271] - 외상력: 다음과 같은 부위를 치료받은 적이 있다면 적어주세요. (회음부, 복부 등)
- [0272] - 사회력: 최근 성관계를 가진 적이 있다면 적어주세요.
- [0273] 위험한 성관계를 가졌다면 함께 적어주세요.
- [0274] (1주일 전 성관계/3일전 위험한 성관계 등)
- [0276] (9) 소변찢끔증
- [0277] Course: 점점 증상이 심해지나요?(심해짐 등)
- [0278] Character: (소변에 관한 것들 다 넣음)
- [0279] - 한번 증상이 나타날 때 소변이 새는 양을 알려주세요.
- [0280] (속옷을 약간 적시는 정도/속옷이 젖을 정도 등)
- [0281] - 하루에 몇번정도 증상이 나타나요? (하루 10번 정도/하루 팬티라이너 5-6장 교체 등)
- [0282] - 해당되는 증상을 모두 작성해주세요.
- [0283] 배뇨통/혈뇨/잔뇨감/거품뇨/빈뇨/야뇨/절박뇨/요 정체/소변이 갑자기 마려움/화장실에 가기도 전에 소변이 샘. 성관계 도중에 소변이 샘/대변이 함께 샘/일상 생활 지장 등
- [0284] (빈뇨: 소변을 자주 봄, 야뇨: 밤 중에 자다가도 일어나서 소변을 봄, 절박뇨: 소변이 마렵기 시작하면 참기 힘들, 요 정체: 소변을 보려고 하면 소변이 잘 안 나와서 시간이 걸림)
- [0285] Associated Sx: 함께 동반된 증상이 있다면 적어주세요.
- [0286] - 허리 통증/체중변화 등
- [0287] Factor: 어떤 경우에 증상이 완화되거나 더 심해지나요?
- [0288] - 운동에 따른 증상 변화, 자세에 따른 증상 변화, 무거운 것을 들때 증상 변화, 기침, 큰소리로 웃을 때 증상 변화 등 ex) 기침을 할 때마다 소변이 샘 등
- [0289] Patient history: 해당되는 내용이 있으면 써 주시고 없으면 그대로 비워주세요.
- [0290] - 복용력: 현재 복용중인 약이 있다면 적어주세요.

- [0291] - 과거력: 다음과 같은 질환을 진단받거나 치료받은 적이 있다면 적어주세요.
- [0292] (고혈압, 당뇨, 척추 디스크, 척수손상, 뇌졸중 등)
- [0293] - 외상력: 비뇨기과 또는 산부인과에서 수술이나 시술을 받은 적이 있다면 적어주세요.
- [0294] - 사회력: 술이나 커피(카페인 음료)는 얼마나 많이 드시나요?
- [0295] - 여성력: 분만 경험이 있다면 분만시 특이사항을 적어주세요
- [0296]
- [0297] (10) 관절 통증
- [0298] Duration: 한번 아프면 얼마나 오랫동안 지속되나요?
- [0299] Character: 증상의 양상이 어떻게 되나요?
- [0300] - NRS(통증의 정도): (0~10)
- [0301] - 육신거리는 느낌/관절 깊은 부위의 통증/관절 주변부 통증 등
- [0302] Associated Sx: 함께 동반된 증상이 있나요?
- [0303] - 관절 부위의 붓기/발열/염발음/불안정/관절이 양쪽 대칭적으로 아픔/입이나 눈이 마름/입 안이 자주 혈음/성기가 혈음/추운곳에 가면 손, 발가락이 하얗게 또는 새파랗게 변함/오한/홍조/발적/저리거나 감각이 둔함/열감/누르면 아픔/눈 충혈 등
- [0304] - 아침에 일어나면 관절이 뻣뻣한 증상, 지속 시간
- [0305] Factor: 어떻게 통증이 시작되었나요?(무리한 운동/몇주전 관절 주변 상처 등)
- [0306] 어떤 경우에 통증이 완화되거나 더 심해지나요?
- [0307] - 특정 시간에 따른 통증 변화, 자세에 따른 통증 변화, 운동 후 통증 변화 등
- [0308] ex) 오전에 심함, 햇빛 받으면 피부발진 생기고 관절통 심해짐 등
- [0309] Patient history: 해당되는 내용이 있으면 써 주시고 없으면 그대로 비워주세요.
- [0310] - 복용력: 현재 복용중인 약이 있다면 적어주세요.
- [0311] - 과거력: 다음과 같은 질환을 진단받거나 치료받은 적이 있다면 적어주세요.
- [0312] 고혈압, 통풍, 류마티스 질환 등
- [0313] 최근 심한 설사를 했거나 목 감기를 앓은 적이 있다면 체크해주세요(설사/목감기)
- [0314] - 가족력: 가족 중에 관절염을 진단 받거나 치료 받은 분이 있다면 적어주세요. (아버지, 관절염)
- [0315]
- [0316] (11) 목 통증/허리 통증
- [0317] Duration: 한번 아프면 얼마나 오랫동안 지속되나요?(하루종일/n시간 이상 지속 등)
- [0318] Character: 증상의 양상이 어떻게 되나요?
- [0319] - NRS(통증의 정도): (0~10)
- [0320] - 둔하고 빠근함/칼로 찌르듯이 날카롭게 아픔/찌릿찌릿함/통증의 퍼짐 등
- [0321] Associated Sx: 함께 동반된 증상이 있나요?
- [0322] - (공통)체중 감소/발열/오한/증상 부위의 힘이 약해짐/증상 부위에 전기가 통하는듯한 감각/증상 부위의 감각이 둔해짐/(목 통증)두통/(허리 통증)옆구리 통증/아침에 일어났을 때 통증과 빠근함/(허리 통증)대변을 보고난 후 항문의 감각이 둔해짐
- [0323] Factor: 어떻게 통증이 시작되었나요?(무리한 운동/ 무거운 물건을 들다가 등) 어떤 경우에 통증이 완화되거나

더 심해지나요?

- 운동 후 통증 변화, 자세에 따른 통증 변화, 기침에 따른 통증 변화 등

Patient history: 해당되는 내용이 있으면 써 주시고 없으면 그대로 비워주세요

- 복용력: 현재 복용중인 약이 있다면 적어주세요.

- 과거력: 다음과 같은 질환을 진단받거나 치료받은 적이 있다면 적어주세요.

- 고혈압, 당뇨, 고지혈증, 류마티스 질환, 골다공증, 근골격계 질환

(12) 혈변

Character:

- 색: 대변의 색깔이 무슨 색이었나요?(검붉은색, 선홍색 등)

- 양상: 피가 어느정도 나는지 체크해주세요

휴지에 묻어나오는 정도/대변과 같이 나옴/피가 변기 가득히 나옴

Associated sx.: 해당되는 증상이 있으면 체크해주세요.

- 변비/ 설사/ 체중감소/ 복통/ 항문통증/ 항문덩이/ 점액/ 복부팽만/ 어지러움/ 오한/ 발열/ 소화불량/ 식욕부진

Patient history: 해당되는 내용이 있으면 써 주시고 없으면 그대로 비워주세요.

- 가족력: 가족 중에 대장암이나 용종을 앓아서 치료받은 분이 있나요?

(아버지, 대장암 등)

- 외상력(수술력): 최근에 항문을 다친 적이 있나요? -> 데이터에는 없음

- 과거력: (내시경 포함)다른 질환을 진단 받거나 병원을 다니고 있는 다른 질환이 있나요? (1년전 내시경에서 대장 폴립 발견)

(13) 구토

Character:

- 색: 무슨 색이었나요?

- 양상: 해당되는 경우를 체크해주세요.

내용물이 저절로 역류함 / 음식물이 섞여 나옴 / 일부러 구토 유발 / (신생아/영아의 경우)토한 후에도 계속 먹으려 함 / (신생아의 경우)수유 후 구토

Associated Sx.: 해당되는 증상을 체크해주세요.

- 복통/ 설사/ 변비/ 속쓰림/ 위산역류/ 토혈/ 복통/ 열/ 오한/ 어지러움/ 체중감소/ 구역질/ 식욕부진/ 무월경/ 무기력함/ 팔다리감각저하/ 의식저하(소실)/ 두통/

Patient history: 해당되는 내용이 있으면 써 주시고 없으면 그대로 비워주세요.

- 약 복용력: 현재 복용중인 약이 있다면 적어주세요.

- 사회력: 잘못 드신 음식이 있거나, 함께 식사한 사람 중 증상을 호소하는 분이 있나요?

- 외상력(수술력): 머리 수술을 받은 적이 있으면 체크해주세요(두부 외상)

- 과거력: 다른 질환을 진단받거나 치료중인 질환이 있나요?

- 여성력: 임신 가능성이 있나요?

(14) 변비

- [0358] Character:
- [0359] - 빈도: 며칠마다 대변을 보시나요?(예시: 3일에 한 번 정도, 최근 3일간 변을 보지 못함)
- [0360] - 모양: 대변이 콩알모양이면 체크해주세요(콩알모양)
- [0361] Associated Sx.: 해당되는 증상을 체크해주세요.
- [0362] - 소화불량/복통/혈변/항문통증/설사/잔변감/체중증가/체중감소/추위를 많이 탐/스트레스/복부팽만
- [0363] Patient history: 해당되는 내용이 있으면 써 주시고 없으면 그대로 비워주세요.
- [0364] - 약복용력: 증상을 완화하기 위해 복용했거나 복용중인 약이 있으면 입력해주세요
- [0365] - 사회력: 식습관/운동습관/직업에 대해 간략히 서술해주세요
- [0366] - 가족력: 가족 중에 대장암이나 용종을 앓아서 치료받은 분이 있나요?
- [0367] - 외상력(수술력): 척추(요추, 경추)를 다친 적이 있으면 입력해주세요(척추 손상)
- [0368] - 과거력: 다른 질환을 진단받거나 병원을 다니고 있는 다른 질환이 있나요?
- [0369]
- [0370] (15) 설사
- [0371] Character:
- [0372] - 빈도: 하루에 몇 번 정도 증상이 나타나나요?
- [0373] - 양상: 증상이 어떻게 나타나는지 서술해주세요(물설사/점액 섞인설사 등)
- [0374] Factor:
- [0375] - 식후 증상이 악화되면 체크해주세요(우유 마신 후 설사/기름진 음식 먹은 후 설사 등)
- [0376] Associated sx: 해당되는 증상이 있으면 체크해주세요
- [0377] - 복통/열/오한/체중감소/갈증/스트레스/변비/복부팽만/갈증/적색변/흑색변
- [0378] Patient history: 해당되는 내용이 있으면 써 주시고 없으면 그대로 비워주세요.
- [0379] - 약복용력: 복용중인 약물이 있으면 입력해주세요.
- [0380] - 사회력: 최근 설사를 유발할 만한 음식을 드신 적이 있나요?
- [0381] - 가족력: 가족중에 복부질환을 앓았던 분이 있나요?(아버지: 대장암)
- [0382] - 외상력(수술력): 과거에 복부수술을 받으신 적이 있나요?
- [0383] - 과거력: 이전에 병을 진단받거나 치료(방사선치료 포함)를 받으신 적이 있나요?
- [0384]
- [0385] (16) 황달
- [0386] Character:
- [0387] - 색: 색깔이 어떤가요?(노란색/어두운 노란색 등)
- [0388] Associated Sx.: 해당되는 증상이 있으면 체크해주세요
- [0389] - 가려움증/ 발열/ 오한/ 체중변화/ 손떨림/ 회색변/ 콜라색/ 스트레스/ 소변/ 혈액응고지연/ 설사/ 호흡 곤란/ 인두통/ 혈뇨/ 피로감/ 구토/ 토혈/ 피로/ 복부팽만/소화불량/가슴에 붉은 반점/구역감/윗배 누르면 아픔/오른쪽 윗배 누르면 아픔/근육통/발음장애/보행장애/삼킴곤란/복통/어지러움/목 부근의 혹
- [0390] Factor:
- [0391] - 운동 후 통증 변화, 식후/식전 증상 변화 변화, 기침에 따른 통증 변화 등

- [0392] Patient history: 해당되는 내용이 있으면 써 주시고 없으면 그대로 비워주세요.
- [0393] - 약 복용력: (한약, 약초, 건강보조식품 포함) 평소 드시는 약이 있나요?
- [0394] - 사회력: 음주/흡연을 얼마나 하시나요? (음주 ~회/주, 흡연 ~갑/일 ~년)
- [0395] - 가족력: 가족 중에 간 질환을 앓고 계신 분이 있나요?
- [0396] - 외상력(수술력): 몸을 다치거나 수술을 받은 적이 있나요?
- [0397] - 과거력: 원래 간질환 혹은 빈혈을 앓았거나 멍이 자주 들었던 적이 있나요?
- [0399] (17) 기침
- [0400] Duration: 한번 증상이 나타나면 얼마나 오랫동안 지속되나요? (하루종일/ n시간 이상 지속/ 주기적으로 등)
- [0401] Character: 증상의 양상이 어떻게 되나요?
- [0402] - 마른 기침/깊은 곳에서 나오는 기침 등
- [0403] Associated Sx: 함께 동반된 증상이 있나요?
- [0404] - 감기/몸살기운/ 침을 삼킬 때 목이 아픔/가래 (색과 양)/ 객혈/ 발열/ 흉통/ 호흡곤란/ 쉼쉼거림/ 발열/ 오한 / 근육통/ 체중 감소/ 속쓰림/ 밤에 자다가 숨이 차서 깸/ 목 뒤로 콧물이 넘어감 등
- [0405] Factor: 어떤 경우에 증상이 완화되거나 더 심해지나요?
- [0406] - 계절에 따른 변화, 공기에 따른 변화, 찬 바람에 따른 변화, 운동 후 변화, 자세에 따른 변화(누워있을 때) 등
- [0407] Patient history: 해당되는 내용이 있으면 써 주시고 없으면 그대로 비워주세요.
- [0408] - 복용력: 최근 복용한 약이 있다면 적어주세요.
- [0409] - 과거력: 이전에 다른 질환을 앓으신 적이 있나요?
- [0410] (알레르기/호흡기 질환/부비동염/고혈압/심혈관 질환 등)
- [0411] - 사회력: 담배를 피우시나요? 하루에 어느 정도 피우시나요?
- [0412]
- [0413] (18) 콧물/코막힘
- [0414] Duration: 한번 증상이 나타나면 얼마나 오랫동안 지속되나요?(하루종일 /n시간 이상 지속 /주기적으로 등)
- [0415] Character: 증상의 양상이 어떻게 되나요?
- [0416] - 콧물의 특징은 어떤가요?
- [0417] (물처럼 맑은 콧물/노란 콧물/피가 섞인 콧물/악취가 나는 콧물 등)
- [0418] - 콧물이 양쪽에서 나오나요? 한쪽에서 나오나요?
- [0419] - 목 뒤로 콧물이 넘어가나요?
- [0420] Associated Sx: 함께 동반된 증상이 있나요?
- [0421] - 감기/몸살기운/기침/코 막힘/재채기/눈 가려움/코 가려움/가래/발열/두통 등
- [0422] Factor: 어떤 경우에 증상이 완화되거나 더 심해지나요?
- [0423] - 특정 계절에 따른 변화/특정 장소에 따른 변화/맵고 뜨거운 음식을 먹을 때/짙은 향수를 맡을 때/기온 차이가 갑자기 변할 때 등
- [0424] Patient history: 해당되는 내용이 있으면 써 주시고 없으면 그대로 비워주세요.
- [0425] - 약 복용력: 최근 복용한 약이 있다면 적어주세요.

- [0426] - 과거력: 이전에 다른 질환을 앓으신 적이 있나요?(천식/알레르기 질환 등)
- [0427] - 가족력: 가족 중 다음과 같은 질환을 진단받은 분이 있나요?
- [0428] (알레르기 비염/아토피 질환 등)
- [0429] - 외상력: 최근 얼굴이나 머리에 수술을 받았거나 다친 적이 있다면 알려주세요.
- [0430]
- [0431] (19) 객혈
- [0432] Character: 증상의 양상이 어떻게 되나요?
- [0433] - 출혈의 색이 어떻게 되나요? (선홍색/빨간색/적갈색)
- [0434] - 출혈의 특징은 어떤가요?
- [0435] (거품이 동반된 출혈/토사물이 섞인 출혈/냄새가 나는 출혈 등)
- [0436] - 한번에 얼마나 나오던가요?
- [0437] (가래에 묻어 나오는 정도/한 스푼/소주컵 한 잔/종이컵 반 컵 등)
- [0438] Associated Sx: 함께 동반된 증상이 있나요?
- [0439] - 감기/ 몸살기운/ 마른 기침/ 가래(색, 양)/ 속 울렁거림/ 복통/ 구토/ 발열/ 오한/ 잘 때 땀을 흘림/체중 감소/피가 나면 잘 못지 않음/가슴 통증/호흡곤란/숨이 참/쉽게 멎이 들 등
- [0440] Patient history: 해당되는 내용이 있으면 써 주시고 없으면 그대로 비워주세요.
- [0441] - 약복용력: 최근 복용한 약이 있다면 적어주세요.
- [0442] - 과거력: 이전에 다른 질환을 앓으신 적이 있나요?(암/결핵/폐렴/심장, 폐 관련 질환 등)
- [0443]
- [0444] (20) 호흡곤란
- [0445] Duration: 한번 증상이 나타나면 얼마나 오랫동안 지속되나요?(하루종일/ n시간 이상 지속/ 주기적으로 등)
- [0446] Character: 증상의 양상이 어떻게 되나요?
- [0447] - 쉼쉼거림/호흡이 빨라짐/호흡이 얕아짐/갑작스럽게 발생/가슴이 답답함/숨을 내쉬기 힘들/숨이 가빠짐/숨을 과도하게 쉼/가슴이 찢어지듯 아픔 등
- [0448] - 숨이 얼마나 차나요?
- [0449] Factor: 어떤 상황에서 숨이 더 차나요?
- [0450] - 운동할 때/누워있을 때/밤에/자세/장소에 따른 변화 등 ex)사람이 많은 곳에서 심해짐
- [0451] Associated Sx: 함께 동반된 증상이 있나요?
- [0452] - 기침/ 가래/ 콧물/ 감기/ 객혈/ 발열/ 어지러움/ 경련/ 두드러기/ 입술 부음/ 두근거림/ 가슴 통증/통증 뻔침/피로/체중 증가/다리가 붓거나 통증을 느낌/공포감/불안감 등
- [0453] Patient history: 해당되는 내용이 있으면 써 주시고 없으면 그대로 비워주세요.
- [0454] - 약 복용력: 최근 복용한 약이 있다면 적어주세요.
- [0455] - 과거력: 이전에 다른 질환을 앓으신 적이 있나요?
- [0456] (꽃가루 알레르기/당뇨/고혈압/심장, 기관지, 폐 관련 질환 등)
- [0457] - 가족력: 가족 중 다음과 같은 질환을 겪은 적이 있다면 알려주세요.
- [0458] (심장, 기관지, 폐 관련 질환 등 ex)아버지 폐렴 등)

- [0459] - 사회력: 담배를 피우시나요? 하루에 어느 정도 피우시나요?
- [0460]
- [0461] (21) 가슴통증
- [0462] Duration: 통증이 한번 발생하면 얼마나 오래 지속되나요?
- [0463] Character: 통증의 양상은 어떻게 되나요? 다음의 사항을 입력해주세요.
- [0464] - 통증의 정도를 1-10점으로 매긴다면 몇 점 정도인가요?
- [0465] - 쥐어짜듯이 아픔 / 묵직하게 누르듯이 아픔 / 통증이 팔 안쪽 또는 턱이나 목까지 뻗치는 느낌 / 통증이 등 뒤로 뻗치는 느낌 / 너무 아파서 식은땀이 남 / 찌르는 듯이 아픔 / 찢어지듯이 아픔 / 답답하고 조이는 느낌 / 쓰림 / 쑤심
- [0466] Associated Sx.: 아래와 같은 증상들 중 나타나는 증상을 모두 입력해주세요. 이외에 동반된 증상이 있다면 자유롭게 입력해주세요.
- [0467] - 발열 / 오한 / 숨이 참 / 오심(구역감) / 신물 / 호흡곤란 / 기침 / 가래 / 두통 / 근육통 / 체중감소 / 구토 / 피부발진 / 복통 / 속쓰림 / 스트레스
- [0468] Factor: 통증이 심해지는 요인이 있나요? 다음 중 해당되는 사항이 있다면 모두 입력해주세요. 이외에 요인이 있다면 자유롭게 입력해주세요.
- [0469] - 운동 후 / 찬바람 쐬 후 / 숨 들이마시면 / 기침 후 / 누우면 / 누르면 / 식전 / 식후
- [0470] Event/외상력: 다음 중 해당되는 사항이 있다면 입력해주세요.
- [0471] - 수술을 받음 / 오랜 기간 누워서 지냄 / 가슴을 다침
- [0472] Patient history: 해당되는 내용이 있으면 써 주시고 없으면 그대로 비워주세요.
- [0473] - 사회력: 식습관/운동습관/음주/흡연에 대해 간략히 서술해주세요(짜게 먹음, 운동안함, ~병/일, ~갑/일)
- [0474] - 가족력: 가족 중에 고혈압, 심근경색, 뇌출혈 등 심혈관질환을 앓고 계신 분이 있나요?
- [0475] - 과거력: 이전에 당뇨, 고혈압, 고지혈증 중에 진단받거나 치료받은 질환이 있나요? 이외에 투병 중이거나 진단받았던 질환이 있다면 자유롭게 입력해주세요.
- [0477] (22) 실신
- [0478] Duration: 얼마 동안 의식을 잃었나요?
- [0479] Character: 증상의 양상이 어떻게 되나요? 다음의 사항을 입력해주세요.
- [0480] - 어떤 상황에서 실신하였나요? 무엇을 하다가 실신했나요?
- [0481] - 목격자가 있다면, 다음의 사항 중 해당되는 사항을 입력해주세요.
- [0482] 팔, 다리를 뻗 / 눈이나 고개가 한 쪽으로 돌아감 / 의식이 없는 동안 대소변이 새어나옴
- [0483] Associated Sx.: 아래와 같은 증상들 중 나타나는 증상을 모두 입력해주세요. 이외에 동반된 증상이 있다면 자유롭게 입력해주세요.
- [0484] - 실신 전 눈 부심 / 어지럼증 / 구역 / 눈앞 캄캄 / 두통 / 가슴 통증 / 불안 / 호흡곤란
- [0485] - 실신 후 한 쪽 팔다리에 힘 빠짐 / 언어장애(말이 어둔해짐) / 구역 / 두통 / 가슴 통증
- [0486] - 불규칙한 심장 박동 / 두근거림 / 피로 / 체중감소 / 탈수
- [0487] Patient history: 해당되는 내용이 있으면 써 주시고 없으면 그대로 비워주세요.
- [0488] - 약 투약력: 복용하고 있는 약물이 있으면 입력해주세요.
- [0489] - 가족력: 가족 중에 실신을 했거나, 급사를 했거나, 심장 질환이 있는 분이 있으면 입력해주세요.
- [0490] - 과거력: 지금 앓고 있는 고혈압이나 심장 질환이 있다면 입력해주세요. 이외에 투병 중이거나 진단받았던 질환

환이 있다면 자유롭게 입력해주세요.

[0491]

[0492] (23) 두근거림

[0493] Duration: 얼마나 자주 발생하나요? 한 번 발생하면 얼마나 오래 지속되나요?

[0494] Character: 증상의 양상이 어떻게 되나요? 다음의 사항을 입력해주세요.

[0495] - 심장이 빨리 뛰는 느낌 / 심장이 불규칙하게 뛰는 느낌 / 심장이 쿵하고 내려앉는 느낌 / 규칙적 맥박 / 불규칙한 맥박 / 서서히 시작 / 서서히 사라짐 / 갑자기 시작 / 갑자기 사라짐

[0496] Associated Sx.: 아래와 같은 증상들 중 나타나는 증상을 모두 입력해주세요. 이외에 동반된 증상이 있다면 자유롭게 입력해주세요.

[0497] - 가슴 통증 / 숨이 참 / 실신 / 공복감 / 식은땀 / 손 떨림 / 두통 / 더위불내성 / 식욕 변화 / 체중 감소 / 어지러움 / 피로 / 발한 / 불안

[0498] Factor: 특별히 가슴이 두근거리는 때가 있나요? 두근거림을 유발하는 것이 있나요?

[0499] Patient history: 해당되는 내용이 있으면 써 주시고 없으면 그대로 비워주세요.

[0500] - 약 투약력: 드시고 있는 약이 있나요?

[0501] - 과거력: 다음 중 해당되는 사항이 있다면 모두 입력해주세요. 이외에 투병 중이거나 진단받았던 질환이 있다면 자유롭게 입력해주세요.

[0502] (심장 질환 / 부정맥 / 고혈압 / 당뇨 / 부정맥 / 정신과적 진료)

[0503] - 가족력: 가족 중에 심장 질환, 부정맥, 또는 급사한 분이 혹시 있나요?

[0504] - (여성) 최근 생리량 또는 생리 주기의 변화가 있나요?

[0505]

[0506] (24) 고혈압

[0507] Course: 어떤 상태에서 혈압을 측정하셨나요?

[0508] - 측정 전에 커피 / 측정 전 담배 / 의자에 앉아서 5분 이상 휴식 후 측정

[0509] Associated Sx.: 아래와 같은 증상들 중 나타나는 증상을 모두 입력해주세요. 이외에 동반된 증상이 있다면 자유롭게 입력해주세요.

[0510] - 몸 등이 부음 / 거품뇨 / 혈뇨 / 두근거림 / 식은땀 / 두통 / 더위불내성 / 식욕 감소 / 체중 증가 / 숨이 참 / 피로 / 소변량 감소 / 다한 / 안면홍조 / 코골이

[0511] Patient history: 해당되는 내용이 있으면 써 주시고 없으면 그대로 비워주세요.

[0512] - 약투약력: 복용하고 있는 약물이 있다면 입력해주세요.

[0513] - 사회력: 식습관/운동습관/음주/흡연에 대해 간략히 서술해주세요(짜게 먹음, 운동안함, ~병/일, ~갑/일)

[0514] - 가족력: 가족 중에 고혈압, 뇌졸중 등으로 치료 받으신 분이 있나요?(아버지: 고혈압)

[0515] - 과거력: 당뇨병, 고지혈증 등 다른 성인병을 앓고 있는 것이 있나요? 이외에 투병 중이거나 진단받았던 질환이 있다면 자유롭게 입력해주세요.

[0516]

[0517] (25) 유방통/유방덩이

[0518] Character:

[0519] - 크기: 유방 덩이가 느껴진다면 크기는 얼마정도 되나요?

[0520] - 통증: 얼마나 아픈가요?(0~10)

- [0521] - 강도: 딱딱한가요? 말랑한가요?
- [0522] Associated Sx.: 해당되는 증상이 있으면 입력(체크)해주세요.
- [0523] - 열/오한/뺨강게 달아 오름/열감/분비물/스트레스/피로/체중감소/유방통증
- [0524] Patient history: 해당되는 내용이 있으면 써 주시고 없으면 그대로 비워주세요.
- [0525] - 약복용력: 약을 복용하고 있다면 어떤 약인지 입력해주세요(피임약)
- [0526] 호르몬 치료를 받고 있다면 체크해주세요(호르몬대체요법).
- [0527] - 가족력: 가족 중에 유방암이나 난소암을 앓으신 분이 있나요?(어머니, 유방암)
- [0528] - 여성력: 마지막 생리를 시작한 날이 언제쯤이었나요?
- [0529] 초경(폐경)은 언제 하셨나요?
- [0530] (수유경험이 있다면) 언제부터 모유수유를 시작하였나요?
- [0531] (n개월 전부터 모유수유중/n개월 전까지 모유수유 했었음)
- [0532]
- [0533] (26) 질 분비물
- [0534] Character:
- [0535] - 색: 색상이 어떤가요?(붉은색/노란색/흰색 등)
- [0536] - 냄새: 냄새가 어떤가요?(없음/고약함/비릿내 등)
- [0537] Associated Sx.:
- [0538] - 외음부 가려움/질 가려움/걸을 때 통증/소변볼 때 통증/식은땀/화끈거림/아랫배 통증/피로/열/ 오한/따가움
- [0539] Factor:
- [0540] - (필요한 질문이니 당황하지 마시고)성관계를 어떤 사람과 하시나요? (평소 남자친구와 성관계/여러 사람과 성관계)
- [0541] - 최근에 성관계를 하셨다면 체크해주세요(최근 성관계), 성관계 시 통증이 생기면 체크해주세요(성관계 시 통증 생김)
- [0542] Patient history: 해당되는 내용이 있으면 써 주시고 없으면 그대로 비워주세요.
- [0543] 여성력: 해당되는 사항을 입력해주세요.
- [0544] - 결혼여부(미혼/기혼), 평소의 생리주기(규칙적/불규칙적), 폐경이시면 체크해주세요(폐경)
- [0545]
- [0546] (27) 월경이상/월경통
- [0547] Onset: 증상이 언제부터 시작되었나요?
- [0548] Character:
- [0549] - 통증: 얼마나 아픈가요?(0~10)
- [0550] - 양: 하루에 패드 몇 개 정도 사용하나요?(팬티라이너 n개, 패드 n개 등)
- [0551] Associated Sx.: 해당되는 증상이 있으면 입력(체크)해주세요.
- [0552] - 두통/체중증가/(의도하지 않은)체중감소/피로/더위/여드름이 많아짐/스트레스/피부 건조/식욕 떨어짐/어렸을 때부터 성장이 느렸음/당이/약취/젖이 흐르거나 가슴에 멍힘/시야가 흐림/
- [0553] Factor:

- [0554] - 성관계 시 통증이 생기면 체크해주세요(성관계 시 통증 생김)
- [0555] - 임신 가능성이 있으면 체크해주세요(모름/임신가능성/현재 임신 n주)
- [0556] - 평소 멍이 자주 들면 체크해주세요(멍이 잘 들)
- [0557] Patient history: 해당되는 내용이 있으면 써 주시고 없으면 그대로 비워주세요.
- [0558] - 약 복용력: 복용하고 있는 약물이 있다면 입력해주세요.
- [0559] - 가족력: 가족 중에 갑상선 관련 질환이나 여성질환을 앓았던 분이 있나요?(어머니, 유방암)
- [0560] - 외상력(수술력): 최근에 자궁이나 질을 다친 적이 있으면 체크해주세요(외상)
- [0561] - 여성력: 해당되는 사항을 입력해주세요.
- [0562] 결혼여부(미혼/기혼)
- [0563] 평소의 생리주기가 어떤가요?(규칙적/불규칙적/초경 안옴)
- [0564] 마지막 생리가 언제인가요?
- [0565] 결과적으로, 위 문진항목 목록을 통해 기존 20여개에 가까웠던 질문들을 주요 증상별 평균 14개의 질문으로 정리하며 핵심적인 정보를 수집할 수 있고, 유저의 편의성 또한 증가할 수 있다.
- [0567] 도 5는 본 발명의 일 실시예에 따라 증상과 질환을 매칭하는 기본 모델을 설명하기 위한 도면이다.
- [0568] 도 5를 참조하면, 사용된 feature는 Chief complaint, Onset, Location이다. 주요 증상 예측은 문진 데이터의 level2에 해당되는 내용이다. 황달, 토혈, 호흡곤란 등 총 27가지 주요 증상이 정리될 수 있다.
- [0569] - TF-IDF 단어 벡터화
- [0570] 텍스트 데이터를 머신러닝에 적용하기 위해서는 단어 벡터화 과정이 필요하다. 주요 증상 예측 모델에서는 TF-IDF (Term Frequency - Inverse Document Frequency) 기반 단어 벡터화를 진행하였다. TF-IDF 기반 단어 벡터화란 Count 기반 단어 벡터화를 보완한 방법으로 다른 문서에서의 단어 빈도를 고려하여 중요도를 표현한다. TF는 특정 단어가 문서 내에서 자주 등장한 값이고 IDF란 역문서 빈도로 단어가 전체 문서 내에서 얼마나 자주 나타나는지 보여주는 값이다.
- [0571] - 베이스라인 비교
- [0572] 단어 벡터화를 적용한 뒤 아래 그림과 같이 여러 머신러닝 모델을 적용하여 베이스라인을 구축하였다. 테스트 데이터에 대한 모델들은 대체로 0.9이상의 Accuracy와 F1-Score를 보였지만, Random Forest Classifier가 다른 모델들보다 약 0.01~0.04의 성능 차이를 보였다.

표 2

Model	param	Test Acc	Test F1-Score
Multinomial Naive Bayes	alpha = 0.001 외 default	0.92	0.92
Support Vector Machine	default	0.95	0.94
Random Forest Classifier	random_state=0 외 default	0.96	0.96
Decision Tree Classifier	random_state=0 외 default	0.94	0.95

- [0573]
- [0574] - 최종 모델 선정
- [0575] 주요 증상 예측 모델의 최종 모델은 위 베이스라인 표에서 성능이 가장 좋은 Random Forest Classifier를 고도화하였다. GridSearchCV를 이용하여 하이퍼 파라미터를 구하였고, Stratified Kfold를 이용하여 각 fold에서 class가 편중되지 않게 조절하였다. 문진 데이터의 level2(주요 증상)에는 라벨들이 고르게 분포되어 있지 않다. 30개로 가장 적은 주요 증상은 ‘변비’와 ‘질 분비물’이고 가장 많은 2개의 값은 ‘황달’이 495개, ‘토혈’이 400개이다. 450개가 넘는 라벨 데이터 분포를 보이는 불균형 데이터이기에 class weight를 이용하여 불균형 라벨 비율을 조정할 수 있다. 그 결과 Random Forest Classifier의 최적 하이퍼 파라미터는 max_depth:

50, min_samples_split: 0.01, n_estimators: 50이고 테스트 데이터로 성능 평가 해본 결과 Accuracy는 0.97, F1-Score는 0.97로 향상될 수 있다.

[0577] 아래에서는 증상과 질환을 예측하는 모델에 관하여 설명하도록 한다.

[0578] 사용된 증상 feature는 Chief complaint, Age, Sex, Onset, Location, Duration, Course, Experience, Character, Associated Sx., Factor, Event, 약물 투약력, 사회력, 가족력, 외상력, 과거력, 여성력, Obesity가 한 문장으로 포함된 'All' 칼럼이다. 질환은 문진 데이터의 level5/diagnosis5에 해당되는 질환들로, 4.데이터 전처리 - 2)질환 관련 작업을 반영한 225가지 질환들이다.

[0579] - 모델1, 모델2 분리

[0580] 해당 데이터셋에는 유방암, 자궁내막증 등 여성에게만 해당되는 질병들이 포함되어 있다. 보다 정확한 모델 성능을 보이기 위해, 성별이 '남자'인 환자에게 적용할 남녀공통 질병만 예측하는 모델1, 성별이 '여자'인 환자에게 적용할 모델2(남녀 공통 질병과 여성력 관련 질병 예측)로 분리하였다. 모델1은 193개 질병, 모델2는 여성력 관련 질병을 포함한 225개의 질병을 예측할 수 있다.

[0581] - Tokenizer 비교

[0582] 데이터 전처리 후, 모델 1과 2를 위한 데이터 분리, 훈련 데이터와 테스트 데이터셋 분리를 거쳤다. 그 후 텍스트를 여러 토큰으로 나누는 토크나이징 작업을 하였다. 한국어 토크나이징에는 Whitespace, Konly Tokenizer, Mecab, Kobert 등이 있다. 적용 가능한 의료 도메인 특화 언어 모델이 없기에 데이터에 기존에 알려진 여러 토크나이저를 적용해서 비교해보았다. 적용해본 결과 토크나이저별 큰 성능 차이는 없지만, LMKor 토크나이저가 약 0.01 정도의 더 우수한 성능을 보이므로 LMKor: funnel-kor-base 토크나이저를 선택하였다. 해당 토크나이저는 국내 주요 커머스 리뷰 1억개와 블로그 형 웹사이트 2000만 개, 모두의 말뭉치, 위키피디아와 나무위키에서 약 70GB 데이터셋 학습으로 이루어져있다. 모델 1의 토크나이저 단어 수는 1600, 모델 2의 단어 수는 1800개이다.

[0583] - 워드 임베딩

[0584] 정수 인코딩 및 패딩 과정을 거친 뒤, Keras 워드 임베딩 과정을 진행하였다. 고차원과 다르게 저차원 밀집 벡터는 특징 되는 표현값들을 더 낮은 차원으로 효율적인 표현이 가능하다. 초기에는 임베딩 벡터는 랜덤값으로 설정되지만, 모델이 학습을 하면서 예측 결과에 따라 가중치가 변할 수 있다.

[0585] - 베이스라인 비교

[0586] 딥러닝 모델을 적용하여 아래 그림처럼 증상-질환 예측 모델 베이스라인을 구축하였다. 자연어 처리에서 많이 쓰는 LSTM, CNN, RNN으로 간단한 모델들을 만들어 보았는데, 이 중 데이터에서 feature를 추출하여 의미를 파악하는 CNN이 모델2 기준으로, 0.74 Accuracy, 0.81 F1-Score로 가장 좋은 성능을 보였다.

표 3

Model	param	Test Acc	Test F1-Score
Multinomial Naive Bayes	alpha = 0.001 외 default	0.92	0.92
Support Vector Machine	default	0.95	0.94
Random Forest Classifier	random_state=0 외 default	0.96	0.96
Decision Tree Classifier	random_state=0 외 default	0.94	0.95

[0587]

[0588] 추가적으로 Random Forest Classifier와 Multinomial Naive Bayes 등의 머신러닝 모델을 도입했지만 약 0.6 Accuracy라는 좋지 못한 성능을 보였다. 또한 사전에 학습된 Word2vec과 Fasttext를 이용한 word embedding의 가중치를 임베딩 레이어에 추가하는 시도를 하였으나, 구별되는 증상 용어를 비슷하게 보아서 성능 개선에 미비하였다.

[0590] 도 6은 본 발명의 일 실시예에 따라 증상과 질환을 매칭하는 최종 모델을 설명하기 위한 도면이다.

[0591] 도 5에 따른 모델들을 반영하여, 도 6에 도시된 바와 같이 최종 모델을 고도화하였다. 모델1, 모델2의 전체적인 구조는 같지만 가중치나 토크나이저에 따른 파라미터의 수 등이 다르다. 라벨은 총 225개 질병(모델2 기준)이지

만, 라벨의 분포가 고르지 않다. 가장 케이스가 많은 질병은 ‘식도 정맥류 파열’ (100개)이고, ‘소화성 궤양’ (95개), ‘역류성 식도염’ (64개) 순으로 많다. 5개의 최소 케이스인 질환들은 ‘미주신경성 실신’, ‘요붕증’, ‘크론병’ 등이다. 라벨의 불균형 분포를 해소하고자 모델에 라벨 개수를 고려한 class weight를 추가할 수 있다.

[0592] LSTM을 추가하여 문맥을 반영하고, Conv1D에서 3개의 다른 필터 (필터 사이즈는 각각 5, 7, 9)를 적용한 뒤, 한 레이어로 합치는 작업을 추가하였다. 또한 오버피팅을 방지하기 위한 여러 Dropout layer와 Dense 레이어를 거친 후, Softmax 활성화 함수로 각 질병별 발생 확률을 계산한다. 추가적으로 relu 활성화 함수와 categorical_crossentropy 손실 함수, adam 옵티마이저를 적용하였다. 배치 사이즈는 216(모델1), 512(모델2)이며 validation set은 train set의 20%에 해당된다.

[0593] 유저에게 증상을 바탕으로 진단 확률이 가장 높은 질환 3개를 보여준다. 때문에 기존 성능지표 외로도 ‘Top3 Hitrate’ 라는 성능 지표를 추가하였다. Top3 Hitrate란, 라벨링된 질병이 Top3 예측 질병에 포함되어 있으면 1을 추가하고 포함되지 않으면 0을 추가하여 평균을 구하는 평가 방법이다. 모델 1은 m_top_k_hitrate, 모델2는 w_top_k_hitrate 함수로 구현하였다. 이렇게 만든 최종 모델의 성능은 아래 표와 같다. 두 모델 모두 테스트 데이터셋에서 0.8이상의 accuracy와 0.9이상의 top3 hitrate라는 좋은 성능을 보였다.

표 4

모델	Test Acc	Test F1-Score	Top3 Hitrate
모델1 (남녀 공통 질병)	0.82	0.88	0.91
모델2 (남녀 공통 + 여성력 관련 질병)	0.80	0.87	0.91

[0594]

[0595] 도 7은 본 발명의 일 실시예에 따른 챗봇의 동작을 나타내는 흐름도이고, 도 8은 본 발명의 일 실시예에 따른 챗봇의 실시예를 나타내는 도면이다.

[0596] 도 7 및 도 8을 참조하면, 본 발명의 일 실시예에 따른 챗봇은 아래와 같이 동작할 수 있다.

[0597] 1. 채팅 메시지 내용 수집

[0598] (1-1) 사용자로부터 질문에 따라 자연어형/선택지형/단답형으로 입력을 받음

[0599] 2. 자연어가 있다면, 수집한 채팅 대화내용에서 핵심 키워드를 추출

[0600] (2-1) 제가 오늘 아까부터 배가 아파서요 밥을 못먹었어요 → 오늘, 배가, 아프다 등

[0601] 3. 키워드의 토큰화(Tokenize) 및 표준어화 진행

[0602] (3-1) 배가 아픈데/아파서/아프다/완전아프네/아프네 등 →아프다

[0603] 4. 추출한 키워드 DB에 저장

[0604] (4-1) 각 증상별로 데이터의 집합을 만들기 - 증상명, 시간, 강도, 빈도, 약 복용 여부 등

[0605] (4-2) 증상의 갯수 = 데이터 집합의 갯수

[0606] 5. 사전에 정의한 형태로 DB 내용 저장

[0607] 6. 저장한 내용을 증상 - 질환 매칭 모델에 적용

[0608] 7. 가장 높은 확률로 나타나는 진단명 세 가지를 사용자에게 알림

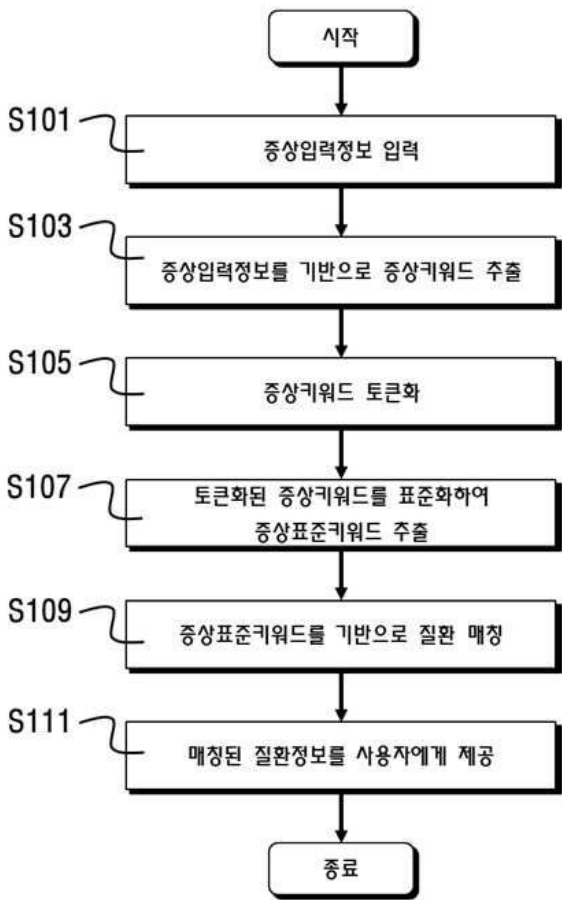
[0610] 또한, 본 명세서와 도면에 개시된 본 발명의 실시예들은 본 발명의 기술 내용을 쉽게 설명하고 본 발명의 이해를 돕기 위해 특정 예를 제시한 것일 뿐이며, 본 발명의 범위를 한정하고자 하는 것은 아니다. 즉 본 발명의 기술적 사상에 바탕을 둔 다른 변형예들이 실시 가능하다는 것은 본 발명의 속하는 기술 분야에서 통상의 지식을 가진 자에게 자명한 것이다. 또한 상기 각각의 실시예는 필요에 따라 서로 조합되어 운용할 수 있다.

[0611] 또한, 본 발명의 일 실시예에 따른 증강현실 기반의 마케팅 시스템의 동작 방법은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능 매체에 기록될 수 있다.

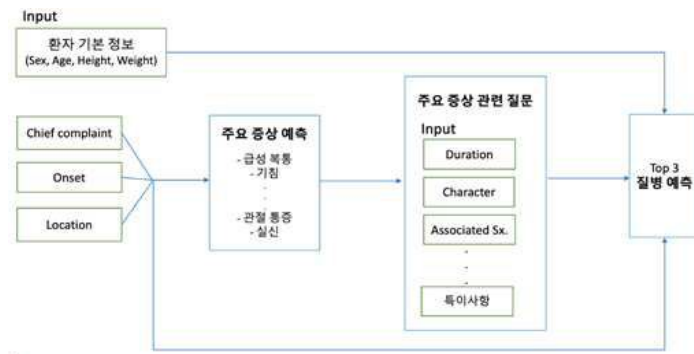
- [0612] 도 9는 본 발명의 일실시예에 따른 컴퓨터 시스템을 나타낸 도면이다.
- [0613] 도 9에 도시된 바와 같이 본 발명의 다양한 실시예들은 특정 관점에서 컴퓨터 리드 가능 기록 매체(computer readable recording medium)에서 컴퓨터 리드 가능 코드(computer readable code)로서 구현될 수 있다. 컴퓨터 리드 가능 기록 매체는 컴퓨터 시스템에 의해 리드될 수 있는 데이터를 저장할 수 있는 임의의 데이터 저장 디바이스이다. 컴퓨터 리드 가능 기록 매체의 예들은 읽기 전용 메모리(read only memory: ROM)와, 랜덤-접속 메모리(random access memory: RAM)와, 콤팩트 디스크-리드 온니 메모리(compact disk-read only memory: CD-ROM)들과, 마그네틱 테이프(magnetic tape)들과, 플로피 디스크(floppy disk)들과, 광 데이터 저장 디바이스들, 및 캐리어 웨이브(carrier wave)들(인터넷을 통한 데이터 송신 등)을 포함할 수 있다. 컴퓨터 리드 가능 기록 매체는 또한 네트워크 연결된 컴퓨터 시스템들을 통해 분산될 수 있고, 따라서 컴퓨터 리드 가능 코드는 분산 방식으로 저장 및 실행된다. 또한, 본 발명의 다양한 실시예들을 성취하기 위한 기능적 프로그램들, 코드, 및 코드 세그먼트(segment)들은 본 발명이 적용되는 분야에서 숙련된 프로그래머들에 의해 쉽게 해석될 수 있다.
- [0614] 또한 본 발명의 다양한 실시예들에 따른 장치 및 방법은 하드웨어, 소프트웨어 또는 하드웨어 및 소프트웨어의 조합의 형태로 실현 가능하다는 것을 알 수 있을 것이다. 이러한 소프트웨어는 예를 들어, 삭제 가능 또는 재기록 가능 여부와 상관없이, ROM 등의 저장 장치와 같은 휘발성 또는 비휘발성 저장 장치, 또는 예를 들어, RAM, 메모리 칩, 장치 또는 집적 회로와 같은 메모리, 또는 예를 들어 콤팩트 디스크(compact disk: CD), DVD, 자기 디스크 또는 자기 테이프 등과 같은 광학 또는 자기적으로 기록 가능함과 동시에 기계(예를 들어, 컴퓨터)로 읽을 수 있는 저장 매체에 저장될 수 있다. 본 발명의 다양한 실시예들에 따른 방법은 제어부 및 메모리를 포함하는 컴퓨터 또는 휴대 단말에 의해 구현될 수 있고, 이러한 메모리는 본 발명의 실시예들을 구현하는 명령들을 포함하는 프로그램 또는 프로그램들을 저장하기에 적합한 기계로 읽을 수 있는 저장 매체의 한 예임을 알 수 있을 것이다.
- [0615] 따라서, 본 발명은 본 명세서의 청구항에 기재된 장치 또는 방법을 구현하기 위한 코드를 포함하는 프로그램 및 이러한 프로그램을 저장하는 기계(컴퓨터 등)로 읽을 수 있는 저장 매체를 포함한다. 또한, 이러한 프로그램은 유선 또는 무선 연결을 통해 전달되는 통신 신호와 같은 임의의 매체를 통해 전자적으로 이송될 수 있고, 본 발명은 이와 균등한 것을 적절하게 포함한다.
- [0616] 본 명세서와 도면에 개시된 본 발명의 실시 예들은 본 발명의 기술 내용을 쉽게 설명하고, 본 발명의 이해를 돕기 위해 특정 예를 제시한 것일 뿐이며, 본 발명의 범위를 한정하고자 하는 것은 아니다. 또한 앞서 설명된 본 발명에 따른 실시예들은 예시적인 것에 불과하며, 당해 분야에서 통상의 지식을 가진 자라면 이로부터 다양한 변형 및 균등한 범위의 실시예가 가능하다는 점을 이해할 것이다. 따라서, 본 발명의 진정한 기술적 보호 범위는 다음의 특허청구범위에 의해서 정해져야 할 것이다.

도면

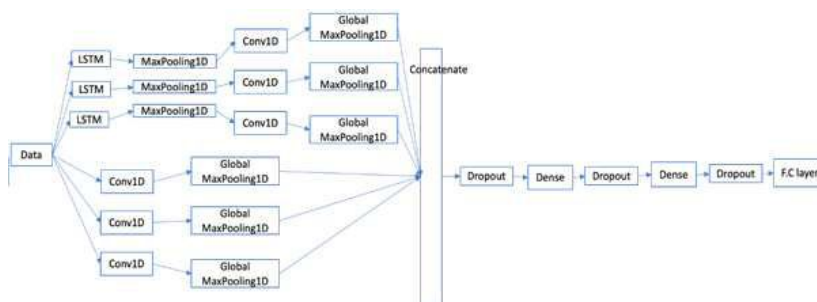
도면1



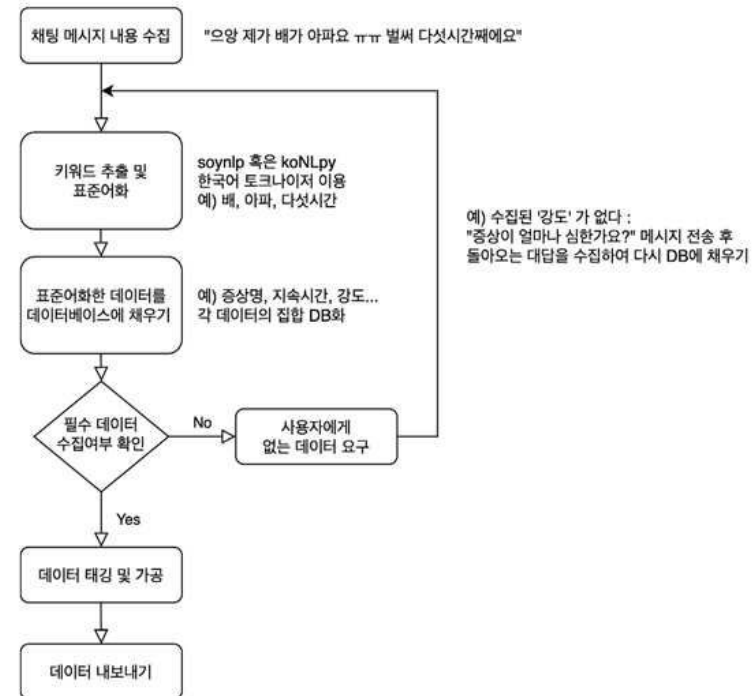
도면5



도면6



도면7



도면8



도면9

